

1.PROCESOS DE MARKOV

Un proceso de Markov tiene la propiedad de que la probabilidad de comportamiento futuro está totalmente definida si se conoce el estado actual. El conocimiento de estados previos al actual no altera la probabilidad de comportamiento futuro.

$$\begin{aligned} \Pr[a \leq X(t) \leq b \mid X(t_1) = X_1, X(t_2) = X_2, \dots, X(t_n) = X_n] = \\ = \Pr[a \leq X(t) \leq b \mid X(t_1) = X(t)] \\ \text{para cualesquiera } t_1, t_2, \dots, t_n \end{aligned}$$

Si el conjunto de posibles estados del proceso es numerable, el proceso recibe el nombre de cadena de Markov.

Por otra parte, si todas las posibles realizaciones del proceso Markoviano son funciones continuas del tiempo el proceso se denomina *proceso de difusión*.

Para los procesos Markovianos se puede definir una *Probabilidad de transición*

$$\Pr(Z, t; Y, s) = \Pr[X(t) = Z \mid X(s) = Y]$$

En el caso de cadenas de Markov, es posible definir una matriz de probabilidades de transición:

$$P_{i,j}(s,t) = \Pr[X(t) = X_j \mid X(s) = X_i]$$

1.1 La ecuación de Chapman-Kolmogorov

Es una de las relaciones más importantes en las cadenas de Markov. Se expresa:

$$P_{i,j}(s,t) = \sum_{\text{todo-}r} \Pr[X(t) = X_j \mid X(u) = X_r, X(s) = X_i] \Pr[X(u) = X_r \mid X(s) = X_i]$$

siendo $s \leq u \leq t$

Por ser un proceso de Markov

$$\Pr[X(t) = X_j \mid X(u) = X_r, X(s) = X_i] = \Pr[X(t) = X_j \mid X(u) = X_r] = P_{r,j}(u,t)$$

por otra parte

$$\Pr[X(u) = X_r \mid X(s) = X_i] = P_{i,r}(s,u)$$

de donde

$$P_{i,j}(s,t) = \sum_{\text{todo-}r} P_{i,r}(s,u) P_{r,j}(u,t)$$

La ecuación de Chapman-Kolmogorov permite obtener las probabilidades de transición entre dos estados en un

intervalo de tiempo dado, a través de las probabilidades de transición de los posibles estados intermedios.

1.2 La matriz de tasa de transición:

Se define:

$$H(s,t) \equiv [P_{i,j}(s,t)]$$

la ecuación de Chapman-Kolmogorov adopta la forma:

$$H(s,t) = H(s,u) H(u,t)$$

y por tanto

$$\begin{aligned} H(s,t+\Delta t) &= H(s,t) H(t, t+\Delta t) \Rightarrow \\ H(s,t+\Delta t) - H(s,t) &= H(s,t) [H(t, t+\Delta t) - I] \end{aligned}$$

Dividiendo por Δt y tomando el límite cuando $\Delta t \rightarrow 0$

$$\lim_{\Delta t \rightarrow 0} \frac{H(s,t+\Delta t) - H(s,t)}{\Delta t} = H(s,t) \lim_{\Delta t \rightarrow 0} \frac{H(t, t+\Delta t) - I}{\Delta t}$$

Definiendo

$$Q(t) \equiv \lim_{\Delta t \rightarrow 0} \frac{H(t, t+\Delta t) - I}{\Delta t}$$

se obtiene la forma diferencial de la ecuación de Chapman-Kolmogorov:

$$\frac{\partial H(s,t)}{\partial t} = H(s,t) Q(t)$$

y por tanto

$$H(s,t) = \exp \left[\int_s^t Q(\tau) d\tau \right]$$

La matriz $Q(s,t)$ se denomina *matriz de tasa de transición* y representa la variación instantánea de las probabilidades de transición de un estado a otro.

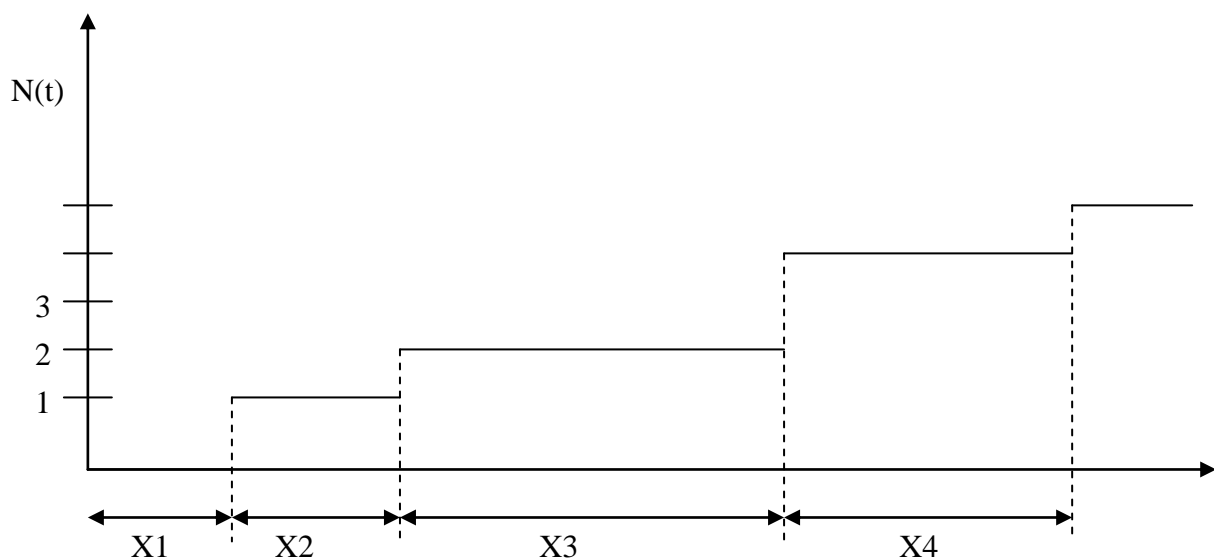
- La forma diferencial de la ecuación de Chapman-Kolmogorov establece una evolución dinámica de las probabilidades de transición y por tanto, de la probabilidad de encontrar el proceso en un determinado estado a medida que transcurre el tiempo.
- La evolución de las probabilidades de estado viene dada por un conjunto de ecuaciones diferenciales lineales que pueden ser resueltas obteniendo la solución y realizar un análisis del transitorio y del estacionario de la probabilidad de estado

2. PROCESOS DE RENOVACIÓN

Un proceso de renovación $N(t)$ se define como un proceso estocástico que toma valores enteros no negativos que contabiliza el número de veces que ocurre un cierto evento durante el intervalo $(0,t]$, donde los intervalos de tiempo entre dos eventos consecutivos son variables aleatorias, positivas, independientes e idénticamente distribuidas.

Se utilizan con frecuencia para modelar la aparición de averías que implican la reparación o sustitución de un componente defectuoso por uno nuevo.

El elemento k -ésimo de la sucesión de variables aleatorias que da el intervalo de tiempo entre dos eventos consecutivos, $\{X_k\}$, se denomina *tiempo de vida* del evento $k-1$.



Asociado al proceso de renovación $N(t)$, se considera el *tiempo de espera* hasta el n -ésimo evento

$$S_n = X_1 + X_2 + X_3 + \dots + X_n \quad n \geq 1$$

y la *función de renovación*, que contabiliza el número esperado de eventos durante el intervalo de tiempo $(0, t]$

$$M(t) = E[N(t)]$$

2.1 Teorema fundamental de la renovación

Es el teorema más importante de la teoría de la renovación. Se puede enunciar del siguiente modo:

Sea X_k un proceso de renovación con $\mu = E[X_1] < \infty$, entonces:

$$\lim_{t \rightarrow \infty} \frac{1}{t} M(t) = \frac{1}{\mu}$$

este teorema relaciona la esperanza matemática de las variables aleatorias con la función de renovación.

Si la varianza del proceso de renovación es $V(t) = \text{var}(N(t))$ y $E[X_1] = \mu$ y $\text{var}[X_1] = \sigma^2$, se verifica:

$$\lim_{t \rightarrow \infty} \frac{V(t)}{t} = \frac{\sigma^2}{\mu^3}$$

3. PROCESO DE POISSON

Un proceso de Poisson $N(t)$, $t \geq 0$, con parámetro λ es un proceso de renovación que contabiliza eventos, siendo la función de distribución de probabilidad del tiempo entre eventos una función de distribución exponencial:

$$F(t) = 1 - e^{-\lambda t}$$

la función de densidad de probabilidad del tiempo entre eventos es:

$$f(t) = \lambda e^{-\lambda t}$$

cuya media y varianza son:

$$E(t) = \frac{1}{\lambda} \qquad \text{Var}(t) = \frac{1}{\lambda^2}$$

La probabilidad de que el número de eventos ocurridos en el intervalo $(0, t]$ sea k , viene dada por la expresión:

$$\Pr[N(t) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

y con esta propiedad, es posible calcular los momentos:

$$E[N(t)] = M(t) = \lambda t$$

$$\text{Var}[N(t)] = E[N^2(t) - E[(N(t))]^2] = \lambda t$$

El proceso de Poisson es un modelo aceptable de un proceso en el que se pueda asumir:

- En cada instante puede tener lugar como máximo un evento.
- $N(t), N(t, t_1), \dots, N(t_{k-1}, t_k)$ son variables aleatorias independientes
- $\Pr[N(t_{k-1}, t_k) = n], n = 0, 1, 2, \dots$ depende de la longitud del intervalo (t_{k-1}, t_k) , pero no depende de los instantes t_{k-1}, t_k .

3.1 Propiedades relevantes

Los procesos de Poisson tienen unas propiedades de indudable interés. De entre ellas podemos destacar

- Carece de memoria, es decir, la función de distribución de probabilidad del tiempo que debe transcurrir hasta el próximo evento es siempre la misma, con independencia del intervalo de tiempo transcurrido desde el último evento:

$$\Pr[V \leq Z + T \mid V \geq z] = 1 - e^{-\lambda T}$$

- La superposición de varios procesos de Poisson caracterizado cada uno de ellos por su

parámetro λ_i es también un proceso de Poisson de parámetro $\sum \lambda_i$

4. PROCESOS DE ACUMULACIÓN

Sea el proceso de renovación cuyos tiempos de vida vienen dados por la sucesión de variables aleatorias $\{X_k\}$. Asociado a X_k se considera la variable aleatoria Y_k , de tal manera que la sucesión $\{Y_k\}$ es una secuencia de variables aleatorias independientes e idénticamente distribuidas. Al proceso dado por:

$$W(t) = \sum_{k=1}^{N(t)+1} Y_k$$

se lo denomina proceso de acumulación.

La variable Y_k se puede interpretar como un coste asociado al intervalo de tiempo x_k , de tal manera que el proceso de acumulación es el coste acumulado durante el intervalo de tiempo $(0,t]$, durante el cual ha habido $N(t)$ eventos.

Si se considera:

$$\begin{aligned} E[X_1] &= \mu \\ E[Y_1] &= \nu \end{aligned}$$

aplicando la ecuación de renovación, se obtiene:

$$t \xrightarrow{\text{lim}} \infty \frac{1}{t} E[W(t)] = \frac{\nu}{\mu}$$

El cociente $\frac{\nu}{\mu}$ se puede interpretar como un coste promedio por unidad de tiempo.

Por otra parte si $\{X_k\}$ y $\{Y_k\}$ son independientes, se verifica:

$$t \xrightarrow{\text{lim}} \infty \frac{\text{Var}[W(t)]}{t} = \frac{\sigma_\nu^2}{\mu} + \nu^2 \frac{\sigma_\mu^2}{\mu^3}$$

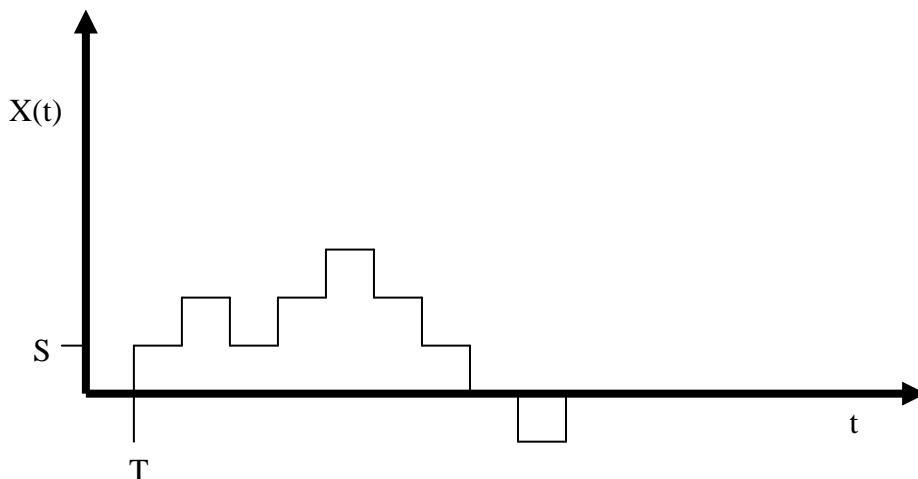
siendo $\sigma_\mu = \text{Var}[X_1]$ y $\sigma_\nu = \text{Var}[Y_1]$

Los procesos de acumulación tienen aplicación en el modelado de procesos en los cuales se ha de evaluar el coste asociado a una serie de eventos; por ejemplo modelos de reemplazo de componentes defectuosos en una instalación.

5. CAMINO ALEATORIO

Se va a considerar una sucesión de variables aleatorias independientes, donde cada variable aleatoria sólo puede tomar dos valores (0 ó 1) equiprobables.

Si se interpreta este proceso como el sentido (derecha o izquierda) de los desplazamientos de longitud s que realiza una partícula cada T segundos, se obtiene un proceso estocástico $X(t)$ que da la posición unidimensional de la partícula en cada instante que se denomina *camino aleatorio*.



Suponiendo que para el instante $t=nT$, el resultado de una realización del camino aleatorio arroja un resultado de k desplazamientos a la derecha y $n-k$ hacia la izquierda:

$$X(nT) = ks - (n - k)s = (2k - n)s = rs$$

$$\Pr[X(nT) = rs] = \binom{n}{\frac{n+r}{2}} \frac{1}{2^n}$$

de donde se obtiene:

$$E[X(nT)] = 0$$

$$E[X^2(nT)] = n \cdot s^2$$

6. PROCESOS DE WIENER-LEVY

El proceso de Wiener-Levy se puede presentar como el caso límite del camino aleatorio cuando $s \rightarrow 0$ y $T \rightarrow 0$

Si en el camino aleatorio se toma $t = nT$

$$E[X(t)] = 0$$

$$E[X^2(t)] = \frac{t \cdot s^2}{T}$$

haciendo tender s y T a cero, la varianza de $X(t)$ se mantiene finita y diferente de cero si y sólo si s tiende a cero como \sqrt{T} . Asumiendo que esto es cierto, se verifica:

$$s^2 = \alpha \cdot t$$

y se define el proceso $W(t) = \lim_{T \rightarrow 0} X(T)$

La función de densidad de probabilidad es:

$$f(W, t) = \frac{1}{\sqrt{2 \cdot \pi \cdot \alpha \cdot t}} \exp\left[-\frac{\omega^2}{2 \cdot \alpha \cdot t}\right]$$

si se fija el valor de t , la función de densidad de probabilidad es la de una v.a. con distribución normal y, por tanto, si $\varpi = r \cdot s$

$$E[W(t)] = 0$$

$$E[W^2(t)] = \alpha \cdot t$$

fijado t , $W(t)$ es una v.a. con distribución normal, de media 0 y varianza $\alpha \cdot t$.

El proceso de Wiener-Levy se denomina también Movimiento Browniano.

El movimiento Browniano puede ser definido como un proceso estocástico que posee las propiedades:

- Cada incremento $W(t_1 + d) - W(t_1)$ es una v.a. con distribución normal de media cero y varianza $\alpha \cdot d$ donde α es fijo.
- Para cada par de intervalos de tiempos disjuntos $[t_1, t_2], [t_3, t_4]$ con $t_4 \geq t_3 \geq t_2 \geq t_1$, los incrementos $W(t_4) - W(t_3)$ y $W(t_2) - W(t_1)$ son v.a. independientes.
- $W(0) = 0$ siendo $W(t)$ continua en $t = 0$

El proceso $\frac{W(t)}{\sqrt{\alpha}}$ se denomina Movimiento Browniano Standard y tiene como particularidad que la varianza es t

7. MOVIMIENTO BROWNIANO REFLEJADO EN EL ORIGEN

Dado un Movimiento Browniano $W(t), t \geq 0$, el proceso estocástico $Y(t)$ que tiene la distribución de:

$$Y(t) = |W(t)| = W(t) = \begin{cases} W(t) \text{ si } W(t) \geq 0 \\ -W(t) \text{ si } W(t) \leq 0 \end{cases}$$

se denomina Movimiento Browniano Reflejado en el origen.

Intuitivamente se puede concebir como el movimiento de una partícula confinada en el semieje positivo. Cuando la partícula alcanza la frontera (0) es reflejada, alejándose de dicha frontera.

La función de distribución es:

$$\begin{aligned} \Pr[Y(t_n + d) \leq z | Y(t_0) = X_0 \cdots Y(t_n) = X_n] &= \Pr[-z \leq X(t_n + d) \leq z | X(t_n)] = \\ &= \int_{-z}^{+z} f(y - X_n, d) dy \end{aligned}$$

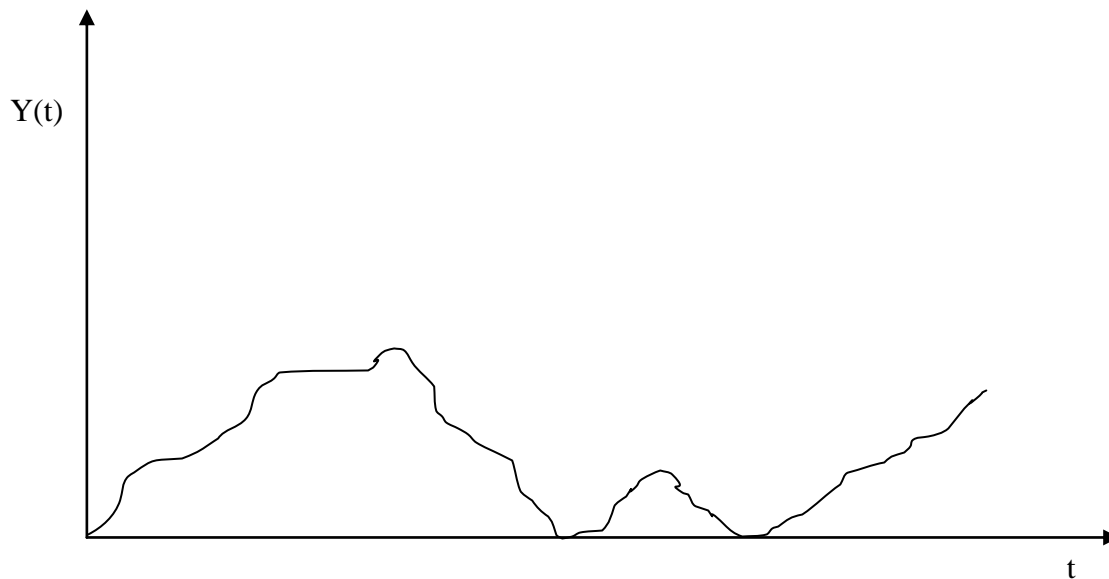
donde,

$$f(W, t) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2 \cdot t}} \exp \left[-\frac{w^2}{2 \cdot \sigma^2 \cdot t} \right]$$

se verifica que:

$$E[Y(t)] = \int_{-\infty}^{+\infty} |X(t)| \cdot f(X, t) \cdot dx = \sqrt{\frac{2t}{\pi}}$$

$$E[Y^2(t)] = E\left[|X(t)|^2\right] - \frac{2t}{\pi} = \left(1 - \frac{2}{\pi}\right) \cdot t$$



8.MOVIMIENTO BROWNIANO CON DESPLAZAMIENTO

Dado un movimiento Browniano $\tilde{W}(t)$, un Movimiento Browniano con Desplazamiento es un proceso estocástico que tiene la distribución de:

$$W(t) = \tilde{W}(t) + \mu \cdot t \quad t \geq 0$$

donde μ es una constante que se denomina parámetro de desplazamiento.

De forma intuitiva, se puede concebir como un Movimiento Browniano al cual se superpone un desplazamiento de velocidad constante.

Posee las propiedades:

- Cada incremento $W(t_1 + d) - W(t_1)$ es una v.a. con distribución normal de media $\mu \cdot t$ y varianza $\sigma^2 \cdot t$; siendo μ y σ constantes.

- Para cada par de intervalos de tiempos disjuntos $[t_1, t_2], [t_3, t_4]$ con $t_4 \geq t_3 \geq t_2 \geq t_1$, los incrementos $W(t_4) - W(t_3)$ y $W(t_2) - W(t_1)$ son v.a. independientes.
- $W(0)=0$ siendo $W(t)$ continua en $t=0$

La v.a dada por:

$$X = \max_{0 \leq t < \infty} W(t) - W(0)$$

tiene la distribución de probabilidad dada por:

$$\Pr\{X \geq \varpi\} = e^{-\lambda\varpi}$$

siendo

$$\lambda = \frac{2|\mu|}{\sigma^2}$$

la función de densidad de probabilidad para el instante en el que se alcanza por primera vez un valor $z \geq W(0)$ condicionado a que $W(0)=\varpi$ es:

$$f(t | \varpi, z) = \frac{z - \varpi}{\sqrt{2 \cdot \alpha \cdot \pi \cdot t^3}} \exp \left[-\frac{(z - \varpi - \mu \cdot t)^2}{2 \cdot \alpha \cdot t} \right]$$

9. APROXIMACIONES DE PROCESOS ESTOCÁSTICOS

Dentro del estudio de procesos estocásticos, existe una parte dedicada al análisis de las propiedades de un proceso estocástico $X(t)$, a través de las propiedades de una sucesión de procesos estocásticos que convergen hacia él $\{X_n(t)\}$

$${}_n \lim_{\rightarrow \infty} X_n(t) = X(t)$$

El estudio formal da lugar a una serie de teoremas límite para procesos estocásticos que son de gran utilidad, ya que en ciertos casos es posible encontrar una sucesión de procesos estocásticos más simples y de propiedades conocidas que permiten inferir las propiedades de un proceso más difícil de analizar y hacia el cual convergen.

9.1 Ley funcional fuerte de los grandes números

Se va a considerar el proceso estocástico

$\{X_i, i = 1, 2, 3, \dots\}$, formado por una sucesión de v.a. independientes e idénticamente distribuidas. A partir de este proceso podemos definir la variable aleatoria

$$S(t) \equiv \sum_{i=1}^{\lfloor t \rfloor} X_i \quad t \geq 0$$

donde

$$\lfloor t \rfloor = j, j \in \{1, 2, 3, \dots\} \quad |, j \leq t \leq j + 1$$

Si se interpreta X_i como la demanda de un producto en el periodo de tiempo i -ésimo, $S(t)$ es la demanda acumulada en el intervalo de tiempo $[0, \lfloor t \rfloor]$.

Sea $E[X_1] = \mu$, entonces la ley funcional fuerte de los grandes números establece que:

$$\bar{S}^n(t) \equiv \frac{1}{n} S(nt) \rightarrow \mu \cdot t \equiv \bar{S}(t)$$

con probabilidad 1 cuando $n \rightarrow \infty$.

Esta ley establece que la media de múltiples realizaciones del proceso $S(t)$ converge hacia una función lineal del tiempo.

9.2 TEOREMA FUNCIONAL DEL LÍMITE CENTRAL.

Se va a considerar el proceso estocástico $\{X_i, i = 1, 2, 3, \dots\}$, formado por una sucesión de v.a. independientes e idénticamente distribuidas. A partir de este proceso podemos definir la variable aleatoria

$$S(t) \equiv \sum_{i=1}^{\lfloor t \rfloor} X_i \quad t \geq 0$$

donde

$$\lfloor t \rfloor = j, j \in \{1, 2, 3, \dots\} \quad |, j \leq t \leq j+1$$

siendo Γ la matriz de covarianzas de x_1 ; el teorema funcional del límite central establece que: \rightarrow

$$\hat{S}^n(t) \equiv \sqrt{n}[\bar{S}^n(t) - \bar{S}(t)] \rightarrow \Gamma^{\frac{1}{2}}W(t) \equiv \hat{S}(t)$$

cuando $n \rightarrow \infty$ siendo $W(t)$ un Movimiento Browniano Standard K -dimensional.

La convergencia en este caso es en distribución, es decir, la función de distribución de probabilidad de $\hat{S}^n(t)$ tiende a la de $\hat{S}(t)$ cuando $n \rightarrow \infty$. El teorema funcional del límite central constituye una extensión aplicada a procesos estocásticos del teorema del límite central para variables aleatorias

9.3 APROXIMACIONES FUERTES.

El teorema Funcional de las Aproximaciones Fuertes engloba bajo una misma perspectiva la Ley Funcional de los Grandes Números y el teorema Funcional del Límite Central. Básicamente establece que:

$$\|S(t) - \tilde{S}(t)\|_T \equiv \sup_{0 \leq t \leq T} |S(t) - \tilde{S}(t)| = o\left(T^{\frac{1}{r}}\right)$$

cuando $T \rightarrow \infty$, siendo $r \geq 2$ el orden de los momentos finitos de la variable aleatoria x_1 y se define:

$$\tilde{S}(t) = \mu \cdot t + \Gamma^{\frac{1}{2}} W(t) \quad t \geq 0$$

De esta manera emergen tres niveles de aproximación al proceso

- 1.- $S(t)$ Aproximación detallada del proceso (aproximación microscópica)
- 2.- $\tilde{S}(t)$ Aproximación intermedia del proceso (aproximación mesoscópica)
- 3.- $\bar{S}(t)$ Aproximación agregada del proceso (aproximación macroscópica)

9.4 APROXIMACIONES DE LOS PROCESOS DE RENOVACIÓN

Dado que el número de eventos que han tenido lugar hasta el instante t ($N(t)$) no es suma de variables aleatorias, no se puede aplicar las aproximaciones del apartado 9.3.

No obstante dado que $S_n = X_1 + X_2 + \dots + X_n$ es suma de variables aleatorias independientes e idénticamente distribuidas y que $N(t) \leq n$ si y sólo si $S_n(t)$, entonces es posible aplicar el Teorema Funcional del Límite Central para procesos de renovación.

Si $E[X_1] = \mu$ y $\text{var}[X_1] = \sigma^2$, se obtiene:

$${}_n \lim_{\rightarrow \infty} \frac{N(nt)}{n} \rightarrow \frac{1}{\mu} t$$

$${}_n \lim_{\rightarrow \infty} \frac{N(n \cdot t) - n \cdot t \frac{1}{\mu}}{\sqrt{n}} \rightarrow \frac{\sigma}{\mu^{\frac{3}{2}}} W(t)$$

de donde la aproximación fuerte de un proceso de renovación se puede expresar:

$$\sup_{0 \leq t \leq T} \left| N(t) - \frac{1}{\mu} t - \frac{\sigma}{\mu^{\frac{3}{2}}} W(t) \right| = o \left[T^{\frac{1}{r}} \right]$$

cuando $T \rightarrow \infty$

10. TEORÍA DE COLAS DE ESPERA

La teoría de colas de espera tiene por objeto:

- modelado
- análisis
- control

de sistemas formados por varios *servidores* que atienden a unos *clientes* que aguardan formando una cola de espera a ser atendidos

Son muchos los sistemas que se ajustan a la idea de cola-cliente-servidor:

- vehículos ante un semáforo
- redes telemáticas
- sistemas de fabricación ...

El gran ámbito de aplicación de la teoría de colas ha hecho que exista un gran número de publicaciones sobre el tema.

Recientemente el interés que ha suscitado el estudio de los sistemas de producción como sistemas de eventos discretos ha provocado que muchos investigadores se dediquen al modelado, análisis y control de sistemas de producción utilizando las redes de colas.

10.1 ELEMENTOS DE LOS SISTEMAS DE COLAS

En los sistemas de colas podemos diferenciar:

- **clientes** : Elementos que requieren unos servicios
- **servidores**: Elementos capaces de proporcionar un servicio
- **cola**: En la cual se acumulan los clientes, en espera de que un servidor les pueda atender.

Si bien en un sistema físico estos elementos se pueden reconocer con facilidad, la característica utilizada para el modelado de cada uno de ellos no se reconoce tan fácilmente:

- Los *clientes* se caracterizan por el *proceso de tiempos de llegada*, pudiendo haber un proceso por cada tipo de clientes
- Los *servidores* se caracterizan por el *proceso* formado por los *tiempos de servicio*; cada servidor tiene un proceso para cada una de las clases de clientes.
- La *cola* se caracteriza por el número de *clientes que contiene en cada instante*. También se puede considerar el número máximo de clientes que admite la cola como otro parámetro característico de la cola.

10.2 NOTACIÓN

Para referirse a una cola está bastante extendida una notación cuya principal ventaja es que describe de forma explícita y concisa las propiedades de dicha cola. Esta notación tiene la forma:

$$A/B/m/K$$

- A es la distribución del tiempo de llegada de los clientes
- B es la distribución de los tiempos de servicio
- m es el número de servidores que atienden a dicha cola
- K es el número máximo de clientes que puede haber en la cola

Para las distribuciones de tiempos de llegada y servicio se utiliza la notación:

- **G**: Distribución general de los tiempos de servicio/llegada. No se establece ninguna característica sobre el proceso
- **GI**: Distribución general de un proceso de renovación; los tiempos entre llegadas/servicios están idénticamente distribuidos.
- **D**: Los tiempos entre llegadas/servicios son deterministas
- **M**: Los tiempos entre llegadas/servicios están distribuidos exponencialmente.

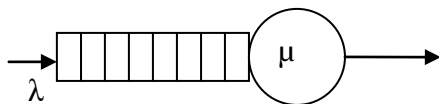
Si se omite K , se sobreentiende que la cola es de capacidad infinita.

Ejemplos:

- $M/M/1$: cola con un único servidor y tiempos de llegada y servicio distribuidos exponencialmente
- $M/G/2/10$: cola con dos servidores que admite como máximo dos clientes. Los intervalos de tiempo entre llegadas siguen una distribución exponencial y los tiempos de servicio una general.

Cola Markoviana: tiempos de llegada y de servicio distribuidos exponencialmente.

Gráficamente una cola de espera se representa:



10.3 DINÁMICA DE UNA COLA

Para el k-ésimo cliente vamos a considerar:

- A_k : Instante en el que *llega* a la cola
- D_k : Instante en el que *abandona el sistema* después de haber recibido servicio
- W_k : *Tiempo de espera* hasta que comienza el servicio
- S_k : Tiempo de permanencia en el sistema ($S_k = D_k - A_k$)
- Z_k : Tiempo de *servicio*.
- Y_k : Intervalo de tiempo transcurrido desde la llegada del cliente k-1 hasta la llegada del cliente k-ésimo.

Vamos a considerar los procesos:

- $X(t)$: Número de clientes en la cola en el instante t
- $U(t)$: Tiempo necesario para vaciar la cola si no legase ningún cliente más (carga del sistema)

$X(t)$ es el proceso que describe el estado de la cola, sin embargo se presta una especial atención a W_k . Se verifican las relaciones:

$$S_k = W_k + Z_k$$
$$D_k = A_k + W_k + Z_k$$

Cuando llega el k -ésimo cliente hay dos posibles situaciones:

1. – *La cola está vacía*

$$D_k \leq A_k$$

$$D_{k-1} - A_k \leq 0 \Leftrightarrow W_k = 0$$

2. – *cola no está vacía*

$$D_{k-1} - A_k \leq 0 \Leftrightarrow W_k = D_{k-1} - A_k$$

Los dos casos se pueden reunir en una única expresión:

$$W_k = \max\{0, D_{k-1} - A_k\}$$

y, por tanto

$$W_k = \max\{0, W_{k-1} + Z_{k-1} - Y_k\}$$

$$S_k = \max\{0, S_{k-1} - Y_k\} + Z_k$$

Finalmente, la dinámica de la cola queda descrita por:

$$W_k = D_k - A_k - Z_k$$

$$D_k = \max\{A_k, D_{k-1}\} + Z_k$$

En general la función de distribución de probabilidad de W_k depende de K , pero, con frecuencia existe una distribución de probabilidad estacionaria cuando $K \rightarrow \infty$

$$\lim_{k \rightarrow \infty} \Pr[W_k \leq t] = \Pr[W \leq t]$$

Cuando este límite existe la variable aleatoria W describe el tiempo de espera de un cliente; entonces $E[W]$ es el tiempo medio de espera de un cliente y $E[S]$ es el tiempo medio de permanencia del cliente en el sistema.

FABRICACIÓN AVANZADA

Dpto. Ingeniería de sistemas y automática

ETS. Ingenieros Industriales, Valladolid

Algo similar puede suceder con los procesos estocásticos $\{X(t)\}$ y $\{U(t)\}$ para los cuales puede existir una distribución de probabilidad estacionaria cuando $K \rightarrow \infty$, de tal manera que las variables aleatorias X y U describen la longitud de la cola y la carga del sistema en estacionario; $E[X]$ y $E[U]$ son la longitud media de la cola y la carga media del sistema en estado estacionario.

En estado estacionario se definen los parámetros:

- Intensidad de tráfico: $\rho = \frac{\text{Tasa de llegada}}{\text{Tasa de servicio}} = \frac{\lambda}{\mu}$
- Utilización: Fracción de tiempo que el servidor se encuentra ocupado $= 1 - \Pr[X=0]$
- Tasa de salida= tasa de partida de los clientes tras recibir servicio $= \mu(1 - \Pr[X = 0])$

En estado estacionario la tasa de salida ha de ser igual a la tasa de llegada ya que si no fuese así o bien la cola crece indefinidamente (tasa salida \leq tasa de llegada) o bien se generan clientes dentro de la cola (tasa salida \geq tasa de llegada).

Uno de los resultados generales más importante en la teoría de colas es la *Ley de Little*:

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

Establece una relación entre las funciones de distribución del tiempo de permanencia en la cola y el número de clientes en el sistema a través de sus esperanzas matemáticas.

11. REDES DE COLAS DE ESPERA:

Las redes de colas son sistemas en los cuales se puede distinguir varios servidores que pueden dar diferentes servicios y clientes que requieren estos servicios, de tal forma que tras recibir servicio en un servidor, el cliente puede circular hasta otra cola donde recibirá otro servicio.

Existen multitud de sistemas físicos que se ajustan a esta idea:

- Sistemas de producción
 - Máquina (servidor)
 - Pieza (cliente)

- Sistemas de transmisión de datos
 - Computador (servidor)
 - Mensaje (cliente)

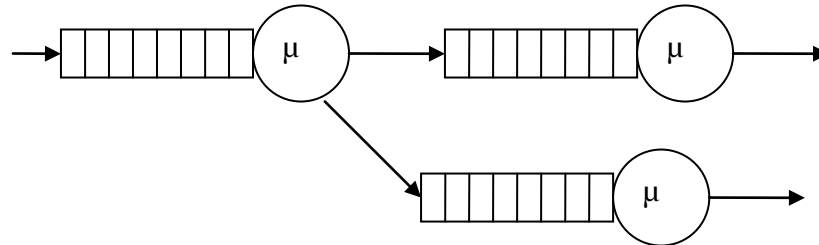
El estado de una red de colas viene dado por el vector:

$$X(t) = [X_1(t), X_2(t), \dots, X_m(t)]$$

que en cada componente tiene el número de clientes en espera en la cola asociada a dicha componente.

La clasificación más simple de las redes de colas es:

- *Redes de colas abiertas*: En las cuales, los clientes proceden del exterior de la red y tras recibir los servicios requeridos, abandonan el sistema.
- *Redes de colas cerradas*: Los clientes ni llegan del exterior ni salen del sistema; circulan de una cola a otra requiriendo diferentes servicios



RED DE COLAS ABIERTA

11.1 REDES ABIERTAS DE COLAS MARKOVIANAS

Las redes abiertas de colas Markovianas están formadas por m nodos, donde cada nodo constituye una entidad cola-servidor con un tiempo de servicio distribuido exponencialmente

El estudio de la cola M/M/1 arroja un importante resultado que recibe el nombre de teorema de Burke:

“El proceso de los tiempos de salida del servidor de los clientes es exponencial y con un parámetro igual al del proceso de los tiempos de llegada”

este resultado es de extrema importancia ya que en una red los clientes que salen de una cola entran en otra para la que el proceso de llegada será markoviano.

En estas redes se considera que el camino de un cliente no es determinista definiendo una matriz $P = [p_{j,i}]$ cuyos elementos son la probabilidad de que el cliente que abandona el nodo j se dirija al nodo i ; de esta forma la tasa de llegada de clientes al nodo i , λ_i , se expresa:

$$\lambda_i = r_i + \sum_{j=1}^m \lambda_j p_{j,i}$$

donde r_i es la tasa de llegada de clientes desde el exterior al nodo i .

Si $p_{i,i} = 0$, la probabilidad estacionaria del número de clientes en las colas, es el producto de las probabilidades estacionarias del número de clientes en cada cola:

$$\Pr[X = [X_1 = N_1, X_2 = N_2, \dots, X_m = N_m]] = \Pr[X_1 = N_1] \cdot \Pr[X_2 = N_2] \cdot \dots \cdot \Pr[X_m = N_m]$$

es por esta razón que se dice que la cola tiene solución en *forma de producto*.

En el caso $p_{i,i} \neq 0$, el proceso de llegada al nodo i -ésimo no es de Poisson, J.R. Jackson estableció que el nodo se comporta como si tuviese un proceso de llegada de Poisson y se mantiene la solución en forma de producto.

11.2 REDES CERRADAS DE COLAS MARKOVIANAS

En este tipo de redes, el número de clientes en la red se mantiene constante y los clientes circulan de un nodo a otro requiriendo los diferentes servicios que dan los servidores.

Las redes cerradas se pueden concebir como un caso particular de las redes abiertas en las cuales la tasa externa de clientes es cero en todos los nodos; por tanto, se restringe el espacio de posibles estados de la red, de tal manera que:

$$\sum_{i=1}^m N_i = N_T$$

siendo N_T el número total de clientes en la red.

La distribución de probabilidad estacionaria admite una solución en forma de producto:

$$\Pr[X_1 = N_1, X_2 = N_2, \dots, X_m = N_m] = \frac{1}{C(N_T)} \left(\frac{\lambda_1}{\mu_1}\right)^{N_1} \left(\frac{\lambda_2}{\mu_2}\right)^{N_2} \dots \left(\frac{\lambda_m}{\mu_m}\right)^{N_m}$$

donde $C(N_T)$ es una constante que depende del número de clientes, N_T , y puede ser obtenida mediante la condición:

$$\frac{1}{C(N_T)} \sum_{N_1, N_2, \dots, N_m} \left(\frac{\lambda_1}{\mu_1}\right)^{N_1} \left(\frac{\lambda_2}{\mu_2}\right)^{N_2} \dots \left(\frac{\lambda_m}{\mu_m}\right)^{N_m} = 1$$

La obtención de $C(N_T)$ no es una tarea sencilla cuando el número de posibles estados es elevado, sin embargo existen técnicas computacionales eficientes y si tan solo se está interesado en la longitud media de cada cola, se puede utilizar el método del valor medio.

11.3 REDES DE COLAS NO MARKOVIANAS

Cuando en una cola el proceso de los tiempos de llegada o el proceso de los tiempos de servicio no es de Poisson, la cola no es markoviana.

El análisis de las colas no Markovianas es mucho más complejo, debido , principalmente a que el sistema no carece de memoria y no basta con conocer el estado actual para determinar la probabilidad de transición a un nuevo estado; se debe conocer la evolución hasta el estado actual.

Si sólo influye el tiempo transcurrido hasta el instante actual la red cae dentro de los procesos *semi-markovianos generalizados*.

Para el estudio de colas no markovianas, existen dos procedimientos generales:

- Utilizar procesos markovianos como elementos básicos para construir un proceso no markoviano
- Estudio de propiedades estructurales del sistema que permitan obtener información sobre su comportamiento.

Existe una abundante bibliografía dedicada al estudio de colas no markovianas con resultados en forma cerrada en algunos casos; en otros casos se ha obtenido aproximaciones que permiten un tratamiento computacional del problema.

12- REDES DE FLUIDOS

En una red de colas el estado viene dado por el número de clientes en cada cola -> el espacio de estado es discreto

Si se considera un sistema formado por nodos, donde cada nodo consta de:

- Un depósito que almacena fluido que va ser procesado.
 - Dispositivo que procesa fluido.
- y, además
- Conducciones que permiten circular el fluido de unos nodos a otros.

Se observa un cierto paralelismo con una red de colas:

- Fluido->clientes
- Depósito->cola
- Procesador de fluido ->servidor

Este nuevo sistema se denomina red de fluidos y su principal diferencia con respecto a las redes de colas es que su espacio de estado es continuo

12.1 MODELO DE UNA RED DE FLUIDOS

Se va a considerar una red de fluido con K nodos; cada nodo consta de un buffer que almacena fluido y una estación de procesamiento de fluido.

Se define:

- $Z_k(i)$ Nivel de fluido en el buffer de la estación k -ésima al final del periodo i -ésimo.
- $\alpha_k(i)$ Volumen de fluido que entra en el nodo k -ésimo durante el periodo i -ésimo.
- $\delta_k(i)$ Volumen de fluido que sale del nodo k -ésimo durante el periodo i -ésimo.
- $P=[p_{j,k}]$ Matriz que da la proporción de fluido que, saliendo del nodo k , va directamente al nodo j .
- $\mu_k(i)$ Volumen máximo de fluido que se puede procesar en el nodo k durante el periodo i -ésimo (el buffer no se vacía)

El volumen de fluido en el depósito k -ésimo al final del periodo i -ésimo se obtiene aplicando la ecuación de balance de fluido en el nodo

$$Z_k(i) = Z_k(i-1) + \alpha_k(i) + \sum_{j=1}^K \delta_j(i) p_{j,k} - \delta_k(i)$$

si se define

$$y_k(i) = \mu_k(i) - \delta_k(i)$$

$y_k(i)$ representa la pérdida de volumen procesado debido a que el buffer está vacío durante algunos instantes del periodo i -ésimo.

Si se considera:

$$X_k = +\alpha_k(i) + \sum_{j=1}^K \delta_j(i) p_{j,k} - \mu_k(i)$$

$X_k(i)$ será el balance de fluido si al comenzar el periodo pésimo el buffer esta vacío y todos los nodos procesan el volumen máximo de fluido (los buffer siempre están llenos).

Si ahora se considera los vectores columna $Z(i)$, $X(i)$, $Y(i)$ en los cuales las componentes pésima son $Z_k(i)$, $X_k(i)$, $Y_k(i)$, la ecuación de balance de fluido es:

$$Z(i) = Z(i-1) + X(i) + [I - P']Y(i)$$

$$Y(i) \geq 0$$

siendo

$$X(i) = +\alpha(i) + \sum_{j=1}^K [P' - I] \mu_j(i)$$

El menor de los $Y(i)$ tal que verifique las relaciones anteriores es el que da lugar al modo de operación más eficiente de la red de fluido, ya que desperdicia el mínimo de capacidad potencial de procesado.

Se verifica que:

$$\mathfrak{S}(x) = \{y \geq 0 : x + [I - P']y \geq 0\} \quad x \in \mathfrak{R}^k$$

tiene un extremo inferior único que, además es la solución de:

$$z = x + [I - P']y$$

$$z \geq 0$$

$$z'y = 0$$

La solución de este problema daría el valor y que origina el modo de operación más eficiente

12.2 FORMULACIÓN EN TÉRMINOS DE ACUMULACIÓN

Se define:

$$A(t) = \sum_1^{\lfloor t \rfloor} \alpha(i)$$

$$M(t) = \sum_1^{\lfloor t \rfloor} \mu(t)$$

$$X(t) = Z(0) + A(t) + [P' - I]M(t)$$

$$Y(t) = \sum_1^{\lfloor t \rfloor} y(t)$$

donde $\lfloor t \rfloor$ es el número entero más próximo a t .

Existe un único par (Y, Z) que satisface las ecuaciones:

$$Z(t) = X(T) + [I - P']Y(t)$$

$$Y(t) \geq 0$$

$$Y(0) = 0$$

$Y(t)$ es no decreciente

$$Z_k(t)[Y_k(t) - Y_k(t-1)] = 0; \quad t = 1, 2, \dots; \quad k = 1, 2, \dots, K$$

donde $Y(t)$ es el menor de los posibles que satisface las ecuaciones anteriores.

Dada $X(t)$, quedan fijadas $Y(t)$ e $Z(t)$ a través de las relaciones indicadas. $Z(t)$ e $Y(t)$ pueden ser expresadas como funciones de $X(t)$:

$$Z(t) = \Phi(X(t))$$

$$Y(t) = \Psi(X(t))$$

se verifica que las funciones Φ, Ψ son lipschitzianas, es decir, definida la norma:

FABRICACIÓN AVANZADA

Dpto. Ingeniería de sistemas y automática

ETS. Ingenieros Industriales, Valladolid

$$\|\xi(t)\|_T = \sup_{0 \leq t \leq T} (\max_{1 \leq k \leq K} |\xi(t)|)$$

existe una constante $h \geq 0$, tal que:

$$\begin{aligned} \|\Phi(X^1(t)) - \Phi(X^2(t))\|_T &\leq h \|X^1(t) - X^2(t)\| \\ \|\Psi(X^1(t)) - \Psi(X^2(t))\|_T &\leq h \|X^1(t) - X^2(t)\| \end{aligned} \quad \text{para todo } X^1 \text{ y } X^2$$

Hasta ahora se ha considerado un modelo de tiempo discreto; un desarrollo análogo se puede realizar cuando la red se considera un sistema de tiempo continuo; en este caso el modelo es:

$$\begin{aligned} Z(t) &= X(T) + [I - P]Y(t) \\ dY(t) &\geq 0 \\ Y(0) &= 0 \\ Z(t)dY(t) &= 0 \end{aligned}$$

y, fijada $X(t)$ existe un único par (Y, Z) que verifica todas las condiciones:

$$\begin{aligned} Z(t) &= \Phi(X(t)) \\ Y(t) &= \Psi(X(t)) \end{aligned}$$

siendo $X(t)$ una función K -dimensional del tiempo.

Definida $X(t)$, el sistema evoluciona según las ecuaciones anteriores; teniendo en cuenta que $X(t)$ representa el balance de fluido en la situación ficticia en la cual todos los nodos procesan el máximo volumen posible, $X(t)$ no depende de la dinámica de la res de fluidos, sino de la máxima capacidad de procesamiento de cada nodo. Si es posible formular dicha capacidad de procesamiento como

función sólo del tiempo $X(t)$ dispara la dinámica de la red a través de los modelos obtenidos.

Una posibilidad es considerar que $X(t)$ es aproximada por una función de la forma:

$$\bar{X}(t) = Z(0) + g \cdot t$$

se obtiene así una red de fluido determinista.

12.3 DINÁMICA ANTE MOVIMIENTO BROWNIANO

Vamos a considerar una red de fluidos cuya dinámica viene dada por las ecuaciones:

$$Z(t) = X(T) + [I - P]Y(t)$$

$$dY(t) \geq 0$$

$$Y(0) = 0$$

$$Z_k(t)dY_k(t) = 0$$

cuya solución, dado $X(t)$, es:

$$Z(t) = \Phi(X(t))$$

$$Y(t) = \Psi(X(t))$$

y se considera el caso en que $X(t)$ es un movimiento Browniano k -dimensional con desplazamiento g y matriz de covarianzas Γ ; en este caso el proceso:

$$Z(t) = \Phi(X(t))$$

es un movimiento Browniano k -dimensional reflejado en el origen que evoluciona en el interior del ortante no negativo

Si se verifica:

$$[I - P']^{-1}(\alpha + [P' - I]\mu) \leq 0$$

$$P_{k,k} = 0 \quad k = 1, 2, \dots, K$$

$$2\Gamma_{j,k} = (P_{j,k} \Gamma_{k,k} + P_{j,k} \Gamma_{j,j}) \quad j \neq k$$

entonces existe una densidad de probabilidad estacionaria para Z en forma de producto:

$$f(Z) = \prod_{k=1}^K \eta_k e^{-\eta_k z_k} \quad Z = z_1, z_2, \dots, z_k \geq 0$$

siendo

$$\eta = (\eta_k) = -2 \cdot \text{diag}(\Gamma)^{-1} [I - P]^{-1} [\alpha + [P' - I]\mu]$$

Por ser $X(t)$ un proceso estocástico no influenciado por la propia dinámica del sistema, provoca una evolución de la red de fluidos que no es determinista, pero que permite obtener una función de densidad de probabilidad estacionaria del volumen de fluido en cada depósito.

La aproximación fuerte de un proceso estocástico permite obtener un modelo aproximado, que es un movimiento Browniano; por tanto la red de fluido estocástica aproxima el comportamiento de una red de colas.