Líneas de Espera: Teoría de Colas

Curso Métodos Cuantitativos

Prof. Lic. Gabriel Leandro

http://www.auladeeconomia.com

AuladeEconomía.com

Las colas…

- Las colas son frecuentes en nuestra vida cotidiana:
  - En un banco
  - En un restaurante de comidas rápidas
  - Al matricular en la universidad
  - Los autos en un lavacar

Las colas…

- En general, a nadie le gusta esperar
- Cuando la paciencia llega a su límite, la gente se va a otro lugar
- Sin embargo, un servicio muy rápido tendría un costo muy elevado
- Es necesario encontrar un balance adecuado

Teoría de colas

- Una cola es una línea de espera
- La teoría de colas es un conjunto de modelos matemáticos que describen sistemas de líneas de espera particulares
- El objetivo es encontrar el estado estable del sistema y determinar una capacidad de servicio apropiada

Teoría de colas

- Existen muchos sistemas de colas distintos
- Algunos modelos son muy especiales
- Otros se ajustan a modelos más generales
- Se estudiarán ahora algunos modelos comunes
- Otros se pueden tratar a través de la simulación

http://www.auladeeconomia.com

Sistemas de colas: modelo básico

- Un sistema de colas puede dividirse en dos componentes principales:
  - La cola
  - La instalación del servicio
- Los clientes o llegadas vienen en forma individual para recibir el servicio

Sistemas de colas: modelo básico

- Los clientes o llegadas pueden ser:
  - Personas
  - Automóviles
  - Máquinas que requieren reparación
  - Documentos
  - Entre muchos otros tipos de artículos

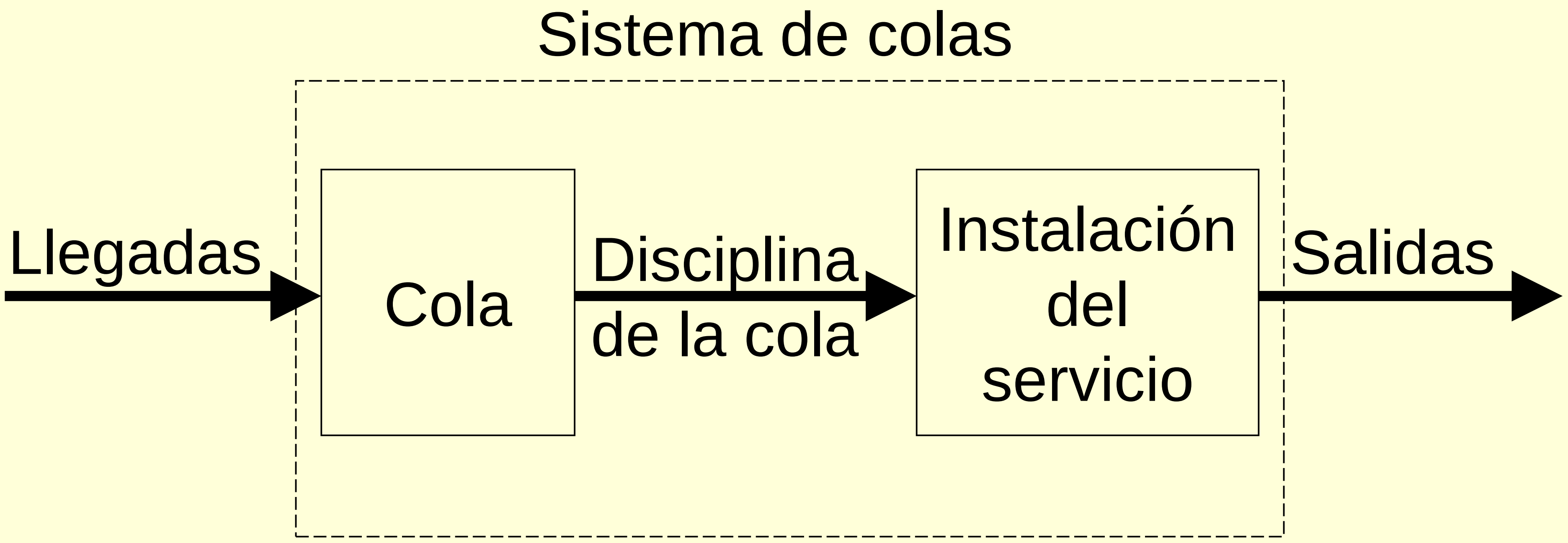
Sistemas de colas: modelo básico

- Si cuando el cliente llega no hay nadie en la cola, pasa de una vez a recibir el servicio
- Si no, se une a la cola
- Es importante señalar que la cola no incluye a quien está recibiendo el servicio

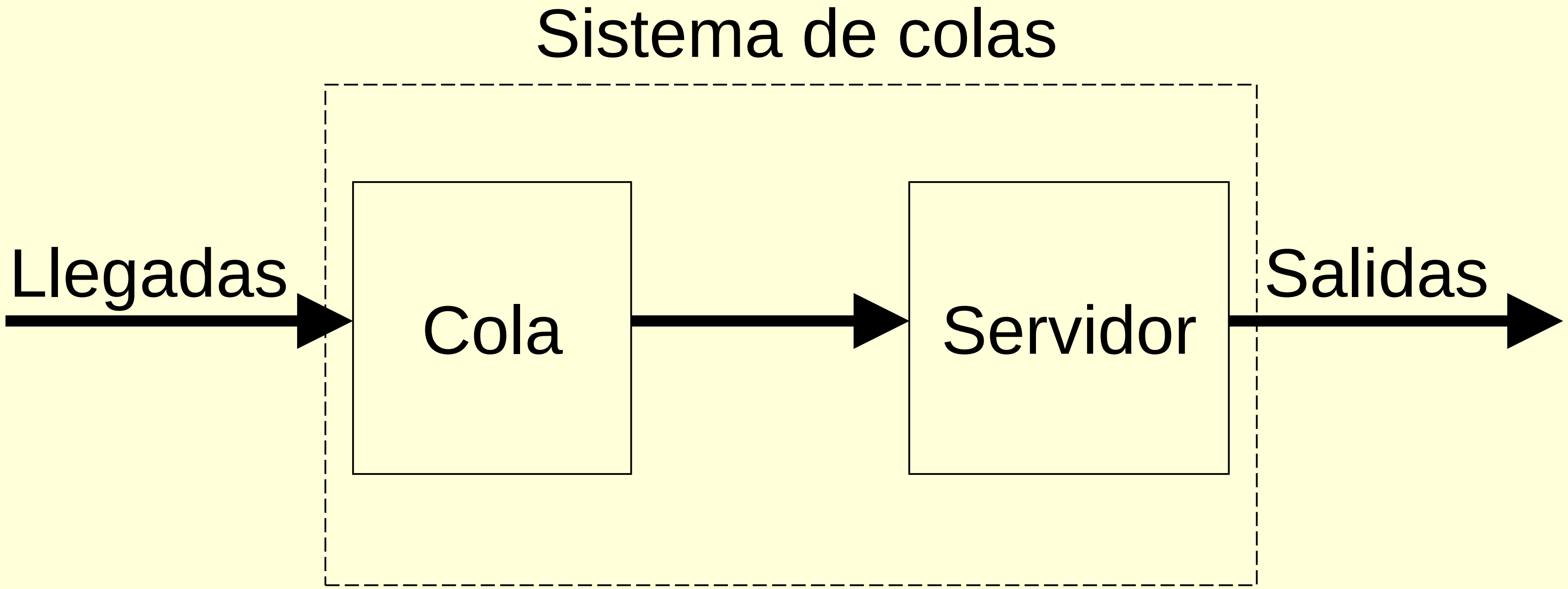
Sistemas de colas: modelo básico

- Las llegadas van a la instalación del servicio de acuerdo con la disciplina de la cola
- Generalmente ésta es *primero en llegar, primero en ser servido*
- Pero pueden haber otras reglas o colas con prioridades

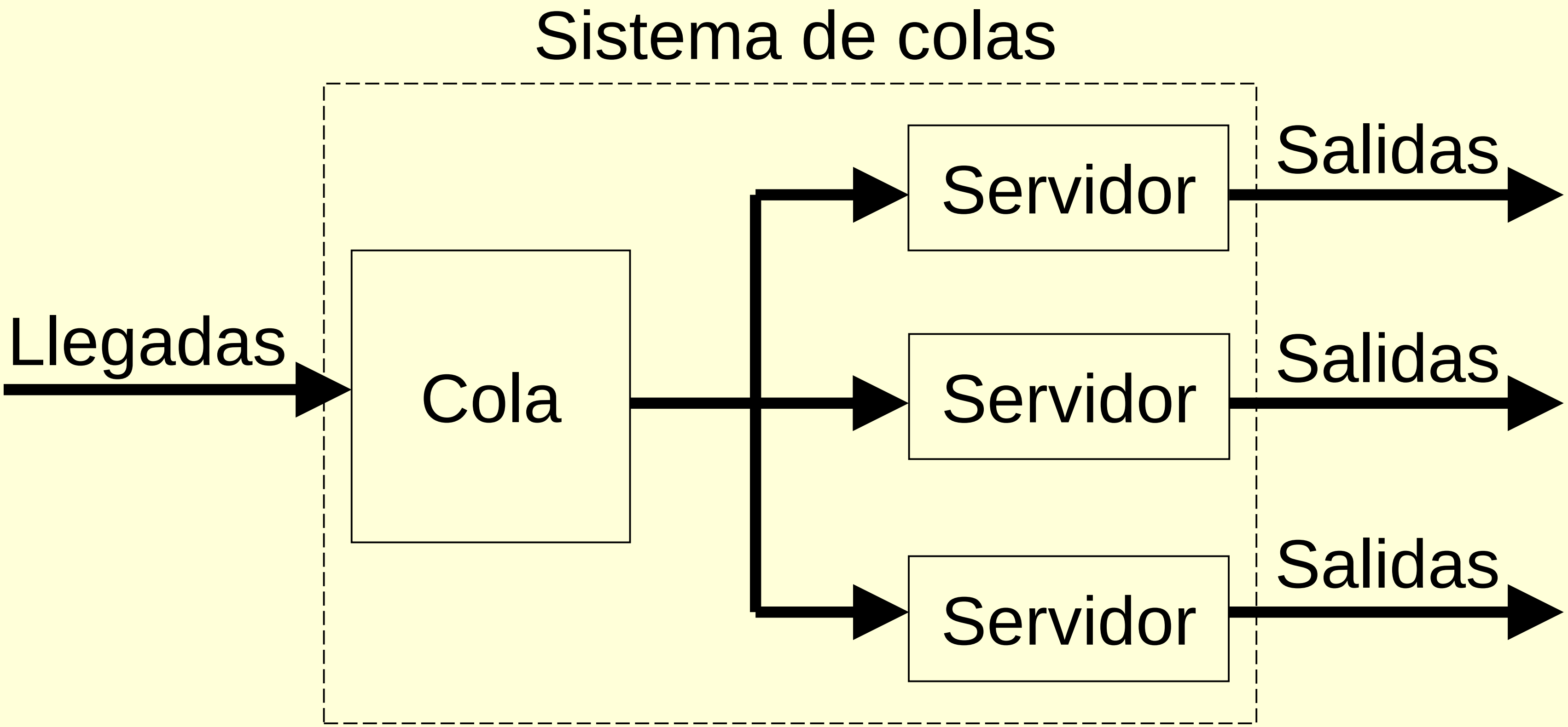
Sistemas de colas: modelo básico



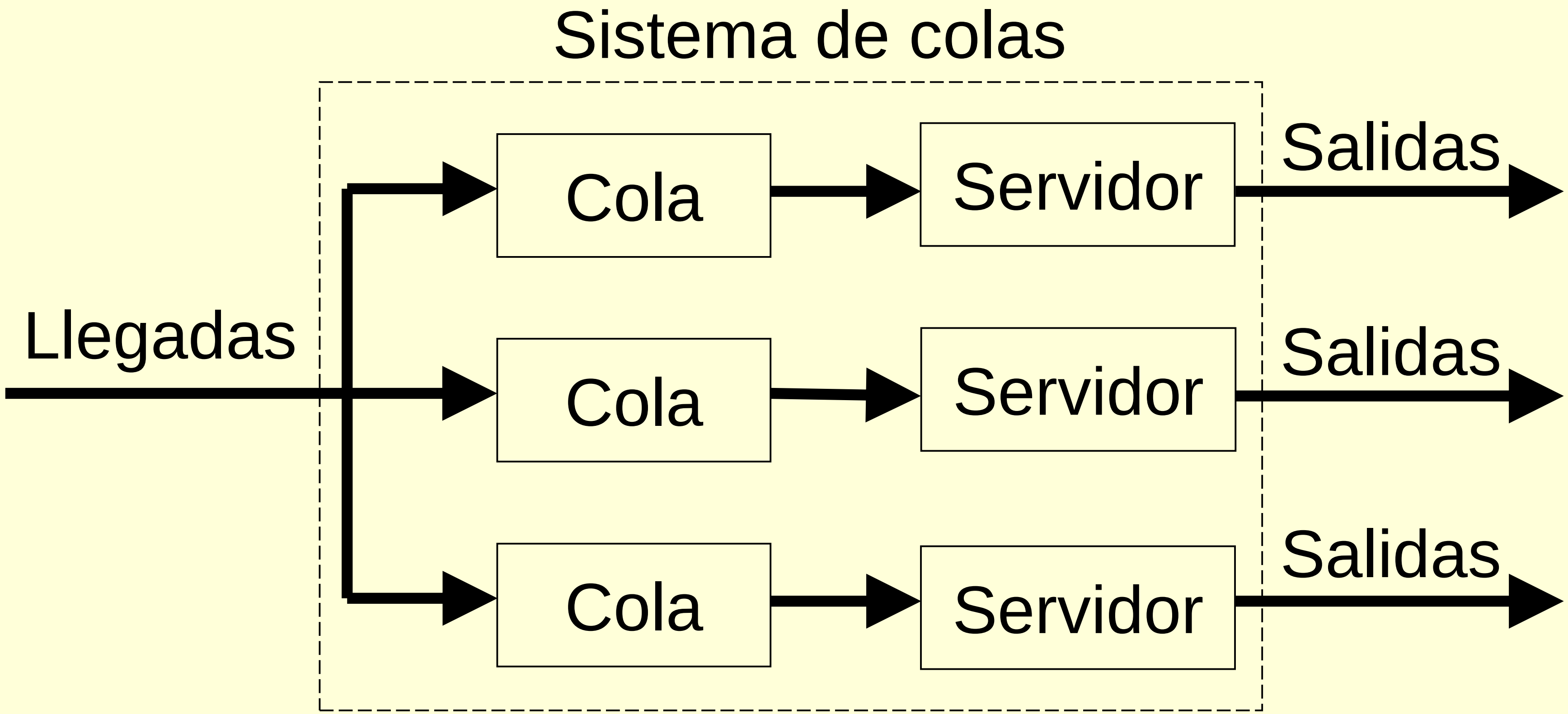
Estructuras típicas de sistemas de colas: una línea, un servidor



Estructuras típicas de sistemas de colas: una línea, múltiples servidores

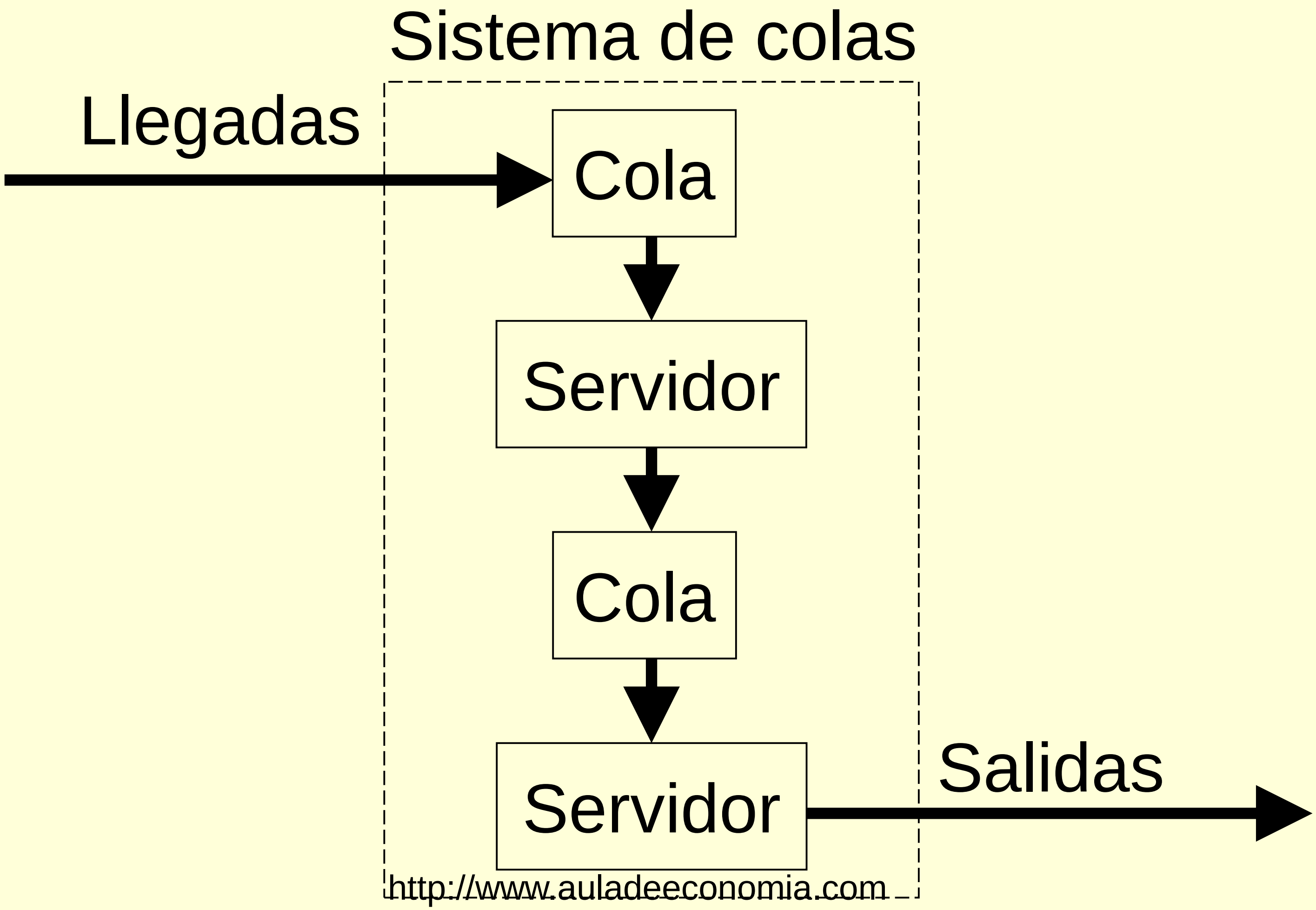


Estructuras típicas de colas: varias líneas, múltiples servidores



http://www.auladeeconomia.com

Estructuras típicas de colas: una línea, servidores secuenciales



http://www.auladeeconomia.com

Costos de un sistema de colas

1. Costo de espera: Es el costo para el cliente al esperar

- Representa el costo de oportunidad del tiempo perdido
- Un sistema con un bajo costo de espera es una fuente importante de competitividad

Costos de un sistema de colas

2. Costo de servicio: Es el costo de operación del servicio brindado

- Es más fácil de estimar
  - El objetivo de un sistema de colas es encontrar el sistema del costo total mínimo

Sistemas de colas: Las llegadas

- El tiempo que transcurre entre dos llegadas sucesivas en el sistema de colas se llama tiempo entre llegadas
- El tiempo entre llegadas tiende a ser muy variable
- El número esperado de llegadas por unidad de tiempo se llama tasa media de llegadas (λ)

Sistemas de colas: Las llegadas

- El tiempo esperado entre llegadas es $1/\lambda$
- Por ejemplo, si la tasa media de llegadas es $\lambda = 20$ clientes por hora
- Entonces el tiempo esperado entre llegadas es $1/\lambda = 1/20 = 0.05$ horas o 3 minutos

Sistemas de colas: Las llegadas

- Además es necesario estimar la distribución de probabilidad de los tiempos entre llegadas
- Generalmente se supone una distribución exponencial
- Esto depende del comportamiento de las llegadas

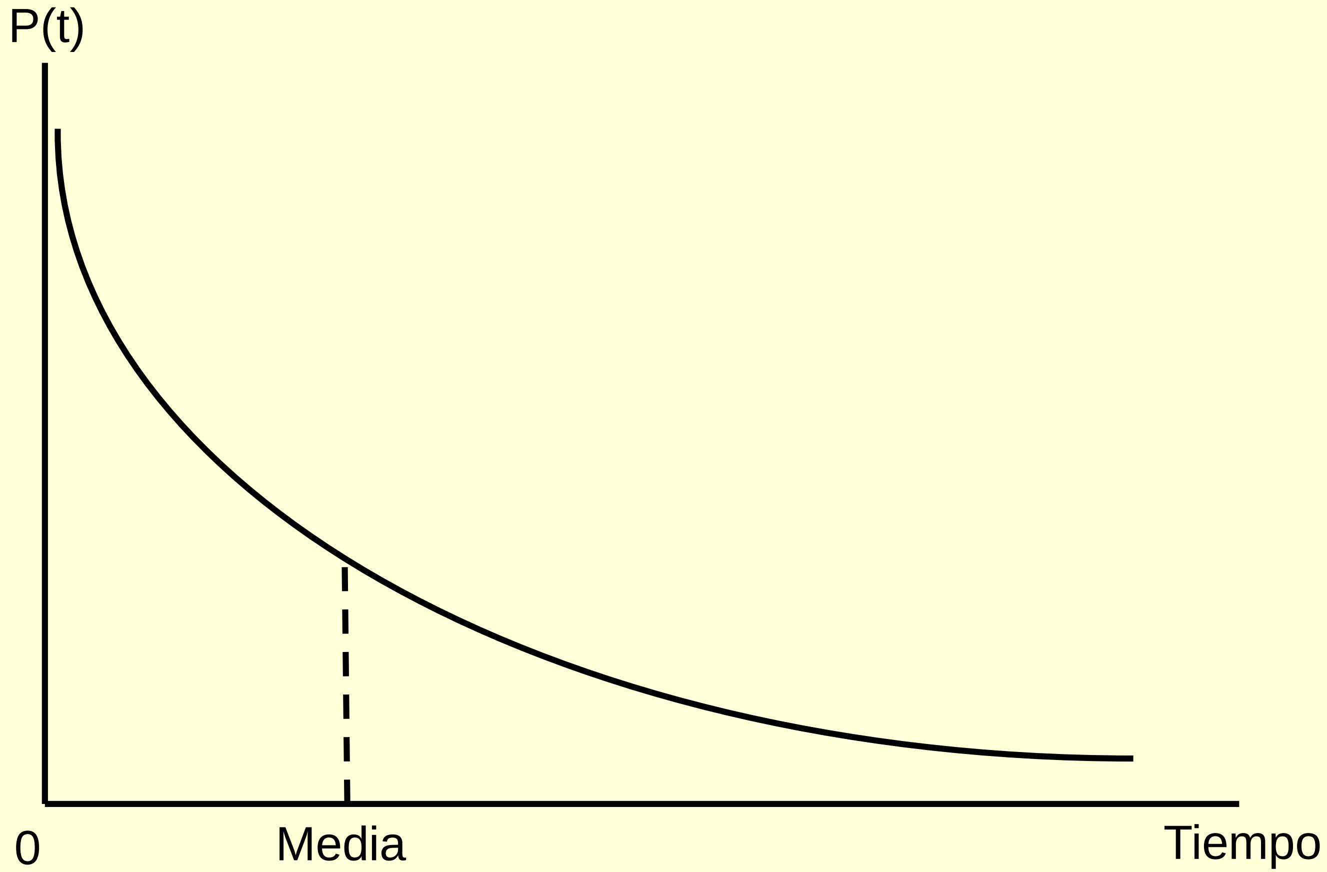
Sistemas de colas: Las llegadas – Distribución exponencial

- La forma algebraica de la distribución exponencial es: ????

$$P(tiempo\, de\, servicio \leq t) = 1 - e^{-\mu t}$$

- Donde ***t*** representa una cantidad expresada en de tiempo unidades de tiempo (horas, minutos, etc.)

Sistemas de colas: Las llegadas – Distribución exponencial



http://www.auladeeconomia.com

Sistemas de colas: Las llegadas – Distribución exponencial

- La distribución exponencial supone una mayor probabilidad para tiempos entre llegadas pequeños
- En general, se considera que las llegadas son aleatorias
- La última llegada no influye en la probabilidad de llegada de la siguiente

Sistemas de colas: Las llegadas - Distribución de Poisson

- Es una distribución discreta empleada con mucha frecuencia para describir el patrón de las llegadas a un sistema de colas
- Para tasas medias de llegadas pequeñas es asimétrica y se hace más simétrica y se aproxima a la binomial para tasas de llegadas altas

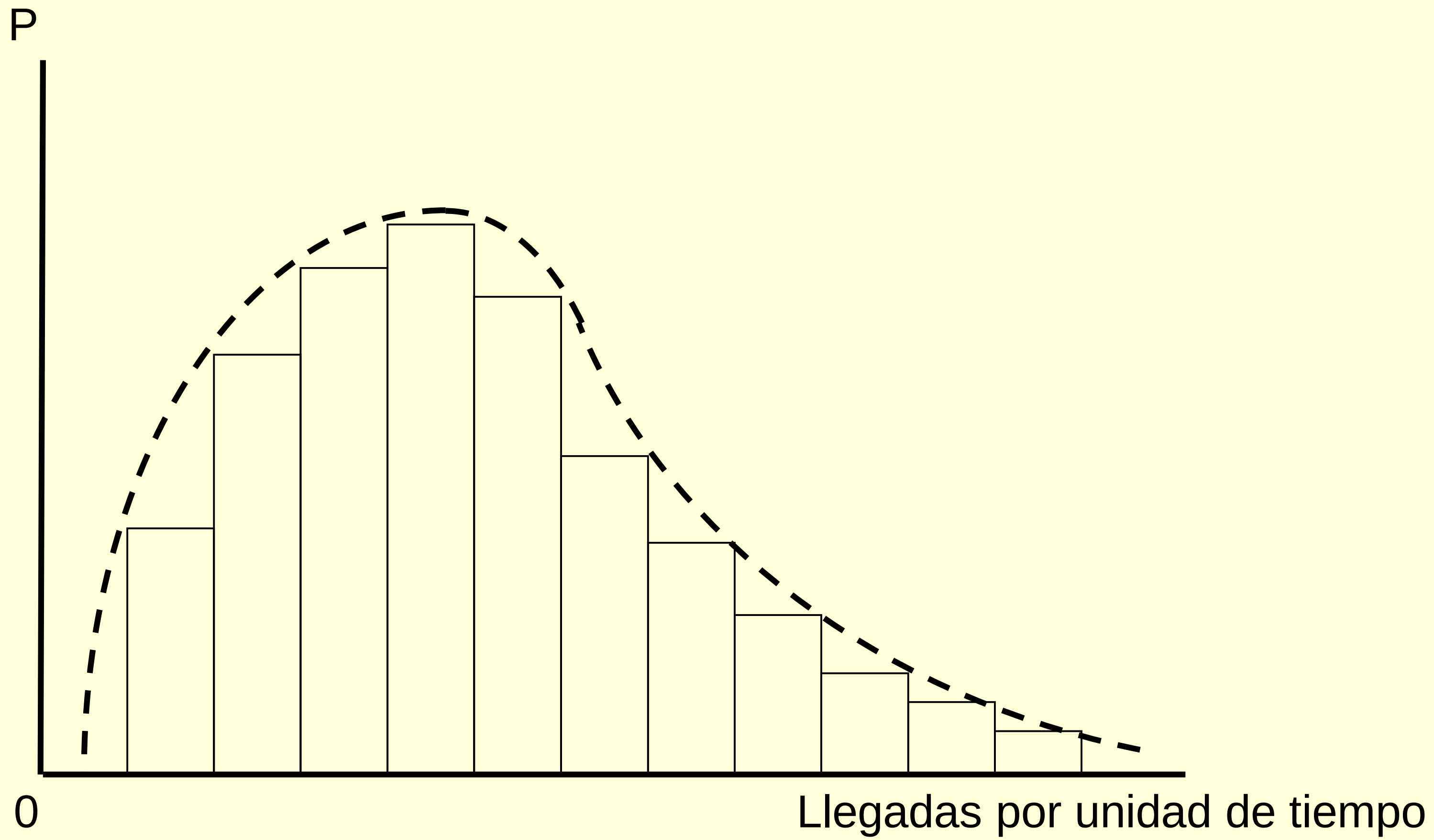
Sistemas de colas: Las llegadas - Distribución de Poisson

- Su forma algebraica es:

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

- Donde:
  - $\boldsymbol{P(k)}$: probabilidad de k llegadas por unidad de tiempo
  - λ : tasa media de llegadas
  - $\boldsymbol{e} = 2{,}7182818...$

Sistemas de colas: Las llegadas - Distribución de Poisson



Sistemas de colas: La cola

- El número de clientes en la cola es el número de clientes que esperan el servicio
- El número de clientes en el sistema es el número de clientes que esperan en la cola más el número de clientes que actualmente reciben el servicio

Sistemas de colas: La cola

- La capacidad de la cola es el número máximo de clientes que pueden estar en la cola
- Generalmente se supone que la cola es infinita
- Aunque también la cola puede ser finita

Sistemas de colas: La cola

- La disciplina de la cola se refiere al orden en que se seleccionan los miembros de la cola para comenzar el servicio
- La más común es PEPS: primero en llegar, primero en servicio
- Puede darse: selección aleatoria, prioridades, UEPS, entre otras.

Sistemas de colas: El servicio

- El servicio puede ser brindado por un servidor o por servidores múltiples
- El tiempo de servicio varía de cliente a cliente
- El tiempo esperado de servicio depende de la tasa media de servicio (μ)

Sistemas de colas: El servicio

- El tiempo esperado de servicio equivale a $1/\mu$
- Por ejemplo, si la tasa media de servicio es de 25 clientes por hora
- Entonces el tiempo esperado de servicio es $1/\mu = 1/25 = 0.04$ horas, o 2.4 minutos

Sistemas de colas: El servicio

- Es necesario seleccionar una distribución de probabilidad para los tiempos de servicio
- Hay dos distribuciones que representarían puntos extremos:
  - La distribución exponencial (σ=media)
  - Tiempos de servicio constantes (σ=0)

Sistemas de colas: El servicio

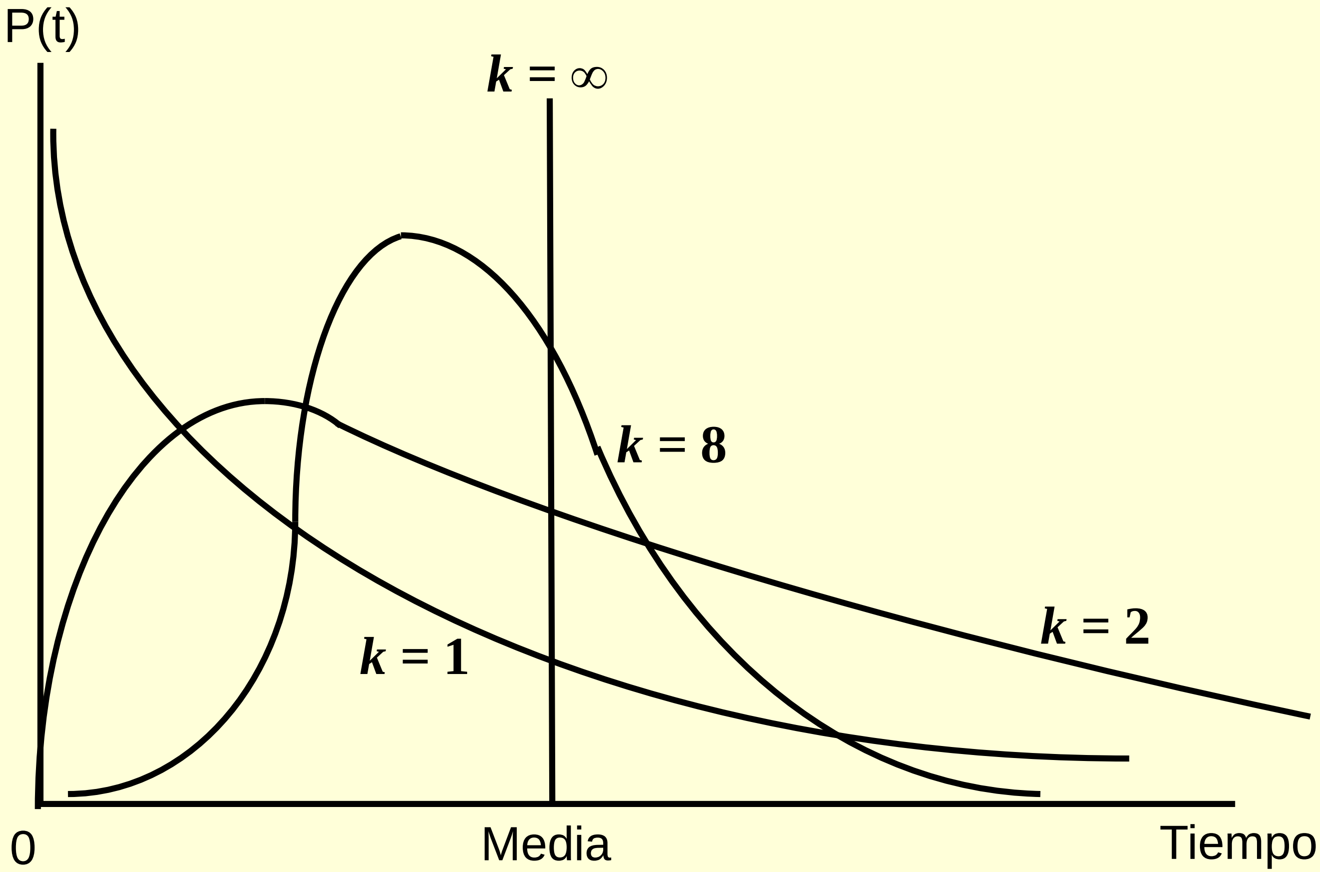
- Una distribución intermedia es la distribución Erlang
- Esta distribución posee un parámetro de forma $\boldsymbol{k}$ que determina su desviación estándar:

$$\sigma = \frac{1}{\sqrt{k}} media$$

Sistemas de colas: El servicio

- Si $k = 1$, entonces la distribución Erlang es igual a la exponencial
- Si $k = \infty$, entonces la distribución Erlang es igual a la distribución degenerada con tiempos constantes
- La forma de la distribución Erlang varía de acuerdo con k

Sistemas de colas: El servicio



http://www.auladeeconomia.com

Sistemas de colas: Distribución Erlang

Distribución	Desviación estándar
Constante	0
Erlang, $\boldsymbol{k}$ = 1	$media$
Erlang, $\boldsymbol{k}$ = 2	$1/\sqrt{2}\, media$
Erlang, $\boldsymbol{k}$ = 4	$1/2\ media$
Erlang, $\boldsymbol{k}$ = 8	$1/\sqrt{8}\, media$
Erlang, $\boldsymbol{k}$ = 16	$1/4\ media$
Erlang, cualquier $\boldsymbol{k}$	$1/\sqrt{k}\ media$

Sistemas de colas: Etiquetas para distintos modelos

Notación de Kendall: $A/B/c$

- A: Distribución de tiempos entre llegadas
- B: Distribución de tiempos de servicio
  - M: distribución exponencial
  - D: distribución degenerada
  - E_k: distribución Erlang
- c: Número de servidores

Estado del sistema de colas

- En principio el sistema está en un estado inicial
- Se supone que el sistema de colas llega a una condición de estado estable (nivel normal de operación)
- Existen otras condiciones anormales (horas pico, etc.)
- Lo que interesa es el estado estable

Desempeño del sistema de colas

- Para evaluar el desempeño se busca conocer dos factores principales:

1. El número de clientes que esperan en la cola
2. El tiempo que los clientes esperan en la cola y en el sistema

Medidas del desempeño del sistema de colas

1. Número esperado de clientes en la cola L_q
2. Número esperado de clientes en el sistema L_s
3. Tiempo esperado de espera en la cola W_q
4. Tiempo esperado de espera en el sistema W_s

Medidas del desempeño del sistema de colas: fórmulas generales

$$W_s = W_q + \frac{1}{\mu}$$

$$L_s = \lambda W_s$$

$$L_q = \lambda W_q$$

$$L_s = L_q + \frac{\lambda}{\mu}$$

Medidas del desempeño del sistema de colas: ejemplo

- Suponga una estación de gasolina a la cual llegan en promedio 45 clientes por hora
- Se tiene capacidad para atender en promedio a 60 clientes por hora
- Se sabe que los clientes esperan en promedio 3 minutos en la cola

Medidas del desempeño del sistema de colas: ejemplo

- La tasa media de llegadas λ es 45 clientes por hora o 45/60 = 0.75 clientes por minuto
- La tasa media de servicio μ es 60 clientes por hora o 60/60 = 1 cliente por minuto

Medidas del desempeño del sistema de colas: ejemplo

$$W_q = 3\,\text{min}$$

$$W_s = W_q + \frac{1}{\mu} = 3 + \frac{1}{1} = 4\,\text{min}$$

$$L_s = \lambda W_s = 0.75 \times 4 = 3\,clientes$$

$$L_q = \lambda W_q = 0.75 \times 3 = 2.25\,clientes$$

Medidas del desempeño del sistema de colas: ejercicio

- Suponga un restaurant de comidas rápidas al cual llegan en promedio 100 clientes por hora
- Se tiene capacidad para atender en promedio a 150 clientes por hora
- Se sabe que los clientes esperan en promedio 2 minutos en la cola
- Calcule las medidas de desempeño del sistema

Probabilidades como medidas del desempeño

- Beneficios:
  - Permiten evaluar escenarios
  - Permite establecer metas
- Notación:
  - P_n : probabilidad de tener n clientes en el sistema
  - $P(W_s \leq t)$: probabilidad de que un cliente no espere en el sistema más de t horas

Factor de utilización del sistema

- Dada la tasa media de llegadas λ y la tasa media de servicio μ, se define el factor de utilización del sistema ρ.
- Generalmente se requiere que $\rho < 1$
- Su fórmula, con un servidor y con s servidores, respectivamente, es:

$$\rho = \frac{\lambda}{\mu} \qquad \rho = \frac{\lambda}{s\mu}$$

Factor de utilización del sistema - ejemplo

- Con base en los datos del ejemplo anterior, $\lambda = 0.75$, $\mu = 1$
- El factor de utilización del sistema si se mantuviera un servidor es

$$\rho = \lambda/\mu = 0.75/1 = 0.75 = 75\%$$

- Con dos servidores ($s = 2$):

$$\rho = \lambda/s\mu = 0.75/(2*1) = 0.75/2 = 37{,}5\%$$

Modelos de una cola y un servidor

- $M/M/1$: Un servidor con llegadas de Poisson y tiempos de servicio exponenciales
- $M/G/1$: Un servidor con tiempos entre llegadas exponenciales y una distribución general de tiempos de servicio
- $M/D/1$: Un servidor con tiempos entre llegadas exponenciales y una distribución degenerada de tiempos de servicio
- $M/E_k/1$: Un servidor con tiempos entre llegadas exponenciales y una distribución Erlang de tiempos de servicio

Modelo M/M/1

$$L_s = \frac{\lambda}{\lambda - \mu} \qquad L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

$$W_s = \frac{1}{\mu - \lambda} \qquad W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

$$P_n = (1 - \rho)\rho^n \qquad P(L_s > n) = \rho^{n+1}$$

$$P(W_s > t) = e^{-\mu(1-\rho)t} \qquad P(W_q > t) = \rho e^{-\mu(1-\rho)t}$$

$$t \geq 0, \rho < 1$$

Modelo M/M/1: ejemplo

- Un lavacar puede atender un auto cada 5 minutos y la tasa media de llegadas es de 9 autos por hora
- Obtenga las medidas de desempeño de acuerdo con el modelo M/M/1
- Además la probabilidad de tener 0 clientes en el sistema, la probabilidad de tener una cola de más de 3 clientes y la probabilidad de esperar más de 30 min. en la cola y en el sistema

Modelo M/M/1: ejemplo

$$\lambda = 9, \mu = 12, \rho = \frac{9}{12} = 0.75$$

$$L_s = \frac{\lambda}{\lambda - \mu} = 3\, clientes \qquad L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = 2.25\, clientes$$

$$W_s = \frac{1}{\mu - \lambda} = 0.33\, hrs = 20 \min$$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = 0.25\, hrs = 15 \min$$

$$P_0 = (1 - \rho)\rho^0 = 0.25 \qquad P(L_s > 3) = \rho^{3+1} = 0.32$$

$$P(W_s > 30/60) = e^{-\mu(1-\rho)t} = 0.22$$

$$P(W_q > 30/60) = \rho e^{-\mu(1-\rho)t} = 0.17$$

http://www.auladeeconomia.com

Modelo M/M/1: ejercicio

- A un supermercado llegan en promedio 80 clientes por hora que son atendidos entre sus 5 cajas.
- Cada caja puede atender en promedio a un cliente cada 3 minutos
- Obtenga las medidas de desempeño de acuerdo con el modelo M/M/1
- Además la probabilidad de tener 2 clientes en el sistema, la probabilidad de tener una cola de más de 4 clientes y la probabilidad de esperar más de 10 min. en la cola

Modelo M/G/1

$$L_s = L_q + \rho \qquad\qquad L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)}$$

$$W_s = W_q + \frac{1}{\mu} \qquad\qquad W_q = \frac{L_q}{\lambda}$$

$$P_0 = 1-\rho \qquad\qquad P_w = \rho$$

$$\rho < 1$$

Modelo M/G/1: ejemplo

- Un lavacar puede atender un auto cada 5 min. y la tasa media de llegadas es de 9 autos/hora, σ = 2 min.
- Obtenga las medidas de desempeño de acuerdo con el modelo M/G/1
- Además la probabilidad de tener 0 clientes en el sistema y la probabilidad de que un cliente tenga que esperar por el servicio

Modelo M/G/1: ejemplo

$$L_s = L_q + \rho = 1.31 + .75 = 2.06\, clientes$$

$$L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)} = 1.31\, clientes$$

$$W_s = W_q + \frac{1}{\mu} = 0.228\, hrs = 13.7 \min$$

$$W_q = \frac{L_q}{\lambda} = 0.145\, hrs = 8.7 \min$$

$$P_0 = 1 - \rho = 0.25 \qquad P_w = \rho = 0.75$$

Modelo M/G/1: ejercicio

- A un supermercado llegan en promedio 80 clientes por hora que son atendidos entre sus 5 cajas.
- Cada caja puede atender en promedio a un cliente cada 3 minutos. Suponga $\sigma = 5$ min
- Obtenga las medidas de desempeño de acuerdo con el modelo M/G/1
- Además la probabilidad de tener 0 clientes en el sistema y la probabilidad de que un cliente tenga que esperar por el servicio

http://www.auladeeconomia.com

Modelo M/D/1

$$L_s = \lambda W_s \qquad L_q = \frac{\rho^2}{2(1-\rho)}$$

$$W_s = W_q + \frac{1}{\mu} \qquad W_q = \frac{L_q}{\lambda}$$

$$\rho < 1$$

Modelo M/D/1: ejemplo

- Un lavacar puede atender un auto cada 5 min.
- La tasa media de llegadas es de 9 autos/hora.
- Obtenga las medidas de desempeño de acuerdo con el modelo M/D/1

Modelo M/D/1: ejemplo

$$L_s = \lambda W_s = 1.875\,clientes$$

$$L_q = \frac{\rho^2}{2(1-\rho)} = 1.125\,clientes$$

$$W_s = W_q + \frac{1}{\mu} = 0.21\,hrs = 12.5\min$$

$$W_q = \frac{L_q}{\lambda} = 0.125\,hrs = 7.5\min$$

Modelo M/D/1: ejercicio

- A un supermercado llegan en promedio 80 clientes por hora que son atendidos entre sus 5 cajas.
- Cada caja puede atender en promedio a un cliente cada 3 minutos.
- Obtenga las medidas de desempeño de acuerdo con el modelo M/D/1

Modelo M/E$_k$/1

$$L_s = \lambda W_s \qquad L_q = \frac{\rho^2 (k+1)}{2k(1-\rho)}$$

$$W_s = W_q + \frac{1}{\mu} \qquad W_q = \frac{L_q}{\lambda}$$

$$\rho < 1$$

Modelo $M/E_k/1$: ejemplo

- Un lavacar puede atender un auto cada 5 min.
- La tasa media de llegadas es de 9 autos/hora. Suponga σ = 3.5 min (aprox.)
- Obtenga las medidas de desempeño de acuerdo con el modelo $M/E_k/1$

Modelo M/E$_k$/1: ejemplo

$$L_s = \lambda W_s = 2.437\,clientes$$

$$L_q = \frac{\rho^2(k+1)}{2k(1-\rho)} = 1.6875\,clientes$$

$$W_s = W_q + \frac{1}{\mu} = 0.2708\,hrs = 16.25\min$$

$$W_q = \frac{L_q}{\lambda} = 0.1875\,hrs = 11.25\min$$

Modelo M/E$_k$/1: ejercicio

- A un supermercado llegan en promedio 80 clientes por hora que son atendidos entre sus 5 cajas.
- Cada caja puede atender en promedio a un cliente cada 3 minutos. Suponga k= 4
- Obtenga las medidas de desempeño de acuerdo con el modelo M/E$_k$/1

Modelos de un servidor: Ejercicio: complete el cuadro ejemplo lavacar

Modelo	L_s	W_s	L_q	W_q
M/M/1				
M/G/1				
M/D/1				
M/E_k/1				

Modelos de varios servidores

- ***M/M*/s**: *s* servidores con llegadas de Poisson y tiempos de servicio exponenciales
- ***M/D*/s**: *s* servidores con tiempos entre llegadas exponenciales y una distribución degenerada de tiempos de servicio
- M/E_k/s: *s* servidores con tiempos entre llegadas exponenciales y una distribución Erlang de tiempos de servicio

M/M/s, una línea de espera

$$P_0 = \frac{1}{\frac{\rho^s}{s!}\left(\frac{s\mu}{s\mu-\lambda}\right)+\sum_{n=0}^{s-1}\frac{\rho^n}{n!}}$$

$$L_q = \frac{\rho^s \lambda\mu}{(s-1)!(s\mu-\lambda)^2}P_0 \qquad L_s = L_q + \frac{\lambda}{\mu} \qquad W_q = \frac{L_q}{\lambda}$$

$$W_s = W_q + \frac{1}{\mu} \qquad P_n = \frac{\rho^n}{n!}P_0, si\ \ n \le k$$

$$P_n = \frac{\rho^n}{s!s^{n-s}}P_0, si\ \ n > k \qquad P_w = \frac{1}{s!}\rho^s\left(\frac{s\mu}{s\mu-\lambda}\right)P_0$$

M/M/s, una línea de espera

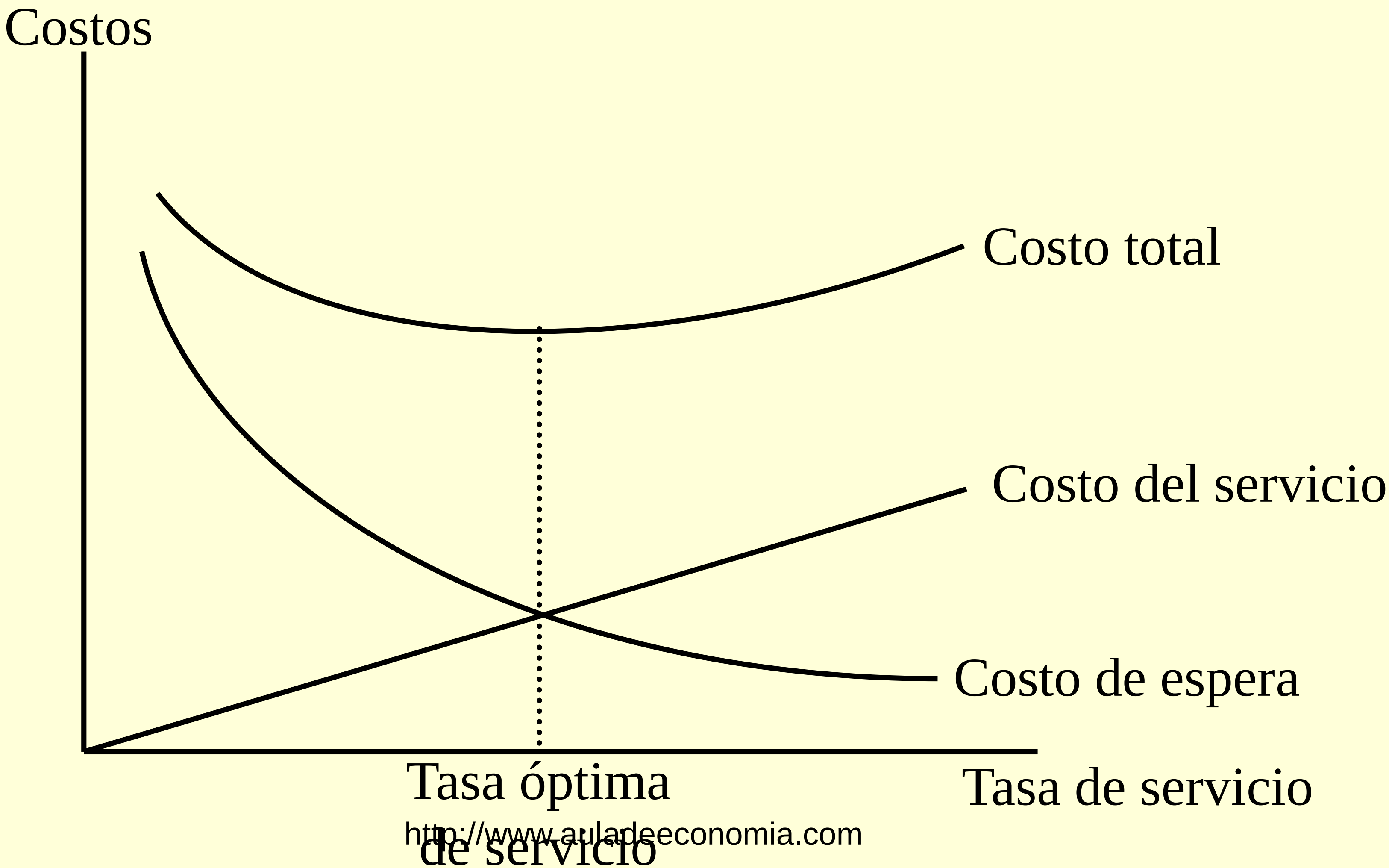
$$Si\ s = 2$$

$$L_q = \frac{\rho^3}{4-\rho^2}$$

$$Si\ s = 3$$

$$L_q = \frac{\rho^4}{(3-\rho)(6-4\rho+\rho^2)}$$

Análisis económico de líneas de espera



http://www.auladeeconomia.com

Si desea más información visite
www.auladeeconomia.com

Le invitamos a leer nuestros artículos y matricular nuestros cursos

AuladeEconomía.com