Procesos de LLegadas

Jhon Jairo Padilla A., PhD.

Introducción

- En teletráfico es importante describir los procesos de llegadas en situaciones como:
  - Las llegadas de llamadas en una central telefónica
  - Las llegadas de paquetes a un enlace de salida de un Router
  - etc.
- Estos procesos se describen matemáticamente mediante Procesos estocásticos puntuales.
- Para que haya un proceso puntual, se debe poder distinguir una llegada de otra.
- La teoría matemática de procesos puntuales fue desarrollada por Conny Palm durante los años 40's.
- Esta teoría fue refinada posteriormente por Khintchine (1968), y ha sido ampliamente aplicada en muchos campos.

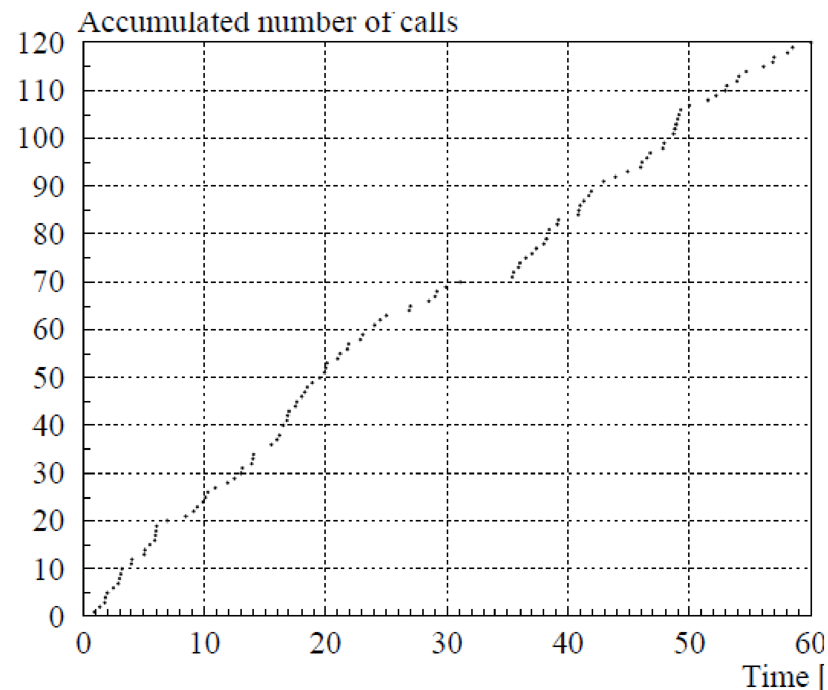
Descripción de los procesos puntuales

- Ahora consideraremos los procesos puntuales simples.
- Se excluyen llegadas múltiples (sólo hay una llegada a la vez).
- En llamadas telefónicas esto se puede obtener eligiendo una escala de tiempo lo suficientemente detallada.
- Considérense tiempos de llegada donde la llamada i-ésima llega en el tiempo i (la primera observación tiene lugar en el tiempo T0=0):

$$0 = T_0 < T_1 < T_2 < \ldots < T_i < T_{i+1} < \ldots .$$

Proceso de llegadas puntual

- Supóngase la v.a. N_t que cuenta el número de llamadas que llegan a una central telefónica en el intervalo semiabierto [0,t)
- N_t será una v.a. con parámetros de tiempo contínuo y espacio discreto.
- Cuanto t se incrementa, N_t nunca decrementa.



Proceso de llegadas puntual

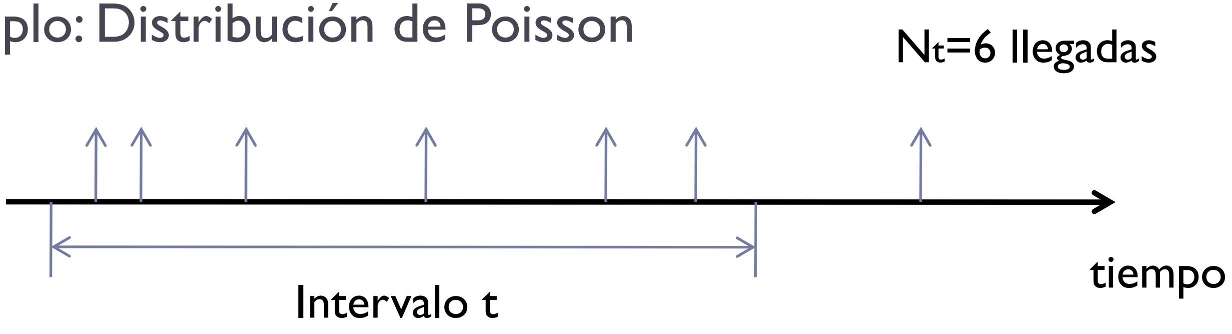
- La distancia en tiempo entre dos llegadas sucesivas es,

$$X_i = T_i - T_{i-1}, \qquad i = 1, 2, \ldots .$$

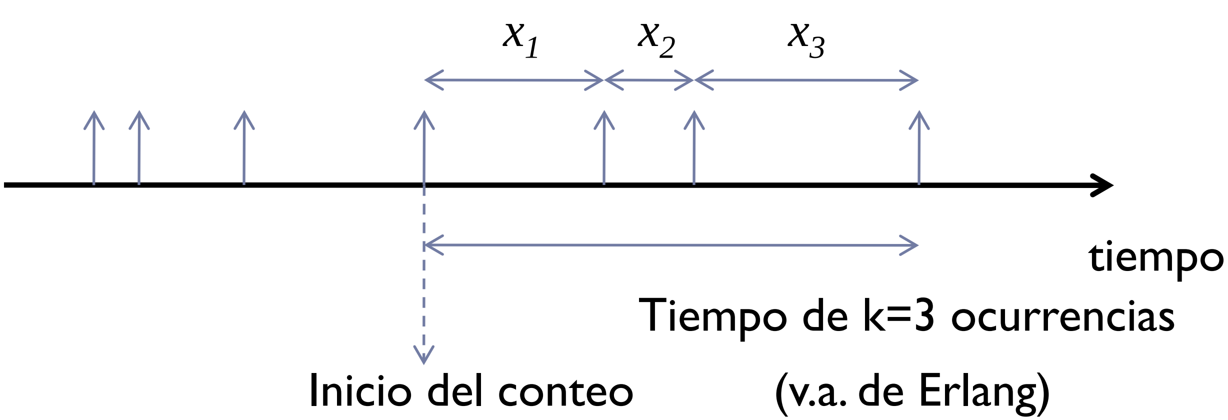
- Este es llamado Tiempo entre-llegadas. La distribución de este intervalo es llamada Distribución del tiempo entre-llegadas.
- Un proceso puntual puede ser caracterizado de dos maneras:
  - **Representación por el Número N_t:** Se mantiene el intervalo t constante y se observa la variable aleatoria N_t que cuenta el número de llamadas en el tiempo t.
  - **Representación por Intervalo T_i:** El número de llegadas de llamadas se mantiene constante y se observa la v.a. T_i para el intervalo hasta que ha habido n llegadas (T_i=X_i)

Proceso de llegadas puntual

- Representación de Número:
  - Ejemplo: Distribución de Poisson



- Representación de tiempo:
  - Ejemplo: Distribución de Erlang



Proceso de llegadas puntual

- La relación entre las dos representaciones está dada por,
- $N_t < n$ si y solo si $T_n = \sum_{i=1}^{n} X_i \geq t\,, \quad n = 1, 2, \ldots$

- Esto se expresa también por la identidad de Feller-Jensen:

$$p\left\{N_t < n\right\} = p\left\{T_n \geq t\right\}, \qquad n = 1, 2, \ldots$$

Proceso de llegadas puntual

- El análisis de un proceso puntual puede estar basado en ambas representaciones. Ellas son equivalentes en principio.
- La *representación de Intervalo* corresponde al análisis de series de tiempo usual. P.ej. si hacemos $i=1$, obtendremos promedios por llamada.
- La *representación de número* no tiene paralelo en las series de análisis de tiempo. Las estadísticas obtenidas son promedios sobre tiempo y obtendremos medias de tiempo. Es decir, estadísticas por unidad de tiempo.
- Las estadísticas de interés nuestro cuando se estudian procesos puntuales se pueden clasificar de acuerdo a estas dos representaciones

Propiedades básicas de la representación por número

- Hay tres propiedades que son de interés teórico:
  1. El número total de llegadas en el intervalo $[t_1,t_2)$ es igual a N_{t2}-N_{t1}.
     - El número medio de llamadas en dicho intervalo se conoce como la *función de renovación, H*:

$$H\left(t_1, t_2\right) = E\left\{N_{t_2} - N_{t_1}\right\}$$

  2. La *densidad de llamadas entrantes* en el tiempo t es:

$$\lambda_t = \lim_{\Delta t \to 0} \frac{N_{t+\Delta t} - N_t}{\Delta t} = N_t'$$

Esta variable mide la tasa de llegada de llamadas en un tiempo t.

  3. *Índice de dispersión de conteos (IDC):* Describe las variaciones de los procesos de llegada durante un intervalo de tiempo t y se define como:

$$IDC = \frac{\mathrm{Var}\{N_t\}}{\mathrm{E}\{N_t\}}$$

Propiedades básicas de la representación por Intervalo

- La distribución acumulada de los intervalos Ti=Xi es,

$$F_i(t) \;=\; p\{X_i \leq t\}$$

$$E\{X_i\} \;=\; m_{1,i}\,.$$

- La distribución V(t) describe el intervalo de tiempo que transcurre hasta que ocurre la próxima llegada
- Índice de dispersión por Intervalos (IDI): Describe las propiedades de segundo orden (variación) para la representación del intervalo. Se define como,

$$IDI = \frac{\mathrm{Var}\{X_i\}}{\mathrm{E}\{X_i\}^2}$$

  - Donde Xi es el tiempo entre-llegadas.
  - IDI es igual al factor de forma menos uno.
  - IDI es más difícil de obtener a partir de las observaciones que el IDC. Además es más sensitivo a la aproximación de las mediciones.
  - Las aplicaciones de datos son más adecuadas para la observación del IDC, mientras que se complica la observación del IDI.

Ejemplos de Aplicación de las representaciones

- Medidas Pasivas:
  - Los equipos de medición graban a intervalos de tiempo regulares el número de llegadas desde la última grabación. Esto corresponde a la *representación por número*, donde el intervalo es fijo.
- Mediciones Activas:
  - El equipo de medición graba un evento en el instante en que éste tiene lugar. Se mantiene el número de eventos fijo y se observa el intervalo que tomó la medición. Esto corresponde a la *representación de Intervalo* y se obtienen estadísticas por cada llamada.

Ejemplos de Aplicación de las representaciones

- La investigación de la calidad del tráfico puede hacerse de dos maneras en la práctica:
  - La Calidad del tráfico es estimada recolectando estadísticas de los resultados de las llamadas de prueba hechas a suscriptores específicos. Las llamadas de prueba son generadas durante la hora pico independientemente del tráfico real. Los equipos de prueba graban el número de llamadas bloqueadas, etc. Las características obtenidas corresponden a tiempos medios de las mediciones de rendimiento. Desafortunadamente, este método incrementa la carga ofrecida al sistema. Teóricamente, las mediciones de rendimiento diferirán de los valores reales.
  - Los equipos de prueba recolectan datos del número de llamadas (N, 2N, 3N,..., donde por ejemplo N=1000). El proceso de tráfico no es cambiado, y las estadísticas de rendimiento arrojan la media de llamadas. (*Representación por número*)

Ejemplos de Aplicación de las representaciones

- Un suscriptor evalúa la calidad por la fracción de llamadas que son bloqueadas. (*Representación por número*).
- El operador evalúa la calidad por la proporción de tiempo que las troncales están ocupadas. (*Representación por Intervalo*).

Propiedades de los procesos puntuales

- Para aplicaciones específicas, debemos introducir las propiedades siguientes.
- Se considerará la representación de número, pero se pueden hacer las mismas consideraciones con la representación de Intervalo.

Propiedades de los procesos puntuales

- Estacionariedad (Homogeneidad en el tiempo):
  - Esta propiedad puede ser descrita como que no interesa la posición en el eje del tiempo, las distribuciones de probabilidad que describen el proceso puntual son independientes del instante de tiempo. Es decir,

$$p\{N_{t_1+t_2} - N_{t_1} = k\} = p\{N_{t_1+t_2+t} - N_{t_1+t} = k\}$$

- Simplicidad:
  - Se refiere a que la probabilidad de que haya más de un evento de llegada en un instante es cero.
  - Esta propiedad se asume para todos los procesos considerados en este curso.

Propiedades de los procesos puntuales

- Independencia:
  - Esta propiedad puede ser expresada como el requisito de que la evolución futura del proceso sólo depende del estado presente. Es decir,

$$p\left\{N_{t_2}-N_{t_1}=k|N_{t_1}-N_{t_0}=n\right\}=p\left\{N_{t_2}-N_{t_1}=k\right\}$$

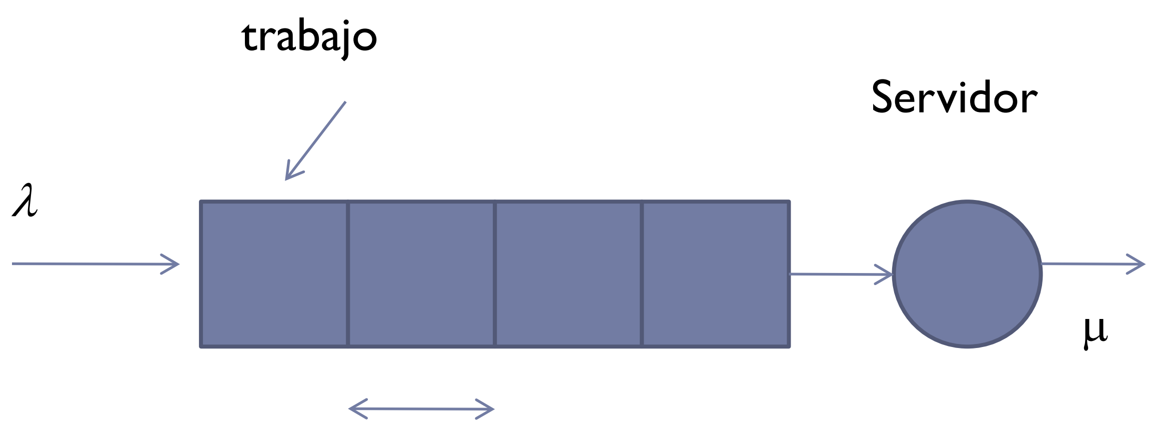
  - Si esto se cumple para todo t, entonces el proceso es un *Proceso de Markov*: Es decir, la evolución futura sólo depende del estado presente, pero es independiente de cómo éste fue obtenido. Esta es la *propiedad de falta de memoria*.
  - Si esta propiedad sólo se mantiene para ciertos puntos en el tiempo (tiempos de llegada), estos puntos son llamados *puntos de equilibrio o puntos de regeneración*. En este caso el proceso tiene una memoria limitada y sólo requiere recordar el último punto de regeneración.

Teorema de Little

- Este es el único resultado general que es válido para todos los sistemas de colas.
- Fue publicado por Little en 1961. Posteriormente fue probado por Eilon (1969) aplicando la teoría de procesos estocásticos.

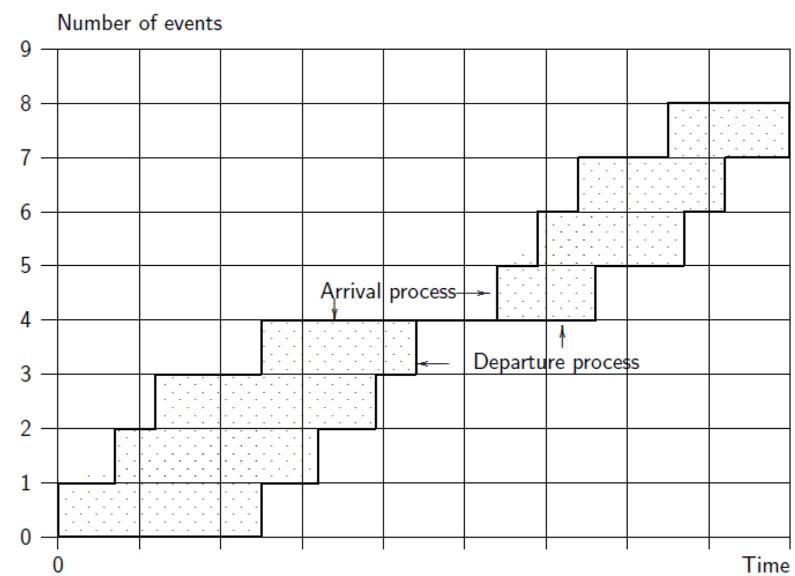
Teorema de Little

- Consideremos un sistema de colas donde los clientes llegan de acuerdo con un proceso estocástico.
- Los clientes entran al sistema en un tiempo aleatorio y esperan para ser servidos.
- Después de ser servidos, los clientes dejan el sistema.



Teorema de Little

- Los procesos de llegada y salida son considerados procesos estocásticos que miden el número de usuarios que llegan/salen (en el eje Y), con respecto al tiempo que ha transcurrido (eje X).



Teorema de Little

- Suponemos que el sistema está en equilibrio en el tiempo t=0.
- Usaremos la siguiente notación:
  - N(T): Número de llegadas en un período T
  - A(T): Los tiempos totales en el sistema de todos los clientes en el período T (*tiempo en el sistema* = *tiempo en cola* + *tiempo de servicio*). Si es una cola FIFO (First-In-First-Out), A(T) es igual al área sombreada entre las curvas. También es igual al volúmen de tráfico transportado.
  - $\lambda(T)=N(T)/T$: Es la intensidad media de llegadas en el período T
  - $W(T)=A(T)/N(T)$: Tiempo medio en el sistema por cada llamada en el período T.
  - $L(T)=A(T)/T$: Número medio de llamadas en el sistema en el período T.

Teorema de Little

- Relación entre variables:

$$L(T) = \frac{A(T)}{T} = \frac{W(T) \cdot N(T)}{T} = \lambda(T) \cdot W(T)$$

- Si se cumplen los límites de:

$$\lambda = \lim_{T \to \infty} \lambda(T) \quad \text{and} \quad W = \lim_{T \to \infty} W(T)$$

- Entonces el valor límite de L(T) también existe y viene siendo:

$$L = \lambda \cdot W \qquad \text{(Teorema de Little)}$$

- Si consideramos sólo los clientes en espera, la fórmula muestra que: *La longitud media de la cola es igual a la intensidad de llegadas multiplicado por el tiempo medio en el sistema.*
- Si sólo consideramos los servidores, la fórmula muestra que: *El tráfico transportado es igual a la intensidad de llegadas multiplicado por el tiempo medio de servicio.* Esto corresponde a la definición de tráfico descrita al inicio del curso.