Queueing theory Questions

Navid Najafi
mobin jelodar

December 2024

Questions

> ### Theoretical Question 1: Why does the queue grow infinitely?

If the arrival rate of individuals (λ) is higher than the service rate (μ), the queue size will grow indefinitely over time because more individuals are entering the system than can be served.

> ### Theoretical Question 2: What does $C \div B$ represent?

$C \div B$ represents the rate at which individuals leave the queue. This is equivalent to the service rate (μ), which indicates the number of individuals served per unit of time when the system is actively providing service.

> ### Theoretical Question 3: What is the logical interpretation of λ?

Lambda (λ) represents the arrival rate of individuals into the system. It can also be interpreted as the reciprocal of the expected time it takes for one individual to enter the system.

In a steady-state condition, the total number of individuals in the system (assuming there is only one server) is logically equal to the total time the previous individual spent waiting or receiving service, divided by the expected time it takes for one individual to enter the system. This relationship aligns with Little's Law, which states:

$$\mathbb{E}[N] = \lambda . \mathbb{E}[T]$$
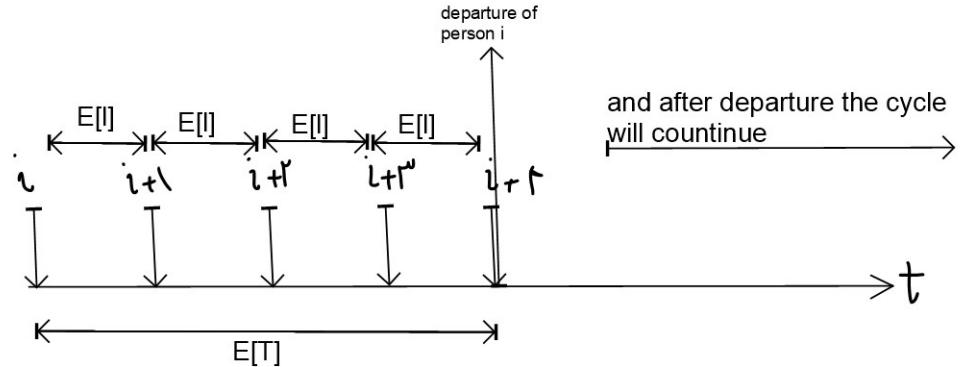
first we define
T := variable of waiting and getting service in the system
l := variable of the time that it takes for a person to arrive
hense:

$$\lambda = \mathbb{E}[l]$$

$$\mathbb{E}[N] = \frac{\mathbb{E}[T]}{\mathbb{E}[l]} = \lambda.\mathbb{E}[T]$$

in steady state situation we have diagram of arrivals like this:

**Theoretical Question 4: prove that: $\rho = \frac{\lambda}{\mu}$?**
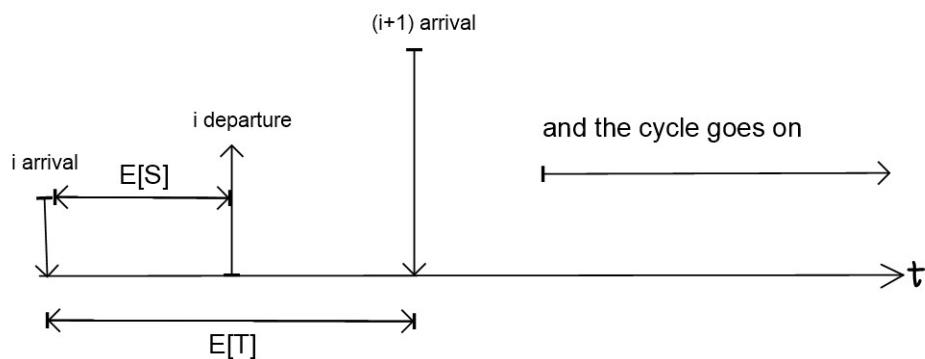
At the outset, we emphasize that if λ is greater than μ, the system will not reach a steady state. As a result, the system's efficiency coefficient will equal 1, because the queue length will continuously increase. This will progress to the point where the system is constantly busy, and the queue length will approach infinity.

if $\mu \leq \lambda$:

$$\lim_{t \to \infty} N(t) = \infty, \rho = 1$$
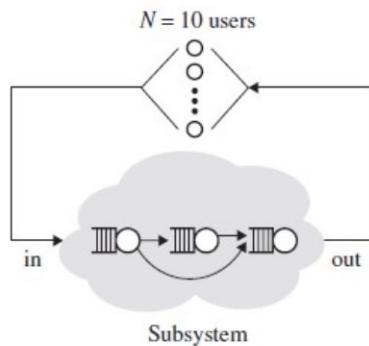
otherwise in steady state:

we expext a cycle of work and rest for system that N(t) constantly toggle from 0 to 1 and 1 to 0:

2

$$\rho = \frac{\mathbb{E}[S]}{\mathbb{E}[T]} = \frac{\frac{1}{\mu}}{\frac{1}{\lambda}} = \frac{\lambda}{\mu}$$

Theoretical Question 5: using the law in the equation

$N = 10$ users

in

out

Subsystem
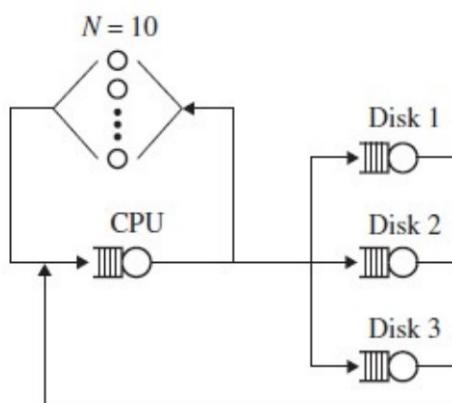
$$\mathbb{E}[T] = \mathbb{E}[R] + \mathbb{E}[Z]$$

$$N = X.\mathbb{E}[T]$$

$$X = \frac{N}{\mathbb{E}[R] + \mathbb{E}[Z]} = \frac{10}{5 + 15} = 0.5$$

Theoretical Question 6: consider the system below:

$N = 10$

Disk 1

CPU

Disk 2

Disk 3

- operational capacity Disk3:

$$\rho = \lambda E[S]$$

$$\rho = 40 \times 0.0225 = 0.9$$

- Waiting time in Disk3:

$$E[T_{Q_3}] = E[T_3] - E[S_3]$$

$$R[T_3] = \frac{E[N]}{8} = \frac{4}{40}$$

$$E[T_{Q_3}] = 0.1 - 0.0225 = 0.0775$$

- jobs in Disk 3:

$$N_{Q_3} = X T_{Q_3}$$

$$E[N_{Q_3}] = 40 \times 0.0775 = 3.1$$

- operational capacity of system:
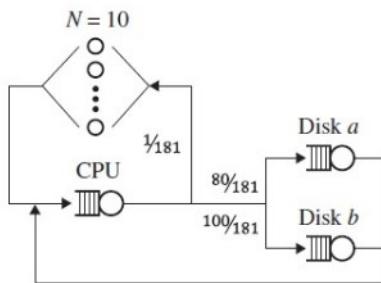
$$E[Z] = \frac{7.5}{10} = 0.75$$

- equation of system operational capacity: you can calculate system operational capacity by this equation:

$$\frac{N}{E[T]}$$

**Theoretical Question 7: consider the Que system**

:
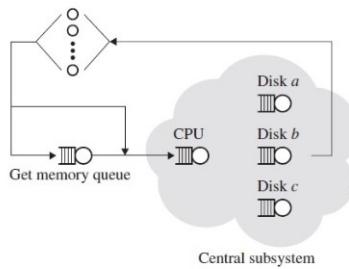
$$\frac{100}{181} = \frac{180}{181} \times E[V_b] \rightarrow E[V_b] = \frac{5}{9}$$

$$\frac{80}{181} = \frac{180}{181} \times E[V_a] \rightarrow E[V_a] = \frac{4}{9}$$

$$\frac{1}{181} = \frac{180}{181} \times E[V_{cpu}] \rightarrow E[V_{cpu}] = \frac{1}{180}$$

$$E[T] = \frac{E[N_{central}]}{X} = \frac{N - E[N_{getmemory}]}{X} = \frac{11.45}{0.45} = 25.44$$

$$D_i = S_i \times V_i$$

we get Expectation of both sides:

$$E[D_i] = E[S_i V_i]$$

we know that S and V are independent (we can prove it but we accept it for now). Thus:

$$E[D_i] = E[S_i] \times E[V_i]$$

bottleneck law express that the most use of a source at a Que system equals to rate of arrivals $\times$ Expected serve Time, which rate fo arrivals represent by X and Expected serve Time represent by $E[D_i]$. intuition of this law is like at a Que system, bottleneck is the step that experience the most Time of serving or rate of Arrivals. this step shows that how much throughput capacity or rate of is in the system, thus the throughput capacity of system limited to the bottleneck capacity.

for example, imagin a producing line in a factory with few steps of producing. if one of the steps be very slower than others. producing line can not be faster

than the speed of that step, even if other steps be very fast, thus that exact step is our bottleneck and determine the whole throughput capacity of system.

for proving the bottleneck law, we use definition of "intensity of usage" ρ:

$$E[D_i].X = \rho_i$$

we want to define "intensity of usage" ρ:

$$\text{intensity of usage} = \rho_i = E[S_i].X_i$$

which X_i is the rate of arrivals, and $E[S_i]$ is the Expectation of serving time.

$$X_i = X.V_i$$

$$\text{intensity of usage} = \rho_i = E[S_i].X.V_i$$

$$D_i = V_i.S_i$$

we get expectation of both sides and S_i, V_i are independent. so:

$$E[D_i] = E[V_i.S_i] = E[V_i].E[S_i] = V_i.E[S_i]$$

after substituting this equation, we have:

$$\rho_i = X.E[D_i]$$

Theoretical Question 11: = proof of two inequalities:

:

---

to proving first inequality we have:
first limitation:

$$X \leq \frac{N}{D + E[Z]}$$

this limitation comes from little's law for closed systems.
little's law for closed systems can be expressed as:

$$(E[Z] + D).X = N$$

Thus:

$$X = \frac{N}{E[Z] + D}$$

and because X can not be higher than $\frac{N}{E[Z]+D}$, we say:

$$X \leq \frac{N}{E[Z] + D}$$

6

second limitation:

$$X \leq \frac{1}{D_{max}}$$

this limitation comes from bottleneck law:

$$X \leq \frac{1}{D_{max}}$$

in this inequality D_{max} equals to longest serving time in all sources.

this shows bottleneck of system and process speed rate of system can not be beyond bottleneck.

we proved 2 limitation, then it is obvious that:

$$X \leq min(\frac{N}{D + E[Z]}, \frac{1}{D_{max}})$$

now we want to prove the second inequality:

first limitation:

$$E[R] \geq D$$

because total response time includes both serving time and waiting time, as D represents only the serving time, response time can not be less than D.

Thus:

$$E[R] \geq D$$

second limitation:

$$E[R] \geq N.D_{max} - E[Z]$$

little's law

$$(E[Z] + E[R]).X = N$$

based on bottleneck law:

$$X \leq \frac{1}{D_{max}}$$

if we substitute in little's law, we have

$$N \leq \frac{E[R] + E[Z]}{D_{max}}$$

thus:
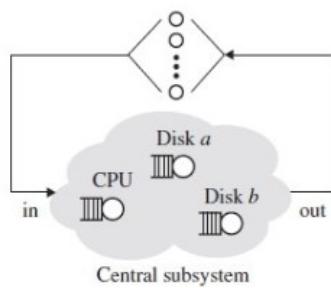
$$E[R] \geq N.D_{max} - E[Z]$$

regrading $E[R]$ should apply in both limitation, then we have:

$$E[R] \geq max(D, N.D_{max} - E[Z])$$

Theoretical Question 12: consider the system below

:

Central subsystem
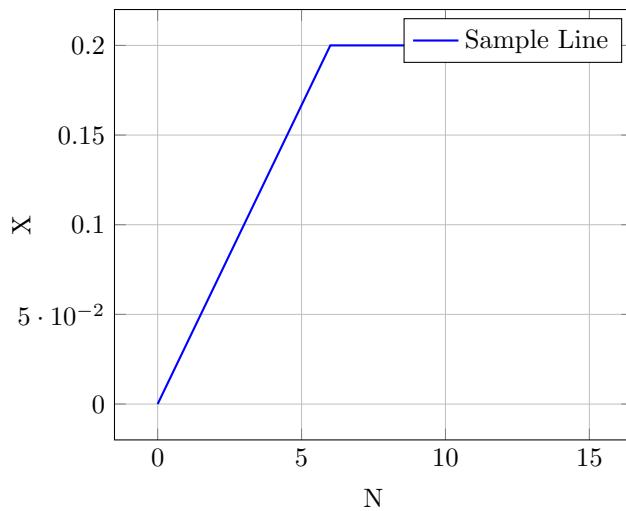
$$E[D_{cpu}] = 5, E[D_{disk,A}] = 4, E[D_{disk,b}] = 3$$
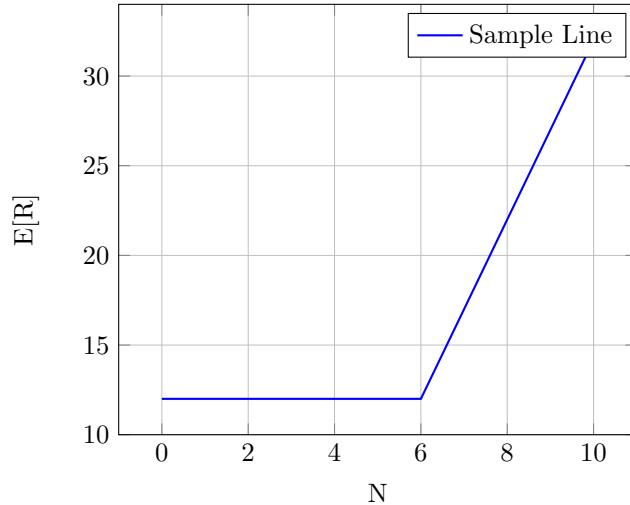
$$D_{total} = E[D_{cpu}] + E[D_{disk,A}] + E[D_{disk,b}] = 12$$

$$\frac{N}{D_{total} + E[Z]} = \frac{N}{C}$$

$$X \leq Min(\frac{N}{C}, 0.2)$$

now we have to plot them:

Theoretical Question 13: comparison of two Que systems

:

we know that $D = \frac{1}{\mu}$ thus:

$$D_i = \frac{1}{\mu_i} \rightarrow D = 2 \times 3 = 6$$

$$D_{server1} = D_{server2} = 3, D_{max} = 3,$$

$$D_{server1} = 2, D_{server2} = 3, D_{max} = 3$$

does throughput of system increases?
yes, because we know:

$$X = \frac{N}{D + E[Z]}$$

because D decreases (from 6 to 5), then the upper bond of throughput increases.

Theoretical Question 14: consider the Que system

:

$$N = 20, E[Z] = 5$$

for the first system:

$$D_{cpu} = 4.6, D_{disk} = 4 \rightarrow D_{total} = 8.6$$

for the second system:

$$D_{cpu} = 4.9, D_{disk} = 1.9 \rightarrow D_{total} = 6.8$$

we knew from the past:

$$X \leq Min(\frac{N}{D + E[Z]}, \frac{1}{D_{max}})$$

for the first system:

$$X \leq min(\frac{20}{8.6 + 5}, \frac{1}{4.6}) = 0.217$$

for the second system:

$$X \leq Min(\frac{20}{5.8 + 5}, \frac{1}{4.9}) = 0.204$$

thus, first system has higher upper-bond for throughput, then the first system has higher throughput.

Theoretical Question 15:

we know these equations, we define a U_i as load on our source!

$$U_i = \frac{B_i}{C_i}$$

$$X = \frac{1}{max(U_i)}$$

first idea, make CPU speed rate, two times faster:

$$C_{CPU_{new}} = 2 \times C_{CPU} = 2 \times 200 = 400 jobs$$

$$U_{CPU_{new}} = \frac{B_{CPU}}{C_{CPU_{new}}} = \frac{400}{400} = 1$$

$$U_{slowDisk} = \frac{100}{2000} = 0.05$$

$$U_{fastDisk} = \frac{600}{2000} = 0.3$$

new system throughput:

$$X = \frac{1}{max(U_i)} = \frac{1}{1} = 1\frac{jobs}{second}$$

second idea, Let's balance the faster and slower disks:

$$B_{slowDisk_{new}} = B_{fastDisk_{new}} = \frac{B_{slowDisk} + B_{fastDisk}}{2} = \frac{100 + 600}{2} = 350$$

$$U_{slowDisk_{new}} = \frac{B_{slowDisk_{new}}}{C_{slowDisk}} = \frac{350}{2000} = 0.175$$

$$U_{fastDisk_{new}} = \frac{B_{fastDisk_{new}}}{C_{fastDisk}} = \frac{350}{2000} = 0.175$$

$$U_{CPU_{new}} = 2$$

$$X = \frac{1}{max(U_i)} = \frac{1}{2} = 0.5\frac{jobs}{second}$$

third idea, add a new fast Disk:

$$C_{fastDisk_{new}} = 2 \times C_{fastDisk} = 2 \times 20000 = 40000 jobs$$

$$U_{fastDisk_{new}} = \frac{B_{fastDisk}}{C_{fastDisk_{new}}} = \frac{600}{40000} = 0.015$$

$$U_{CPU} = 2, U_{slowDisk} = 0.05$$

$$X = \frac{1}{max(U_i)} = \frac{1}{2} = 0.5\frac{jobs}{second}$$

fourth idea, combining all:

$$C_{CPU_{new}} = 2 \times C_{CPU} = 400 jobs$$

$$C_{fastDisk_{new}} = 2 \times C_{fastDisk} = 2 \times 20000 = 40000 jobs$$

$$B_{slowDisk_{new}} = B_{fastDisk_{new}} = \frac{B_{slowDisk} + B_{fastDisk}}{2} = \frac{100 + 600}{2}$$

$$U_{CPU_{new}} = \frac{B_{CPU}}{C_{CPU_{new}}} = 1$$

$$U_{slowDisk_{new}} = \frac{B_{slowDisk_{new}}}{C_{slowDisk}} = \frac{350}{2000} = 0.175$$

$$U_{fastDisk_{new}} = \frac{B_{fastDisk_{new}}}{C_{fastDisk}} = \frac{350}{40000} = 0.00875$$

$$X = \frac{1}{max(U_i)} = \frac{1}{1} = 1\frac{jobs}{second}$$

Theoretical Question 16: calculation of E[T], Var(T)

$$E[T] = \int_0^\infty x\lambda e^{-\lambda x}dx = \lambda \int_0^\infty xe^{-\lambda x}dx = \lambda(\frac{xe^{-\lambda x}}{\lambda}|_\infty^0 + \frac{\int_0^\infty e^{-\lambda x}dx}{\lambda}) = \frac{1}{\lambda}$$

$$Var(T) = \mathbb{E}[T^2] - \mathbb{E}[T]^2 = \int_0^\infty x^2\lambda e^{-\lambda x}dx - \mathbb{E}[T]^2 = \lambda \int_0^\infty x^2 e^{-\lambda x}dx - \mathbb{E}[T]^2$$

Integration by parts method:

$$Var(T) = \lambda \int_0^\infty x^2 e^{-\lambda x}dx - \mathbb{E}[T]^2 = \lambda(\frac{1}{\lambda^3} + \frac{1}{\lambda^2}) - \frac{1}{\lambda} = \frac{1}{\lambda^2}$$

Theoretical Question 17: Proof that the exponential distribution has no memory:

Let $X \sim \text{Exp}(\lambda)$, meaning the probability density function (PDF) of X is:

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

We need to show that:

$$P(X > t + s \mid X > t) = P(X > s)$$

Using the definition of conditional probability:

$$P(X > t + s \mid X > t) = \frac{P(X > t + s \cap X > t)}{P(X > t)}$$

Note that $P(X > t + s \cap X > t) = P(X > t + s)$, so we have:

$$P(X > t + s \mid X > t) = \frac{P(X > t + s)}{P(X > t)}$$

The survival function for the exponential distribution is given by:

$$P(X > x) = e^{-\lambda x}, \quad x \geq 0$$

Thus:

$$P(X > t + s) = e^{-\lambda(t+s)}$$

and

$$P(X > t) = e^{-\lambda t}$$

Substituting these into the conditional probability:

$$P(X > t + s \mid X > t) = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s}$$

Finally, since $P(X > s) = e^{-\lambda s}$, we conclude that:

$$P(X > t + s \mid X > t) = P(X > s)$$

Thus, the exponential distribution has no memory.

Theoretical Question 18: finding the distribution of $\text{Min}(X_1, X_2)$:

To find the distribution of $Y = \min(T_1, T_2)$, where T_1 and T_2 are independent exponential random variables with rate parameters λ_1 and λ_2, respectively, we proceed as follows:

CDF of Y

The CDF of $Y = \min(T_1, T_2)$ is given by:

$$F_Y(y) = P(Y \le y) = P(\min(T_1, T_2) \le y)$$

Since T_1 and T_2 are independent, we can write this as:

$$F_Y(y) = 1 - P(T_1 > y \text{ and } T_2 > y)$$

The probability that both $T_1 > y$ and $T_2 > y$ is:

$$P(T_1 > y \text{ and } T_2 > y) = P(T_1 > y) \cdot P(T_2 > y)$$

For exponential distributions, the survival function (probability that the random variable exceeds a value) is given by:

$$P(T_1 > y) = e^{-\lambda_1 y}, \quad P(T_2 > y) = e^{-\lambda_2 y}$$

Thus, the CDF becomes:

$$F_Y(y) = 1 - e^{-\lambda_1 y} \cdot e^{-\lambda_2 y} = 1 - e^{-(\lambda_1 + \lambda_2)y}$$

PDF of Y

To find the PDF of Y, we differentiate the CDF:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(1 - e^{-(\lambda_1 + \lambda_2)y}\right)$$

This gives:

$$f_Y(y) = (\lambda_1 + \lambda_2)e^{-(\lambda_1 + \lambda_2)y}$$

Final Answer

The PDF of $Y = \min(T_1, T_2)$ is:

$$f_Y(y) = (\lambda_1 + \lambda_2)e^{-(\lambda_1 + \lambda_2)y}, \quad y \geq 0$$

we can coclude that for n random variables $T_1, T_2, \ldots, T_n$ we have:
we define $S = Min(T_1, T_2, \ldots, T_n)$:

$$f_S(s) = (\lambda_1 + \lambda_2 + \cdots + \lambda_n)e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_n)s}, \quad s \geq 0$$

Theoretical Question 19: finding $P(T_1 \leq T_2)$:

we need to integrate the joint distributions of T_1, T_2 under the $T_1 \leq T_2$ region.

$$P(T_1 \leq T_2) = \iint_{T_1 \leq T_2} f_{T_1, T_2}(t_1, t_2)dt_1 dt_2$$

T_1, T_2 are independent, so:

$$P(T_1 \leq T_2) = \iint_{T_1 \leq T_2} \lambda_1 \lambda_2 e^{-(\lambda_1 t_1 + \lambda_2 t_2)}dt_1 dt_2 = \lambda_1 \lambda_2 \int_0^\infty \int_0^{t_1} e^{-(\lambda_1 t_1 + \lambda_2 t_2)}dt_2 dt_1$$

$$P(T_1 \leq T_2) = \lambda_1 \lambda_2 \int_0^\infty e^{-\lambda_1 t_1}\left(\int_0^{t_1} e^{-\lambda_2 t_2}dt_2\right)dt_1 = \lambda_1 \lambda_2 \int_0^\infty e^{-\lambda_1 t_1}\left(\frac{1 - e^{-\lambda_2 t_1}}{\lambda_2}\right)dt_1$$

after calculation of this simple integral we have:

$$P(T_1 \leq T_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

Theoretical Question 20: Maximum of n Exponential distributions:

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables following an exponential distribution with rate parameter λ. We aim to find the expected value of

$$Z = \max\{X_1, X_2, \ldots, X_n\}.$$

1. CDF of Z

The cumulative distribution function (CDF) of the exponential random variable $X \sim \text{Exp}(\lambda)$ is given by:

$$F_X(x) = 1 - e^{-\lambda x} \quad \text{for } x \geq 0.$$

Thus, the CDF of Z is:

$$F_Z(z) = P(Z \leq z) = P(\max\{X_1, \ldots, X_n\} \leq z) = P(X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z).$$

Since the X_i are i.i.d.:

$$F_Z(z) = (F_X(z))^n = (1 - e^{-\lambda z})^n.$$

2. PDF of Z

The probability density function (PDF) is obtained by differentiating the CDF:

$$f_Z(z) = \frac{d}{dz} F_Z(z) = n(1 - e^{-\lambda z})^{n-1} \cdot \lambda e^{-\lambda z}.$$

Therefore, we have:

$$f_Z(z) = n\lambda e^{-\lambda z}(1 - e^{-\lambda z})^{n-1} \quad \text{for } z \geq 0.$$

3. Expected Value $E[Z]$

The expected value is given by:

$$E[Z] = \int_0^\infty z f_Z(z) \, dz.$$

Using the known property of the maximum of i.i.d. exponential random variables, we can express the expected value as:

$$E[Z] = \frac{H_n}{\lambda},$$

where H_n is the n-th harmonic number:

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}.$$

Final Result

Thus, the expected value of Z is:

$$E[Z] = \frac{H_n}{\lambda}.$$
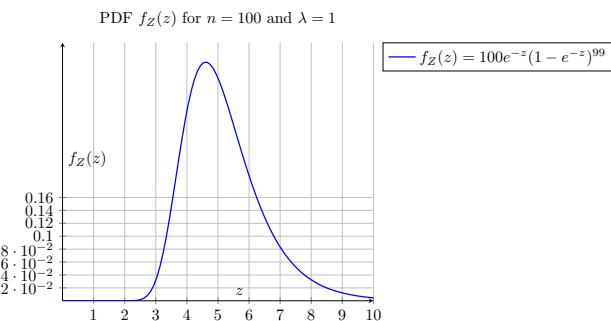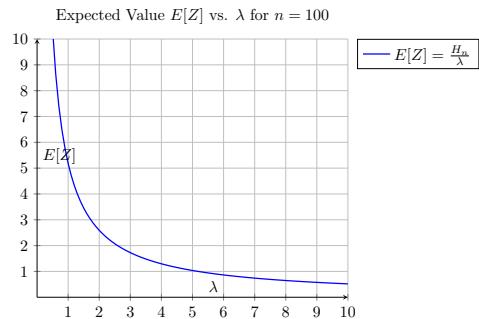
where H_n is the n-th harmonic number.
For large n, the harmonic number can be approximated as:

$$H_n \approx \ln(n) + \gamma,$$

where γ is the Euler-Mascheroni constant (approximately 0.577).
Substituting this approximation into the expression for $E[Z]$, we get:

$$E[Z] \approx \frac{\ln(n) + \gamma}{\lambda}.$$

Expected Value $E[Z]$ vs. λ for $n = 100$

PDF $f_Z(z)$ for $n = 100$ and $\lambda = 1$

> **Theoretical Question 21: proving a prperty about**
> $Min(X, Y)$ **where** $X \sim exp(\lambda_X)$ **and** $Y \sim exp(\lambda_Y)$ **:**

To find $P(X > t, X < Y)$, we can express this probability as:

$$P(X > t, X < Y) = \int_t^\infty P(X < Y | X = x) f_X(x)\, dx$$

Since Y is independent of X, we have:

$$P(X < Y | X = x) = P(Y > x) = e^{-\lambda_Y x}$$

The probability density function (PDF) of X is given by:

$$f_X(x) = \lambda_X e^{-\lambda_X x}, \quad x \geq 0$$

Therefore, we can set up the integral as follows:

$$P(X > t, X < Y) = \int_t^\infty e^{-\lambda_Y x} f_X(x) \, dx$$

Substituting $f_X(x)$:

$$P(X > t, X < Y) = \int_t^\infty e^{-\lambda_Y x} (\lambda_X e^{-\lambda_X x}) \, dx$$

Combining the exponentials:

$$= \lambda_X \int_t^\infty e^{-(\lambda_X + \lambda_Y) x} \, dx$$

Now we evaluate the integral:

$$\int_t^\infty e^{-(\lambda_X + \lambda_Y) x} \, dx = \left[\frac{e^{-(\lambda_X + \lambda_Y) x}}{-(\lambda_X + \lambda_Y)} \right]_t^\infty = \frac{e^{-(\lambda_X + \lambda_Y) t}}{\lambda_X + \lambda_Y}$$

Putting this back into the probability expression, we obtain:

$$P(X > t, X < Y) = \lambda_X \cdot \frac{e^{-(\lambda_X + \lambda_Y) t}}{\lambda_X + \lambda_Y}$$

Thus, we conclude that:

$$P(X > t, X < Y) = \frac{\lambda_X e^{-(\lambda_X + \lambda_Y) t}}{\lambda_X + \lambda_Y}$$

Find $P(Z > t)$

Let $Z = \min(X, Y)$ where $X \sim \text{Exp}(\lambda_X)$ and $Y \sim \text{Exp}(\lambda_Y)$. The cumulative distribution function $F_Z(z)$ is given by:

$$F_Z(z) = P(Z \leq z) = P(\min(X, Y) \leq z) = 1 - P(X > z, Y > z)$$

Since X and Y are independent:

$$P(X > z, Y > z) = P(X > z) P(Y > z)$$

For exponential distributions, we have:

$$P(X > z) = e^{-\lambda_X z}, \quad P(Y > z) = e^{-\lambda_Y z}$$

Thus:

$$P(X > z, Y > z) = e^{-\lambda_X z} e^{-\lambda_Y z} = e^{-(\lambda_X + \lambda_Y)z}$$

Therefore, the expression for $F_Z(z)$ becomes:

$$F_Z(z) = 1 - e^{-(\lambda_X + \lambda_Y)z}$$

Consequently, the probability $P(Z > t)$ is:

$$P(Z > t) = 1 - F_Z(t) = e^{-(\lambda_X + \lambda_Y)t}$$

Thus, we conclude that:

$$P(Z > t) = e^{-(\lambda_X + \lambda_Y)t}$$

Final Expression

Substituting back into the conditional probability, we have:

$$P(X > t | X < Y) = \frac{P(X > t, X < Y)}{P(X < Y)}$$

From our previous calculations, we know that:

$$P(X > t, X < Y) = \frac{\lambda_X e^{-(\lambda_X + \lambda_Y)t}}{\lambda_X + \lambda_Y}$$

And we found:

$$P(X < Y) = \frac{\lambda_X}{\lambda_X + \lambda_Y}$$

Thus, the conditional probability can be expressed as:

$$P(X > t | X < Y) = \frac{P(X > t, X < Y)}{P(X < Y)} = \frac{\frac{\lambda_X e^{-(\lambda_X + \lambda_Y)t}}{\lambda_X + \lambda_Y}}{\frac{\lambda_X}{\lambda_X + \lambda_Y}}$$

Simplifying this gives:

$$P(X > t | X < Y) = e^{(-\lambda_X + \lambda_Y)t} = P(Z > t)$$

Theoretical Question 22: if $X \sim poisson(\lambda)$ prove this :

$$E[X(X - 1)(X - 2) \ldots (X - k + 1)] = \lambda^k$$

we form a random variable like t^X where $t \in R$

$$\frac{d^k}{dt^k}E[t^X]|_{t=1} = E[\frac{d^k}{dt^k}t^X]|_{t=1} = E[X(X-1)(X-2)\ldots(X-k+1)]$$

we know that $\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!} = 1$ so based of Expectation definition and formula we have:

$$E[t^X] = \sum_{x=0}^{\infty}t^x\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(t\lambda)^x}{x!} = e^{-\lambda}e^{\lambda t} = e^{\lambda(t-1)}$$

$$\frac{d^k}{dt^k}E[t^X]|_{t=1} = \frac{d^k}{dt^k}e^{\lambda(t-1)}|_{t=1} = \lambda^k e^{\lambda(t-1)}|_{t=1} = \lambda^k$$

Theoretical Question 23: if $X \sim poisson(\lambda)$ prove this :

we use this property of independent variables like $X_1, X_2, \ldots, X_n$:

$$M_{X_1,X_2,\ldots,X_n}(s_1, s_2, \ldots, s_n) = E[e^{s_1 X_1 + s_2 X_2 + \cdots + s_n X_n}] = E[e^{s_1 X_1}e^{s_2 X_2}\ldots e^{s_n X_n}]$$

because $X_1, X_2, \ldots, X_n$ are independent:

$$E[e^{s_1 X_1}e^{s_2 X_2}\ldots e^{s_n X_n}] = E[e^{s_1 X_1}]E[e^{s_2 X_2}]\ldots E[e^{s_n X_n}] = M_{X_1}(s_1)M_{X_2}(s_2)\ldots M_{X_n}(s_n)$$

so we have this important property:

$$M_{X_1,X_2,\ldots,X_n}(s_1, s_2, \ldots, s_n) = M_{X_1}(s_1)M_{X_2}(s_2)\ldots M_{X_n}(s_n)$$

now we calculate M_X when $X \sim poisson(\lambda)$:

$$E[e^{sX}] = \sum_{x=0}^{\infty}e^{sx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(e^s\lambda)^x}{x!} = e^{-\lambda}e^{e^s\lambda} = e^{\lambda(e^s-1)}$$

so we define $Y = X_1 + X_2 + \cdots + X_n$ and then we calculate $M_Y(s)$:

$$M_Y(s) = M_{X_1}(s)M_{X_2}(s)\ldots M_{X_n}(s) = e^{\lambda_1(e^s-1)}e^{\lambda_2(e^s-1)}\ldots e^{\lambda_n(e^s-1)} = e^{(\sum_{i=1}^{n}\lambda_i)(e^s-1)}$$

then we can conclude Y is a random variable coming from Poisson($\sum_{i=1}^{n}\lambda_i$):

$$X_1 + X_2 + \cdots + X_n \sim poisson(\sum_{i=1}^{n}\lambda_i)$$

Theoretical Question 24: derivation of poisson from binomial distribution :

we have n points that the probability of each on of them to be in the τ interval:

$$p_i = \frac{\tau}{T}$$

if $T, n \to \infty$ then $p_i \to 0$ for $1 \le i \le n$

number of points which are in τ interval is a random variable which comes from binomial distribution:

$$P_X[x] = \binom{n}{x}(\frac{\tau}{T})^x(1 - \frac{\tau}{T})^{n-x}$$

where $\frac{n}{T} = \lambda$:

$$P_X[x] = \binom{n}{x}(\frac{\tau}{T})^x(1 - \frac{\tau}{T})^{n-x} = \binom{n}{x}(\frac{\lambda\tau}{n})^x(1 - \frac{\lambda\tau}{n})^{n-x}$$

as $n \to \infty$ each parts is equal to:

$$\binom{n}{x} \approx \frac{n^x}{x!}$$

$$(1 - \frac{\lambda\tau}{n})^{n-x} \approx (1 - \frac{\lambda\tau}{n})^n \approx e^{-\lambda\tau}$$

Substitute the approximations into the PMF:

$$P_X[x] = \binom{n}{x}(\frac{\lambda\tau}{n})^x(1 - \frac{\lambda\tau}{n})^{n-x} \approx \frac{n^x}{x!}(\frac{\lambda\tau}{n})^x e^{-\lambda\tau} = \frac{(\lambda\tau)^x}{x!}e^{-\lambda\tau}$$

so if $n \to \infty, T \to \infty$ and $\frac{n}{T} = \lambda$ we have:

$$X \sim poisson(\lambda\tau)$$

> **Theoretical Question 25:** relation between Poisson and Exp distribution :

we define a random variable T, which is the time that it takes to see a point
note that the reference of time dosn't matter because the occurrence of points are independent and memoryless

$$F_T(t) = 1 - P[T \ge t] = 1 - P[X = 0]$$

note that in the previous question we find that $X \sim poisson(\lambda t)$, so:

$$F_T(t) = P[T \le t] = P[X = 0] = 1 - e^{-\lambda t}\frac{(\lambda t)^0}{0!} = 1 - e^{-\lambda t}$$

since:

$$f_T(t) = \frac{d}{dt} F_T(t)$$
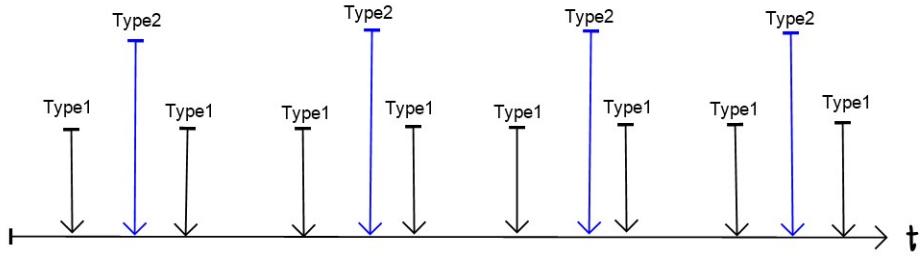
$$f_T(t) = \frac{d}{dt}(1 - e^{-\lambda t}) = \lambda e^{-\lambda t}$$

then we have:

$$T \sim Exp(\lambda)$$

> ## Theoretical Question 26: sum of 2 different types of arrival:

assume we have 2 types of clients (type1, type2) which each one of them have Poisson process model. how we can model the arrival when we consider both of them:

we can easily say that the total number of people which are in que at time t, is the sum of total number of people which are type 1 and total number of people which are type 2:

$$N(t) = N_1(t) + N_2(t)$$

N_1, N_2 are poisson process so:

$$P[N_1(t) = n] = e^{-\lambda_1 t} \frac{(\lambda_1 t)^n}{n!}$$

$$P[N_2(t) = n] = e^{-\lambda_2 t} \frac{(\lambda_2 t)^n}{n!}$$

we know that:

$$X_1 + X_2 + \cdots + X_n \sim poisson(\sum_{i=1}^{n} \lambda_i)$$
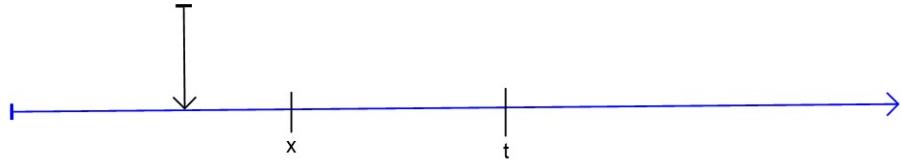
thus:

$$P[N(t) = n] = e^{-(\lambda_1 + \lambda_2)t} \frac{((\lambda_1 + \lambda_2)t)^n}{n!}$$

21

now we can easily (like before) prove that this kind of process has 3 properties of Poisson process

Theoretical Question 27: Uniform distribution!

we want prove that the below probability is actualy the CDF of Uniform destribution. T is the random variable that denotes in what point in the $(0, t)$ interval the person has been arrived:

$$P[T \leq x | N(t) = 1] = \frac{P[T \leq x, N(t) = 1]}{P[N(t) = 1]}$$

the numerator means that, one person arrives at (0, x) and nobody comes at (x, t)

$$P[T \leq x | N(t) = 1] = \frac{P[N(x) = 1]e^{-\lambda(t-x)}}{P[N(t) = 1]} = \frac{e^{-\lambda(t-x)}e^{-\lambda x}\frac{(\lambda x)^1}{1!}}{e^{-\lambda t}\frac{(\lambda t)^1}{1!}} = \frac{x}{t}$$

the result is:

$$T | N(t) = 1 \sim Uniform(t)$$

Theoretical Question 28: $E[T|T \leq c]$!

definition of conditional expectation

by definition we have:

$$E[T|T < c] = \int_0^\infty t f_{T|T<c}(t)dt = \int_0^\infty t \frac{f_T(t).I}{P[T < c]}dt$$

in above equation, I is indicator:

$$I(t) = \begin{cases} 0 & t \geq c, \\ 1 & t < c. \end{cases}$$

$$E[T|T < c] = \int_0^c t \frac{f_T(t)}{\int_0^c \lambda e^{-\lambda t}dt}dt = \frac{1}{\int_0^c \lambda e^{-\lambda t}dt}\int_0^c t\lambda e^{-\lambda t}dt$$

after simplification and calculating above integrals, we have:

$$E[T|T < c] = \frac{1}{\lambda} - \frac{c}{e^{\lambda c} - 1}$$

22

calculating by below formula

$$E[T] = P[T < c] \times E[T|T < c] + P[T > c] \times E[T|T > c].$$

$$\frac{1}{\lambda} = (1 - e^{-c\lambda}) \times E[T|T < c] + e^{-c\lambda} \times E[T|T > c]$$

we know that $\text{Exp}(\lambda)$ is a memoryless distribution(we proved before), so if we know that T is greater than c, nothing would change and the conditional expectation can easily be like:

$$E[T|T > c] = c + \frac{1}{\lambda}$$

so:

$$\frac{1}{\lambda} = (1 - e^{-c\lambda}) \times E[T|T < c] + e^{-c\lambda} \times c + \frac{1}{\lambda}$$

we obtain the conditional Expectation:

$$E[T|T < c] = \frac{1}{\lambda} - \frac{c}{e^{\lambda c} - 1}$$

> ## Theoretical Question 29:when the student wants to cross the street!

first we calculate the probability of one single $\text{Exp}(\lambda)$ to be more than c:

$$P[t_i > c] = \int_c^\infty t\lambda e^{-\lambda t} dt = e^{-\lambda c}$$

so know we have a geometric distribution for J of parampeter $p = e^{-\lambda c}$, because we keep tring every interval between car arrivals until we see an interval wider than c:

now we can find Ecpectation of j:

$$j \sim Geo(e^{-\lambda c}) + 1 \rightarrow E[j] = (\frac{1 - e^{-\lambda c}}{e^{-\lambda c}}) + 1 = e^{\lambda c}$$

now we have:

$$T_{j-1} = t_1 + t_2 + \cdots + t_{j-1}$$

note that every t_i , $1 \leq i \leq j - 1$ are less than c:

$$E[T_{j-1} + c] = E[t_1 + t_2 + \cdots + t_{j-1}] + c = E[j - 1] \times E[t|t < c] + c =$$

$$E[T_{j-1} + c] = (e^{\lambda c} - 1)(\frac{1}{\lambda} - \frac{c}{e^{\lambda c} - 1}) + c$$
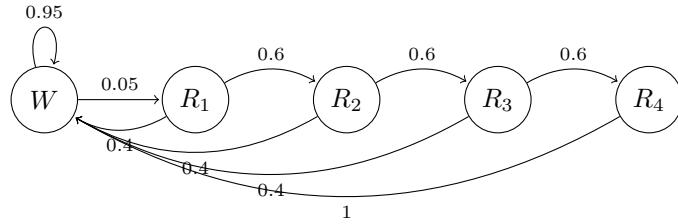
after simplification of above equation, we have:

$$E[T_{j-1} + c] = \frac{e^{\lambda c} - 1}{\lambda}$$

Theoretical Question 30: a Markov chain question:

fist we derive the state diagram:

W represent the working day and R_1, R_2, R_3, R_4 are day one till day 4 in Repair days:

trinsition matrix can be obtained very easily from the diagram:

$$P = \begin{bmatrix} 0.95 & 0.05 & 0 & 0 & 0 \\ 0.4 & 0 & 0.6 & 0 & 0 \\ 0.4 & 0 & 0 & 0.6 & 0 \\ 0.4 & 0 & 0 & 0 & 0.6 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Theoretical Question 31: C-k formula:

first we define a new concept. $P_{ij}^{(m)}$ is the probability of transition from i to j in m discrete steps of time.

$$P_{ij}^{(m)} = P[X_m = j | X_0 = i]$$

and now we difine $P^{(m)}$. a matrix that $P_{ij}^{(m)} = P[X_m = j | X_0 = i]$ and we call it m step trinsition matrix

we want to calculate $P^{(2)}$ from P matrix:

$$P_{ij}^{(2)} = \sum_s P_{is} \times P_{sj}$$

if we look closely to above equation we find out that it is the definition of matrix multiplication. so:

$$P^{(2)} = P \times P = P^2$$

We can prove by inductio:

$$P^{(m)} = P \times P \times \cdots \times P = P^m$$

so the m step transition matrix can be obtained easily by m times multiplication of primary transition matrix.

so we have for any $k \in \{0, 1, 2, \dots\}$:

$$P_{ij}^n = P[X_{k+n} = j | X_k = i]$$

Theoretical Question 32: independence of steady state probabilities of initial state

we say that steady state probability distribution is independent of initial state if for any $i \in \{0, 1, 2, \dots\}$ exists $\pi_k \in (0, 1)$ that holds this equation:

$$\lim_{n \to \infty} P[X_n = k | X_o = i] = \pi_k$$

the conditions for π_k to exists is that the Markov chain has these properties:
- Irreducibility

- Aperiodicity

but we don't want to talk about them here

if we assume X_0 is a vector that its sum of elements equals to 1 and it represents the probability distribution of states initially, we can have:

$$\lim_{n \to \infty} P[X_n = k] = \lim_{n \to \infty} X_0 \times P^n = \pi_k$$

for any X_0 vector (but acceptable one) the above equation should be holden, so for any i;

$$\lim_{n \to \infty} P_{ik}^n = \pi_k = \pi \times \mathbf{1}$$

where:

$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \dots \\ 1 \end{bmatrix}_{n \times 1}$$

$$\pi = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 & \dots \end{bmatrix}$$

Theoretical Question 33: why Irreducibility is necessary

- **Convergence to a Unique Distribution:** For a Markov chain to have a stationary distribution that is independent of the initial state, the system needs to be able to "forget" its starting state and reach an equilibrium where the probabilities of being in each state become independent of the initial conditions. Irreducibility ensures that, from any starting state, there is a non-zero probability of eventually reaching every other state, meaning that all states influence each other in the long run.

- **Without Irreducibility:** If the chain were reducible (i.e., there exist separate subsets of states that cannot reach each other), the chain would get "stuck" in one of these subsets, unable to transition to other subsets. In such cases, different subsets of the states could have different stationary distributions, and the stationary distribution would depend on the initial state (since the chain would not be able to reach all states). This would prevent the convergence to a unique stationary distribution that is independent of the initial state.

- **Transition Probabilities Across All States:** Irreducibility ensures that, for any starting state, the chain can eventually reach any other state. This guarantees that the system's behavior does not depend on where it started. The system "mixes" across all the states, and in the long run, each state has the same probability of being reached, determined by the stationary distribution.

In summary, irreducibility ensures that:

- Every state is reachable from every other state, leading to mixing of probabilities across all states.

- Eventually, as the number of steps increases, the distribution of states approaches the stationary distribution, which is independent of the initial state, because the chain forgets its starting point.

Without irreducibility, the system could get trapped in a subset of states and fail to converge to a steady state that is independent of the initial state. Thus, irreducibility is essential for ensuring the existence of a unique stationary distribution that does not depend on the initial state.

Theoretical Question 34: prove of convergence:

fist we define a vector that has a 1 at position j and all zero at other indexes:

$$\vec{e_j} = \begin{bmatrix} 0 \\ \cdots \\ 0 \\ 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

now we have:

$$[P^n\vec{e_j}]_i = P[X_n = j | X_0 = i]$$

so if all the elements in $P^n\vec{e_j}$ as $n \to \infty$ be the same, we can conclude that the steady state is independent of initial state

if M_n be the largest element in $P^n\vec{e_j}$ and m_n be the smallest, we want to show that:

$$M_n - m_n \leq (1 - 2s)(M_{n-1} - m_{n-1})$$

which s is the smallest element in P.

to prove this, we want to find an upper-bond for M_n and a lower-bond for m_n

consider $P^{n-1}\vec{e_j}$ with its own M_{n-1}, m_{n-1}, then we have:

$$P^n\vec{e_j} = P \times P^{n-1}\vec{e_j}$$

$$P^n\vec{e_j} = \begin{bmatrix} \cdots \\ \cdots \\ m_n \\ \cdots \\ M_n \\ \cdots \\ \cdots \end{bmatrix} = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & s & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \times \begin{bmatrix} \cdots \\ m_{n-1} \\ \cdots \\ \cdots \\ \cdots \\ M_{n-1} \\ \cdots \end{bmatrix}$$

note that sum of elements in each row equals to 1. the optimum case that we can reach the largest element in $P^n\vec{e_j}$ is to look at the row that s exist and in that row all the elements equals to zero else than s and (1-s). and also the positions of m_{n-1} and s be the same and also the positions of M_{n-1} and (1-s) be the same

$$P^n\vec{e_j} = \begin{bmatrix} \cdots \\ \cdots \\ m_n \\ M_n \\ \cdots \\ \cdots \\ \cdots \end{bmatrix} = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & s & 0 & 0 & (1-s) & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \times \begin{bmatrix} \cdots \\ m_{n-1} \\ \cdots \\ \cdots \\ \cdots \\ M_{n-1} \\ \cdots \end{bmatrix}$$

we know it for sure that $s \leq 0.5$ and $s \leq (1-s)$ so ideally $M_n = s \cdot m_{n-1} + (1-s) \cdot M_{n-1}$

$$M_n \leq s \cdot m_{n-1} + (1-s) \cdot M_{n-1}$$

Theoretical Question 35: why $s \leq 0.5$:

we know that the sum of elements in each row equals to 1.N is the number of states. so for $1 \leq i \leq N$:

$$\sum_{j=1}^{N} P_{ij} = 1$$

assume s exists in row number m and column n:

$$\sum_{j=1}^{N} P_{mj} = s + \sum_{j \neq n} P_{mj} = 1$$

we know that for each $(i,j) \neq (m,n)$ $P_{ij} \geq s$
so it is obvious that for $N \geq 2$:

$$s \leq \sum_{j \neq n} P_{mj}$$

so

$$s + \sum_{j \neq n} P_{mj} = 1 \rightarrow 2s \leq 1 \rightarrow s \leq 0.5$$

Theoretical Question 36:proving a Theorem:

Theorem: If $\gcd(n_1, n_2) = 1$, then for any $n \geq n_0$, there exist $\alpha, \beta \in \mathbb{N}$ such that
$$n = \alpha \cdot n_1 + \beta \cdot n_2.$$

Proof:
1. **GCD Property:** Since $\gcd(n_1, n_2) = 1$, Bézout's identity guarantees that there exist integers x and y such that:

$$x \cdot n_1 + y \cdot n_2 = 1.$$

2. **General Form:** Multiplying both sides of the equation by $n \geq n_0$, we get:
$$n = (n \cdot x) \cdot n_1 + (n \cdot y) \cdot n_2.$$

Here, $n \cdot x$ and $n \cdot y$ are integers. To ensure $\alpha, \beta \in \mathbb{N}$, we need to ensure that both coefficients are non-negative.

3. **Non-Negativity Condition:** By the theory of linear Diophantine equations, there are infinitely many solutions of the form:

$$\alpha = n \cdot x + k \cdot n_2, \quad \beta = n \cdot y - k \cdot n_1,$$

where $k \in \mathbb{Z}$ is an integer.

4. **Choosing k:** To ensure $\alpha, \beta \geq 0$, we adjust k such that:

$$n \cdot x + k \cdot n_2 \geq 0 \quad \text{and} \quad n \cdot y - k \cdot n_1 \geq 0.$$

Solving these inequalities, we find that a suitable k can always be chosen if $n \geq n_0$ for some sufficiently large n_0.

5. **Conclusion:** For any $n \geq n_0$, there exist $\alpha, \beta \in \mathbb{N}$ such that:

$$n = \alpha \cdot n_1 + \beta \cdot n_2.$$

we can add that this n_0 can be expressed as:

$$n_0 = n_1 n_2 - n_1 - n_2$$

for each $n \geq n_0$ a solution for (α, β) exists.

Theoretical Question 37:finding steady state probabilities in a simple state machine:

---

we want the solutions of this system of equations

$$\pi_\circ = \pi_\circ(1 - r) + \pi_1 s$$
$$\pi_1 = \pi_\circ r + \pi_1(1 - r - s) + \pi_\Upsilon s$$
$$\pi_\Upsilon = \pi_1 r + \pi_\Upsilon(1 - r - s) + \pi_\Upsilon s$$
$$\pi_\Upsilon = \pi_\Upsilon r + \pi_\Upsilon(1 - r - s) + \pi_\digamma s$$

$$\dots$$

$$\pi_\circ + \pi_1 + \pi_\Upsilon + \dots = 1$$

we derive the recursive equation:

$$\pi_{i+1} = \frac{\pi_i - r\pi_{i-1}}{s}$$

we substitute $\pi_i = m^i$:

$$m^{i+1} = \frac{m^i - rm^{i-1}}{s} \rightarrow m = \frac{1 - \frac{r}{m}}{s}$$

The characteristic equation:

$$sm^2 - m + r = 0 \rightarrow m_1 = 1, m_2 = \frac{r}{s}$$

$$\pi_i = C_1 + C_2(\frac{r}{s})^i$$

C_1 must be zero because otherwise $\sum_{i=0}^{\infty} \pi_i$ would not converge.

$$\sum_{i=0}^{\infty} \pi_i = \sum_{i=0}^{\infty} C_2(\frac{r}{s})^i = C_2 \sum_{i=0}^{\infty} (\frac{r}{s})^i = \frac{C_2}{1 - \frac{r}{s}} = 1$$

Thus:

$$C_2 = 1 - \frac{r}{s}$$

$$\pi_i = (1 - \frac{r}{s}) \cdot (\frac{r}{s})^i$$

Theoretical Question 38:Expectation of steady state:

now we want to calculate the expectation of how many jobs are in Que at steady state condition:

$$E[N] = \sum_{i=0}^{\infty} i \cdot \pi_i = \sum_{i=0}^{\infty} i \cdot (1 - \frac{r}{s}) \cdot (\frac{r}{s})^i = (1 - \frac{r}{s}) \sum_{i=0}^{\infty} i \cdot (\frac{r}{s})^i$$

we know that by derivation of $\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i$:

$$\frac{x}{(1-x)^2} = \sum_{i=0}^{\infty} i \cdot x^i$$

so:

$$E[N] = (1 - \frac{r}{s}) \sum_{i=0}^{\infty} i \cdot (\frac{r}{s})^i = (1 - \frac{r}{s}) \cdot \frac{\frac{r}{s}}{(1 - \frac{r}{s})^2} = \frac{\frac{r}{s}}{(1 - \frac{r}{s})} = \frac{r}{s - r}$$

Theoretical Question 39:Expectation of entries in recurrent states:

we define f_i as:

$$f_i = P[(X_{n+1} = i) \cup (X_{n+2} = i) \cup (X_{n+3} = i) \cup \ldots \ |X_n = i]$$

$$\begin{cases} f_i < 1 & \text{transient state} \\ f_i = 1 & \text{recurrent state} \end{cases}$$

Based on this definition, if state i is recurrent and the system enters this state, it will definitely return to i in the future (near or far). Similarly, from every state, it starts again and returns to the same state. This is also true for every other state where the number of times the system enters is infinite.

so the Expectation of Number of entries is ∞ and does not converge.

> ## Theoretical Question 40:Expectation of entries in transient states:

However, if state i is transient and the system starts from this state, with probability f, it will return to i, and with probability $1 - f$, it will never return.

X is a random variable representing the number of visits. so:

$$P[X = x] = f_i^x \cdot (1 - f_i)$$

so $X \sim Gep(1 - f)$ and the Expectation is:

$$E[X] = \frac{f_i}{1 - f_i}$$

> ## Theoretical Question 41: convergence of $\sum_{n=0}^{\infty} P_{ii}^n$

We define Y_i as:

$$Y_m = \begin{cases} 1 & X_m = i \\ 0 & X_m \neq i \end{cases}$$

note that we start from state i.

now we define N:

$$N_s = Y_1 + Y_2 + \cdots + Y_s$$

$E[N_s] = E[\text{how many times we be in state i} \mid \text{start with state i}] = E[Y_1] + E[Y_2] + \cdots + E[Y_s]$

we have:

$E[Y_n] = 1 \times P[\text{being in state i after n steps} \mid \text{start with state i}] + 0 \times \cdots = P_{ii}^n$

so:

$$E[N_s] = \sum_{n=0}^{s} P_{ii}^n$$

we can easily have $s \to \infty$:

31

$$E[\text{how many times we visit state i}|\text{ start with state i}] = \sum_{n=0}^{\infty} P_{ii}^n$$

> ## Theoretical Question 42: if i is recurrent and $i \leftrightarrow j$ then j is recurrent

If $i \leftrightarrow j$ and i is recurrent, then j is also recurrent.

Proof

Since i and j are related, the following relations hold for positive integers m and m':

$$p_{ij}^{(m')} > 0, \quad p_{ji}^{(m')} > 0$$

Using equation $C - K$, we have the relation below, which holds for all $n \geq m + m'$:

$$p_{jj}^{(n)} \geq p_{ji}^{(m)} p_{ii}^{(n-m-m')} p_{ij}^{(m')}$$

Summing over $n \geq m + m'$, we get:

$$\sum_{n=m+m'}^{\infty} p_{jj}^{(n)} \geq p_{ji}^{(m)} p_{ij}^{(m')} \sum_{n=m+m'}^{\infty} p_{ii}^{(n-m-m')}$$

Since i is recurrent, the summation $\sum_{n=m+m'}^{\infty} p_{ii}^{(n-m-m')} = \infty$, and m, m' are specific and finite integers.

Thus, $\sum_{n=m+m'}^{\infty} p_{jj}^{(n)} = \infty$, which implies that j is also recurrent.

> ## Theoretical Question 43: Given that there is no limiting distribution, show that there is no stationary distribution for an irreducible non-ergodic Markov chain

of there was any stationary distribution $\overrightarrow{S}$ we would say:

$$P \times \overrightarrow{S} = \overrightarrow{S}$$

and for every n (even for $n \to \infty$) we would have:

$$P^n \times \overrightarrow{S} = \overrightarrow{S}$$

we know that for a Markov chain that all the states are transient:

$$\lim_{n \to \infty} P_{ij} = 0, \forall i, j$$

then $\lim_{n \to \infty} P = $ a matrix of zeros

Therefore, we reached a contradiction.as $n \to \infty$

$$P^n \times \vec{S} = \vec{S} \text{ and also } P^n = \text{matrix of zeros}$$

note that $\vec{S}$ can not be a vector of zero because the summation of its elements must be 1 to be acceptable as a probability distribution

Theoretical Question 44: Lavrov lemma

$$\binom{2n}{n} = \frac{(2n)!}{n! \times n!} = \left[\left(\frac{2n}{n}\right) \times \left(\frac{2n-2}{n-1}\right) \times \cdots \times \left(\frac{2}{1}\right)\right]\left[\left(\frac{2n-1}{n}\right) \times \frac{2n-3}{n-1} \times \cdots \times \left(\frac{1}{1}\right)\right]$$

because $\left(\frac{2n-1}{n}\right) \leq 2$ for $n \geq 1$:

$$\binom{2n}{n} = 2^n \times \Pi_1^n \left(\frac{2n-1}{n}\right) \leq 2^n \cdot 2^n = 4^n$$

simillary we have:

$$\binom{2n}{n}(2n+1) = \frac{(2n)!}{n! \times n!} = \left[\left(\frac{2n}{n}\right) \times \left(\frac{2n-2}{n-1}\right) \times \cdots \times \left(\frac{2}{1}\right)\right]\left[\left(\frac{2n+1}{n}\right) \times \frac{2n-1}{n-1} \times \cdots \times \left(\frac{3}{1}\right)\right]$$

because $\left(\frac{2n+1}{n}\right) \geq 2$ for $n \geq 1$:

$$\binom{2n}{n}(2n+1) = 2^n \times \Pi_1^n \left(\frac{2n+1}{n}\right) \geq 2^n \cdot 2^n = 4^n$$

so:

$$\binom{2n}{n} \geq \frac{4^n}{2n+1}$$

we combine the results:

$$\frac{4^n}{2n+1} \leq \binom{2n}{n} \leq 4^n$$

Theoretical Question 45: proove $P[X_n = X_i | X_0 = X_i]$ for $n > 1$

we know that only if p=q=0.5, the probability of visiting initial state in future equals to 1 and also $M_i = \sum_{n=0}^{\infty} P_{ii}^n$ equals to ∞.

in other cases, M_i is a limited value, and we know from basic theorem that the probability of not visiting the initial state in the furue (probability of escape), equals to:

$$P_i^* = \frac{1}{M_i}$$

33

now we want to calculate M_i (for each i, results would be the same):

$$M_i = \sum_{n=0}^{\infty} \binom{2n}{n} (pq)^n$$

we know that the generative function would be:

$$\sum_{n=0}^{\infty} \binom{2n}{n} (x)^n = \frac{1}{\sqrt{1-4x}} \text{ for } |x| < \frac{1}{4}$$

thus:

$$M_i = \frac{1}{\sqrt{1-4(pq)}} = \frac{1}{1-2min(p,q)}$$

then the probability of visiting state i, in future:

$$p_i = 1 - p^* = 1 - \frac{1}{M_i} = 1 - (1 - 2min(p,q)) = 2min(p,q)$$

Theoretical Question 46: Ergodic Markov chains

The mathematical definition for a recurrent and transient states is:

$$\begin{cases} \sum_{n=0}^{\infty} P_{ii}^n < \infty & \text{i is a transient state} \\ \sum_{n=0}^{\infty} P_{ii}^n = \infty & \text{i is a recurrent state} \end{cases}$$

we saw before that $\sum_{n=0}^{\infty} P_{ii}^n$ is the expectation of 'Number of visit of state"
if $\sum_{n=0}^{\infty} P_{ii}^n$ for an state like k be finite be infinite but $\lim_{n\to\infty} P_{ii}^n = 0$ we call it a null recurrent state.

we want to prove that for a finite state Markov chain (if it has steady state), all statea has non zero $\lim_{n\to\infty} P_{ii}^n$:

$$m_i = E[\text{Expectaion of "how many steps is between two folowing visit}] = \frac{1}{\lim_{n\to\infty} P_{ii}^n} = \frac{1}{\pi_i}$$

note that if $\pi_i = 0$ then $m_i = \infty$ .assume j be a arbitary state.

$$m_{ii} = M_{ij} + m_{jj} + M_{ji}$$

M_{ij} is the Expectation of steps between states i and j. we know that for a Irreducible Markov chain, there is a path between every two states.since. if we assume that the chain has only finite states:

$$M_{ij}, M_{ji} \text{ is finite}$$

because $m_i = M_{ij} + m_j + M_{ji}$, we have:

34

$$m_i \leq \infty \leftrightarrow m_j \leq \infty$$

because i and j was arbitary and all states connected together and the number of states is finite. we conclude:

eather all states are transient, or all states are null recurrent or all states are positive recurrent.

we can not have a markov chain that all the states are null recurrent. because if so, all states can be visited infinite times and also the time between two visits is infinite. and that is not possible.

if all states are null recurent think of π vector:

$$\pi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix}$$

and also:

$$\sum_{i=0}^{N} \pi_i = 1$$

Contradiction!

so if we have a Irreducible and aperiodic Markov chain eather the chain is transient with no steady state or all states are positive recurrent.

Theoretical Question 47: proof of everything

in previous questions we discussed about different parts of this theorem.

we showed before that if two states are connected, then if one of them is transient(positive recurrent, null recurrent) then the other one must be transient (positive recurrent, null recurrent) too. look at question 46

the discussion in question 46 can be applied exactly for transient and recurrent property;

we saw that if 2 sates i and j are connected:

$$m_{ii} = M_{ij} + m_{jj} + M_{ji}$$

M_{ij} is the Expectation of steps between states i and j. we know that for a Irreducible Markov chain, there is a path between every two states.since. if we assume that the chain has only finite states:

$$M_{ij}, M_{ji} \text{ is finite}$$

because $m_i = M_{ij} + m_j + M_{ji}$, we have:

$$m_{ii} \leq \infty \leftrightarrow m_{jj} \leq \infty$$

we knew that:

$$\begin{cases} m_{ii} < \infty & \text{state i must be recurrent} \\ m_{ii} = \infty & \text{state i must be null-recurrent or transient} \end{cases}$$

if we assume all states are connected, there is two possibilities:
- all states are either transient or null-recurrent $\rightarrow$ there is no stationary distribution

- all states are positive recurrent $\rightarrow$ there must be a stationary distribution:

if the second possibility happend, we have a staionary distribution. by difinition we have:

$$X_n = P^n X_0$$

which X_0 is the first distribution and X_n is the nth distribution.

$$\lim_{n\to\infty} X_n = \lim_{n\to\infty} P^n X_0 =$$

$$\pi_j = \lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} P^n X_0 =$$

we discussed in question 33 that why in a irreducible and aperiodic Markov chain with stationary distribution, X_n is independent of X_0. then $\lim_{n\to\infty} P^n$ must be like:

$$\lim_{n\to\infty} P^n = \left[\pi_{m\times 1} | \pi_{m\times 1} | \pi_{m\times 1} | \pi_{m\times 1} | \ldots | \pi_{m\times 1} \right]$$

m is the number of states (can be ∞ as well).
the π vector is the stationary state distribution
we can easily show that the only stationary distribution is the steady state condition distribution. the proof exists at page 21 of help document.
then the only stationary distribution is the steady state distribution.

Theoretical Question 48: proof of everything

Consider a Markov chain where the limiting fraction of time spent in state j is given by:

$$p_j = \lim_{t\to\infty} \frac{N_j(t)}{t},$$

where $N_j(t)$ is the number of visits to state j up to time t.
In renewal theory:
- A **renewal event** occurs when the process returns to a specific state (state j in this case).

- **Cycle times**, denoted as $T_1, T_2, \ldots$, are the durations between successive visits to state j.

- The total number of renewals (or visits) by time t is $N_j(t)$.

The total time t can be decomposed as:

$$t = \sum_{k=1}^{N_j(t)} T_k + R(t),$$

where $R(t)$ is the residual time that does not complete a full cycle.

From renewal theory:

1. **Strong Law of Large Numbers for Renewals:**

$$\frac{N_j(t)}{t} \to \frac{1}{\mathbb{E}[T_j]}, \quad \text{as } t \to \infty,$$

where $\mathbb{E}[T_j]$ is the expected time between successive visits to state j.

2. **Stationary Behavior:** For an ergodic Markov chain, the limiting fraction of time spent in state j is the reciprocal of the mean recurrence time:

$$p_j = \frac{1}{\mathbb{E}[T_j]}.$$

Define τ_k as the time of the k-th visit to state j. Then:

$$N_j(t) = \max\{k : \tau_k \leq t\}.$$

The total time spent in state j over $N_j(t)$ cycles approximately equals t, and the **average cycle time** converges to $\mathbb{E}[T_j]$:

$$\lim_{t \to \infty} \frac{t}{N_j(t)} = \mathbb{E}[T_j].$$

Thus, we have:

$$\frac{N_j(t)}{t} \to \frac{1}{\mathbb{E}[T_j]}, \quad \text{as } t \to \infty.$$

The fraction of time spent in state j is given by:

$$p_j = \lim_{t \to \infty} \frac{N_j(t)}{t}.$$

Using renewal theory and the ergodic property of Markov chains:

$$p_j = \frac{1}{\mathbb{E}[T_j]},$$

where $\mathbb{E}[T_j]$ is the mean recurrence time for state j.

we discussed that $\pi_i P_{ij}$ is the probability (in long time) being in state i and have a transition of $i \to j$ in the next step. so if we sum these terms over j, we find the total rate of flow out of state i .

we discussed that $\pi_i P_{ji}$ is the probability(rate) (in long time) of have a transition to state i. and if we sum these terms over j, we find the total rate of flow into state i .

either in continues and discrete model of Markov chain, we have one important property that the history is not important (based on the second property at the help document):

the phrase below indicates that the events in future only depend the state at the present time not before:

$$P[\tau_i > t + s | \tau_i > s] = P[\tau_i > t]$$

it means that the times that has been passed (s) does not effect the 'probability' and 'the time we expect to take' to change the state.

References

[1] Dr. Mohammadmodares Yazdi, *Queueing Theory*, 1st ed. Tehran: Tehran University Publisher, 1991.

[2] Queueing Theory Explained, YouTube, 3 January 2025. [Online]. Available: https://www.youtube.com/watch?v=Jp0lPj2-DQA. [Accessed: 10 January 2025].

[3] *Queueing Theory*, YouTube, 3 January 2025. [Online]. Available: https://www.youtube.com/watch?v=rBIQmwaoZfs&t=1s&pp=ygUOcXVldWluZyB0aGVvcnk