<u>Resumen Teoría Simulación (Prof. Lara)</u>

Hecho por AndyNOB – http://www.foroutnrosario.com.ar/

<u>Capitulo 1 Law & Kelton (Modelado de Simulación Básico).</u>

Sistema: empresa o proceso de interés a modelizar.

Modelo: representación del sistema en términos de relaciones cuantitativas y lógicas.

Los sistemas se categorizan en dos tipos: discretos o continuos. Un sistema discreto es aquel en el que las variables de estado cambian instantáneamente en puntos separados del tiempo. En un sistema continuo en cambio las variables de estado cambian continuamente con respecto al tiempo.

Modelos de Simulación Estocásticos vs. Determinísticos: Si un modelo de simulación no contiene componentes probabilísticas (es decir aleatorias) se conoce como determinístico, en estos modelos la salida se "determina" una vez que se especifica el conjunto de relaciones (ecuaciones) y los valores de entrada. En cambio los modelos estocásticos contienen variables aleatorias de entrada sujetas a una distribución probabilística de algún tipo.

<u>Simulación de Eventos Discretos:</u>

La simulación de eventos discretos comprende el modelado de un sistema a medida que este evoluciona a través del tiempo por medio de una representación en la cual las variables de estado cambian instantáneamente en puntos separados en el tiempo. Estos puntos en el tiempo son aquellos en los cuales un evento ocurre, donde un evento se define como una ocurrencia instantánea que puede cambiar el estado del sistema.

<u>Mecanismo de Avance del Tiempo:</u> Llamamos Reloj de la simulación a la variable de un modelo de simulación que contiene el valor actual del tiempo simulado. La unidad del reloj nunca se enuncia explícitamente y se asume que está en las mismas unidades que los parámetros de entrada.

Existen dos enfoques para el mecanismo de avance del tiempo:

- **Avance del tiempo al siguiente evento**.
  Con este enfoque el reloj de la simulación se inicializa a cero y se determinan los tiempos de ocurrencia de eventos futuros, luego el reloj se avanza al tiempo de ocurrencia del evento futuro más próximo, en este punto el estado del sistema se actualiza para determinar que un evento ha ocurrido y los tiempos de futuros eventos también se actualizan.
  Este proceso continua hasta que se cumple con una condición de parada pre especificada.
- **Avance del tiempo a incrementos fijos.**
  La diferencia con el método anterior es que este enfoque no saltea periodos de inactividad en el sistema, lo que supone una mayor cantidad de cómputo.

Simulación de un Sistema de Colas de un solo Servidor (M/M/1):

En un sistema de colas de un solo servidor, los tiempos entre arribos A_1, A_2, … , A_n (de cada cliente al sistema) son variables aleatorias IID (independientes e idénticamente distribuidas). Un cliente que arriba y encuentra al servidor desocupado se atiende inmediatamente, y los tiempos de servicio S_1, S_2, … , S_n (de cada cliente) son también variables aleatorias IID independientes de los tiempos de arribo. Si un cliente arriba y encuentra al servidor ocupado se une al final de cola. Al producirse una partida (un cliente completa el servicio) el servidor elige un cliente de la cola según la disciplina FIFO. El sistema se simula hasta que un número fijo (n) de clientes hayan completados sus demoras en cola, es decir cuando el n-esimo cliente entre en servicio.

Medidas de Rendimiento: Para medir el rendimiento de este sistema observamos las estimaciones de tres parámetros (más un parámetro opcional que es *w(n)*):

- Demora promedio esperada en cola de los n clientes. Llamada *d(n).*
- Número de clientes promedio esperado en la cola. Denotado por *q(n).*
- Utilización del servidor. Denominada *u(n).*
- Demora promedio esperada en el sistema de los n clientes. Llamada *w(n)*.

Demora promedio esperada en cola de los "n-clientes":

Lo que queremos estimar d(n) es el valor esperado para esta variable aleatoria. d(n) es el promedio de una gran numero de demoras promedio de n clientes. A partir de una sola corrida de la simulación podemos estimar este parámetro a través de:

$$\hat{d}(n) = \frac{\sum_{i=1}^{n} D_i}{n}$$

Esta fórmula es el promedio de las n demoras D_i que fueron obtenidas durante la simulación.

Este estimador está basado en una muestra de tamaño 1 ya que estamos haciendo solamente una sola corrida de la simulación. Un estimador de este tipo no tendrá demasiada precisión, pues el sistema seguramente se encuentra en estado *transitorio*.

Número de clientes promedio esperado en la cola (pero no siendo atendidos):

Este promedio se toma sobre el periodo de tiempo necesarios para observar las n demoras que definen nuestra regla de parada. Esta es una clase diferente de promedio que el anterior, ya que se toma sobre el tiempo (continuo) en lugar de los clientes (discreto).

Definimos Q(t) al número de clientes en cola en el momento t (para cualquier $t \geq 0$) y T(n) al tiempo requerido para observar n demoras en cola. Si llamamos p_i a la proporción esperada (entre 0 y 1) del tiempo en que Q(t) es igual a i, una definición de q(n) sería:

$$q(n) = \sum_{i=0}^{\infty} i p_i$$

Para estimar q(n) en una simulación, simplemente reemplazamos los p_i's con sus respectivas estimaciones y obtenemos:

$$\hat{q}(n) = \sum_{i=0}^{\infty} i \hat{p}_i$$

Donde $\hat{p}_i$ es la proporción observada del tiempo en que hubo i clientes en la cola (en la simulación).

Sin embargo una manera más sencilla de obtener $\hat{q}(n)$ es mediante algunas consideraciones geométricas, llamamos T_i al tiempo total durante la simulación en que la cola es de tamaño i, luego:

$T(n) = \sum_{i=0}^{n} T_i = T_0 + T_1 + T_2 + \ldots + T_n$ y $P_i = T_i / T(n)$

Y el estimador puede escribirse como:

$$\hat{q}(n) = \frac{\sum_{i=0}^{\infty} i T_i}{T(n)}$$

Sin embargo la sumatoria en el numerador de la ecuación anterior es solo el área bajo la curva de Q(t), lo cual puede escribirse como una integral de 0 hasta T(n), quedando finalmente la expresión:

$$\hat{q}(n) = \frac{\int_0^{T(n)} Q(t)\,dt}{T(n)}$$

<u>Utilización esperada del servidor:</u>

La utilización esperada del servidor es la proporción esperada de tiempo durante la simulación en que el servidor está ocupado y por eso es un número entre 0 y 1.

El estimador $\hat{u}(n)$ será la proporción observada de tiempo durante la simulación en que el servidor está ocupado. Para esto definimos la “función ocupado” B(t).

$$B(t) \begin{cases} 1 & \text{si el servidor está ocupado en el tiempo t.} \\ 0 & \text{si el servidor está desocupado en el tiempo t.} \end{cases}$$

De esta manera $\hat{u}(n)$ puede expresarse como la proporción de tiempo en que B(t) es igual a 1.

$$\hat{u}(n) = \frac{\int_0^{T(n)} B(t)\,dt}{T(n)}$$

$\hat{u}(n)$ es el área bajo la función B(t).

<u>Demora o Tiempo de espera promedio esperado en el sistema (cola + servidor):</u>

Esta medida se define como el intervalo de tiempo desde el instante que un cliente arriba a la cola hasta el instante en que el cliente completa el servicio y parte.

El estimador usual de w(n) sería:

$$\widehat{w}(n) = \frac{\sum_{i=1}^{n} D_i}{n} + \frac{\sum_{i=1}^{n} S_i}{n} = \hat{d}(n) + \bar{S}(n)$$

Donde S_i es el tiempo de espera de los n clientes en el servidor y $\bar{S}(n)$ es el promedio de los n tiempos de servicio de los clientes. Ya que el tiempo de servicio medio o esperado E(S) es conocido un estimador alternativo sería:

$$\widetilde{w}(n) = \hat{d}(n) + E(S)$$

En casi todas las simulaciones de colas $\widetilde{w}(n)$ será mejor que $\widehat{w}(n)$. Ambos son estimadores no sesgados.

Capitulo 9 Law & Kelton (Análisis de Datos de Salida).

Existen varias razones por las cuales el análisis de datos de salida de una simulación no pueden ser tratados en forma apropiada. Primero debe considerarse que una simulación es un experimento de muestreo estadístico basado en computadoras, por lo tanto deben usarse las técnicas estadísticas apropiadas para diseñar y analizar los experimentos de simulación. Una segunda razón para análisis estadísticos inadecuados es que los procesos de salida de virtualmente todas las simulaciones son no estacionarios y auto correlacionados. Por este motivo las técnicas estadísticas clásicas basadas en observaciones IID no son aplicables directamente.

Llamamos Y_1, Y_2, ... a un proceso estocástico de salida a partir de una sola corrida de simulación. Los Y_i son variables aleatorias que en general no son IID.

Llamamos $\mathbf{y_{ij}}$ a una observación de la variable $\mathbf{Y_j}$ en la i-ésima corrida o réplica. Si corremos la simulación con un conjunto de números aleatorios diferentes obtendremos distintos valores de $\mathbf{y_{ij}}$

Suponga que hacemos n corridas de la simulación independientes (con distintos números aleatorios en cada corrida) de tamaño m, resultando en las observaciones:

$$y_{11}, y_{12}, \ldots, y_{1m}$$

$$y_{21}, y_{22}, \ldots, y_{2m}$$

$$y_{n1}, y_{n2}, \ldots, y_{nm}$$

Las observaciones de una fila son no IID, además representan los valores de las distintas variables $\mathbf{Y_j}$ en la corrida "i". Las observaciones de una columna en cambio si son IID, y representan los distintos valores que asume una única variable $\mathbf{Y_j}$ a través de las m corridas "i".

Esta independencia a través de las corridas es la clave para los relativamente simples métodos de análisis de datos de salida, el objetivo de este análisis es usar las observaciones $\mathbf{y_{ji}}$ (i= 1,2,...,m ; j=1,2,...n) para trazar inferencias acerca de las distribuciones de las variables aleatorias $\mathbf{Y_j}$.

Análisis Estadístico para Simulaciones Terminales:

Estimación de Medias:

Suponga que nos interesaría obtener una estimación puntual y el intervalo de confianza para la media $\mu = E(X)$ donde X es una variable aleatoria definida sobre una corrida. Si se hacen n corridas independientes de la simulación y siendo X_1, X_2, ... , X_n las variables aleatorias IID resultantes obtenemos que $\bar{X}(n)$ es un estimador puntual no sesgado para μ y un intervalo de confianza aprox del 100(1-α)% con (0<α<1) para μ está dado por:

$$\bar{X}(n) \pm t_{n-1\,;\,1-\frac{\alpha}{2}} \sqrt{\frac{S^2(n)}{n}}$$

Dicho intervalo de confianza se denomina **procedimiento de muestra de tamaño fijo**. La exactitud de este intervalo depende de la suposición de que los $\mathbf{X_j}$ son variables aleatorias normales.

Obtención de una precisión especificada: Una desventaja del procedimiento anterior es que en analista no tiene control sobre la mitad del tamaño del intervalo de confianza (es decir la precisión de $\bar{X}(n)$), por este motivo existen dos métodos que estiman el número de corridas requeridas para estimar la media $\mu = E(X)$ con una precisión o error especificado.

Estos métodos son: el número de corridas para establecer un error absoluto β y el número de corridas para establecer un error relativo γ.

Comenzamos definiendo el primero, si la estimación $\bar{X}$ es tal que $|\bar{X} - \mu| = \beta$, entonces decimos que $\bar{X}$ tiene un error absoluto de β. Si hacemos corridas de una simulación hasta que la mitad del tamaño del intervalo de confianza de 100(1-α)% es menor o igual que β, entonces:

$$1 - \alpha \approx P\left(\bar{X} - \frac{1}{2}tam \le \mu \le \bar{X} + \frac{1}{2}tam\right)^{(*)}$$

$$= P\left(|\bar{X} - \mu| \le \frac{1}{2}tam\right)$$

$$\le P(|\bar{X} - \mu| \le \beta)$$

(*) ½ tam = mitad de tamaño del intervalo de confianza.

$\bar{X}$ tiene un error absoluto de a lo sumo β con una probabilidad de aprox. 1 – α. Si hemos construido un intervalo de confianza para μ basado en un número fijo de corridas n y si asumimos que nuestra estimación $S^2(n)$ de la varianza poblacional no cambiara a medida que el número de corridas se incrementa, entonces obtenemos la siguiente expresión:

$$n_a^*(\beta) = \min\left\{ i \ge n : t_{n-1\,;\,1-\frac{\alpha}{2}} \sqrt{\frac{S^2(n)}{n}} \le \beta \right\}$$

La formula obtenida indica el número total de corridas $n_a^*(\beta)$ que se necesitan para obtener un error absoluto de β. Los dos puntos se leen como "tal que", dicha fórmula se itera para valores de i≥1 hasta obtener un valor de i para el cual el lado a continuación de los 2 puntos de la expresión sea válido. La precisión de la formula depende de cuan cercana es $S^2(n)$ de la Var(x).

Definimos ahora la segunda fórmula, si el valor de $\bar{X}$ estimado es tal que $\left|\frac{\bar{X}-\mu}{\mu}\right| = \gamma$, entonces decimos que $\bar{X}$ tiene un error relativo de γ. Si hacemos corridas de una simulación hasta que la mitad del tamaño del intervalo de confianza de 100(1-α)% divido por $\bar{X}$ es menor o igual que γ, entonces:

$$1 - \alpha \approx P\left(\left|\frac{\bar{X} - \mu}{\mu}\right| \le \gamma\right)$$

$$\le P(|\bar{X} - \mu| \le \gamma|\bar{X}|)$$

$$= P(|\bar{X} - \mu| \le \gamma|\bar{X} - \mu + \mu|)$$

$$\le P(|\bar{X} - \mu| \le \gamma(|\bar{X} - \mu| + |\mu|))$$

$$= P((1 - \gamma)|\bar{X} - \mu| \le \gamma|\mu|)$$

$$= P\left(\left|\frac{\bar{X}-\mu}{\mu}\right| \le \frac{\gamma}{1-\gamma}\right)$$

Así, $\bar{X}$ tiene un error relativo de a lo sumo $\left|\frac{\bar{X}-\mu}{\mu}\right|$ con una probabilidad de aprox. 1-α. Si hemos construido un intervalo de confianza para μ basado en un número fijo de corridas n y si asumimos que nuestras estimaciones de la media y varianza poblacional no cambiarán a medida que el número de corridas se incrementa, entonces obtenemos la siguiente expresión:

$$n_r^*(\gamma) = \min\left\{ i \geq n : \frac{t_{n-1\,;\,1-\alpha/2}\sqrt{\frac{S^2(n)}{i}}}{|\bar{X}(n)|} \leq \gamma' \right\}$$

Esta fórmula indica el número total de corridas $n_r^*(\gamma)$ necesarias para obtener un error relativo de γ. Donde $\gamma' = \frac{\gamma}{1+\gamma}$ es el error relativo ajustado necesario para obtener un error relativo real de γ. Como en la fórmula anterior los dos puntos indica que esta fórmula se itera para valores de i≥1 hasta obtener un valor de i para el cual el lado a continuación de los 2 puntos de la expresión sea válido.

Resumen:

Construcción de Intervalos de Confianza

Si la población es Normal:

Procedimiento de muestra de tamaño fijo.
Sirve para estimar un intervalo de confianza de la media poblacional $\mu = E(X)$.

Si la población no es Normal:

Número de corridas para un error absoluto β.
Sirve para estimar el número de corridas n_a suponiendo que la varianza estimada no cambiará a medida que el número de corridas se incrementa.

Número de corridas para un error relativo γ.
Sirve para estimar el número de corridas n_r suponiendo que la media y varianza estimadas no cambiarán a medida que el número de corridas se incrementa.

Modelo Analítico para una cola M/M/1 (Mc Millan - González).

Tipos de sistemas de colas:

Un sistema de colas se distingue de otro por cierto número de atributos. Los principales son:

1. El número de fases.
2. El número de canales.
3. La disciplina de las colas.

La disciplina de las colas se refiere al hecho de si los clientes se acomodan de acuerdo con una norma de servicio por orden de llegada (FIFO, LIFO, etc.) o se aplica alguna otra regla de prioridad especial.
Entendemos por clientes a entidades cuyas llegadas ejercen demandas sobre alguna instalación (servidor). El patrón de llegadas depende de: a) el tamaño del universo de clientes posibles (que "genera" clientes que necesitan atención) y b) el nivel de sus actividades, que hace que necesiten servicios de vez en cuando.

Caso M/M/1:

En este caso se supone que el tiempo entre llegadas tiene una distribución exponencialmente negativa y que el tiempo de servicio tiene el mismo tipo de distribución. Puesto que la misma es un producto del proceso de Poisson, nuestro sistema de colas será totalmente de Poisson.

Nos interesa desarrollar un modelo para predecir (analíticamente):

1. La probabilidad de varios números de clientes en la cola. (También llamado *número de clientes promedio esperado en la cola q(n)* en Law-Kelton).
2. El tiempo esperado o promedio que pasará un cliente en las instalaciones de servicio.
3. La probabilidad de que las instalaciones de servicio estén ociosas. (También llamado *factor de utilización del servidor [1-u(n)]* en Law-Kelton).

Para empezar suponemos que nuestro sistema puede dar atender (dar servicio) a μ clientes por unidad de tiempo (en promedio). De esta manera μ también es el número esperado de salidas (partidas) del sistema durante cada unidad de tiempo. Llamamos al promedio de llegadas por unidad de tiempo λ.

A continuación consideramos que t es un momento en el tiempo y que $p_n(t)$ es la probabilidad de que haya n clientes en el sistema en el momento t. Si a su vez consideramos una porción de tiempo después de t denominada $(t + \Delta t)$ podríamos pensar que Δt es tan pequeño que aún cuando exista una llegada o una partida durante el intervalo Δt, es imposible más de una llegada o salida durante ese intervalo (modelo de Poisson). Para enfrentarnos a este problema nos preguntamos cuál es la probabilidad de que haya n clientes en el sistema en el intervalo $(t + \Delta t)$. O sea $p_n(t + \Delta t)$. Ahora bien, n clientes en el sistema durante ese intervalo puede presentarse de cuatro modos distintos:

Modo 1:
Tener n clientes en el sistema en el tiempo t, cero llegadas y cero salidas durante el intervalo Δt.
Esto se puede obtener a partir del siguiente razonamiento:
La probabilidad de 1 llegada es: $\mathrm{p}_1 = \lambda\Delta \mathrm{t}$
La probabilidad de 0 llegada es: $\mathrm{p}_0 = 1 - \mathrm{p}_1 = (1 - \lambda\Delta \mathrm{t})$
La probabilidad de 1 partida es: $\mathrm{p}_1 = \mu\Delta \mathrm{t}$
La probabilidad de 0 partida es: $\mathrm{p}_0 = 1 - \mathrm{p}_1 = (1 - \mu\Delta \mathrm{t})$
Entonces la probabilidad del modo uno es: $\mathrm{p}_\mathrm{n}(\mathrm{t})(1 - \lambda\Delta \mathrm{t})(1 - \mu\Delta \mathrm{t})$

Modo 2:
Tener n-1 clientes en el sistema en el tiempo t, una llegada y cero salidas durante el intervalo Δt.
Probabilidad del modo dos: $\mathrm{p}_{\mathrm{n}-1}(\mathrm{t})(\lambda\Delta \mathrm{t})(1 - \mu\Delta \mathrm{t})$

Modo 3:
Tener n+1 clientes en el sistema en el tiempo *t*, cero llegadas y una salida durante el intervalo Δt.
Probabilidad del modo tres: $p_{n+1}(t)(1-\lambda\Delta t)(\mu\Delta t)$

Modo 4:
Tener n clientes en el sistema en el tiempo *t*, una llegada y una salida durante el intervalo Δt.
Probabilidad del modo cuatro: $p_n(t)(\lambda\Delta t)(\mu\Delta t)$

La probabilidad total de tener n clientes en el sistema en el momento t+Δt es la suma de las probabilidades de los cuatro modos anteriores, quedando la expresión como:

$$\mathbf{p_n(t+\Delta t) = p_n(t)(1-\lambda\Delta t)(1-\mu\Delta t) + p_{n-1}(t)(\lambda\Delta t)(1-\mu\Delta t) + p_{n+1}(t)(1-\lambda\Delta t)(\mu\Delta t) + p_n(t)(\lambda\Delta t)(\mu\Delta t)}$$

Ver en carpeta la demostración.

$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} = p_n(t)(-\lambda-\mu) + \lambda p_{n-1}(t) + \mu p_{n+1}(t)$$

Si hacemos que en el límite de la expresión anterior Δt tienda a 0, la misma se transforma en una ecuación diferencial:

$$\boxed{\frac{dp_n(t)}{dt} = -(\lambda+\mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t)}$$

La expresión anterior hace referencia a un sistema de infinitas ecuaciones diferenciales con n+1ecuaciones y n+1 incógnitas, llamado *cadena de Markov*, en la cual la probabilidad de que un evento pueda ocurrir depende del evento anterior. El caso anterior se aplica a n≥1 si hacemos el caso especial en que n=0, entonces $p_{n-1}(t) = 0$, y la expresión se transforma en:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) - \mu p_0(t) + \mu p_1(t)$$

Dado a que cuando n=0 el sistema está vacío y no se producen salidas, el segundo término se desecha y finalmente:

$$\boxed{\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t)}$$

Finalmente obtenemos la expresión que nos dice que las probabilidades de que haya distintos números de clientes en el sistema cambian con el tiempo.
Mientras el sistema se está asentando para llegar a una condición estable decimos que se encuentra en estado transitorio, una vez alcanzado su condición estable se encuentra en estado estacionario.
La condición de régimen estacionario es que la probabilidad de que haya varios números de clientes en el sistema sea cero, o sea: $\frac{dp_n(t)}{dt} = 0$ quedando finalmente las expresiones:

$$\boxed{(\lambda+\mu)p_n = \lambda p_{n-1} + \mu p_{n+1}}$$

$$\boxed{\lambda p_0 = \mu p_1}$$

A partir de estas ecuaciones resulta evidente que: $p_1 = \frac{\lambda}{\mu} p_0$
Si n=1 en la primera expresión entonces:

$$(\lambda+\mu)p_1 = \lambda p_0 + \mu p_2$$

$$0 = \lambda p_0 - (\lambda + \mu)p_1 + \mu p_2$$

$$p_2 = -\frac{\lambda}{\mu}p_0 + \frac{\lambda + \mu}{\mu}p_1$$

$$p_2 = -\frac{\lambda}{\mu}p_0 + \left(\frac{\lambda + \mu}{\mu}\right)\left(\frac{\lambda}{\mu}p_0\right)$$

$$p_2 = -\frac{\lambda}{\mu}p_0 + \left(\frac{\lambda}{\mu}\right)^2 p_0 + \frac{\lambda}{\mu}p_0$$

$$p_2 = \left(\frac{\lambda}{\mu}\right)^2 p_0$$

Para n=2 siguiendo los mismos pasos obtenemos:

$$p_3 = \left(\frac{\lambda}{\mu}\right)^3 p_0$$

Por ende para el n-esimo término la *probabilidad de que haya n elementos en el sistema* resulta:

$$\boxed{p_n = \left(\frac{\lambda}{\mu}\right)^n p_0}$$

Ahora bien si queremos determinar p_0 tenemos que: $p_0 + p_1 + p_2 + p_3 + \cdots + p_n = 1$

Por ende: $p_0 + \left(\frac{\lambda}{\mu}\right)p_0 + \left(\frac{\lambda}{\mu}\right)^2 p_0 + \left(\frac{\lambda}{\mu}\right)^3 p_0 + \cdots + \left(\frac{\lambda}{\mu}\right)^n p_0 = 1$

$$p_0 = \frac{1}{1 + \left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3 + \cdots + \left(\frac{\lambda}{\mu}\right)^n}$$

El denominador de la expresión anterior es una serie geométrica donde a=1 y r= λ/μ, su suma es: $\frac{a}{1-r}$
Finalmente: $p_0 = 1 - \lambda/\mu$

Por lo tanto la expresión final de p_n sería: $\boxed{p_n = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n}$

Este es un modelo general para determinar la probabilidad de que haya n clientes en el sistema de colas de canal simple, en el estado estacionario, donde el ritmo de llegadas (λ) es menor que el índice medio de servicio (μ). Sin embargo nuestro modelo está limitado al hecho de que **λ/μ<1** (esto es una condición para que exista la solución estacionaria). Desde el punto de vista lógico resulta obvia esa condición ya que si **λ/μ≥1** entonces el índice de llegadas siempre será más alto que la capacidad del servidor de atender esas llegadas, con lo cual la cantidad de gente en cola se haría cada vez más grande a medida que pasa el tiempo y la probabilidad de que haya n elementos en el sistema tendería hacia el infinito a medida que n aumenta.

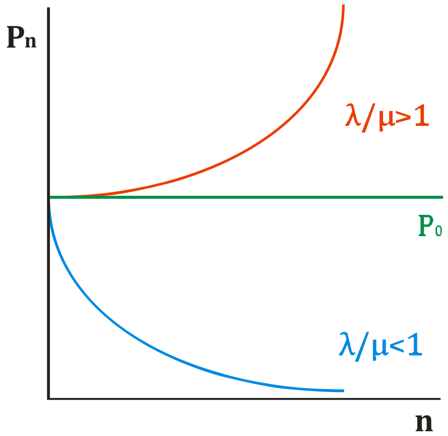

Fig. 1. El estado más probable es que haya 0 clientes en cola si $\lambda/\mu<1$.

Medidas de rendimiento:

A partir del análisis se pueden obtener las siguientes medidas de rendimiento:

Porcentaje de tiempo y ocioso y porcentaje de utilización: λ/μ.

Número de clientes esperado en la cola:

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

Tiempo esperado en el sistema (en la cola): $\frac{1}{\mu-\lambda}$