Práctica de simulación. Cadena de Markov de parámetro continuo.

1. Descripción del modelo

Mediante el algoritmo de Gillespie para Cadenas no Markovianas se trata de obtener una traza de simulación para la cadena no markoviana mostrada en la figura 1. Aparece para el caso particular de los parámetros que se muestran en la misma figura (autobuses con solo capacidad para tres pasajeros) para poder tener una representación gráfica.

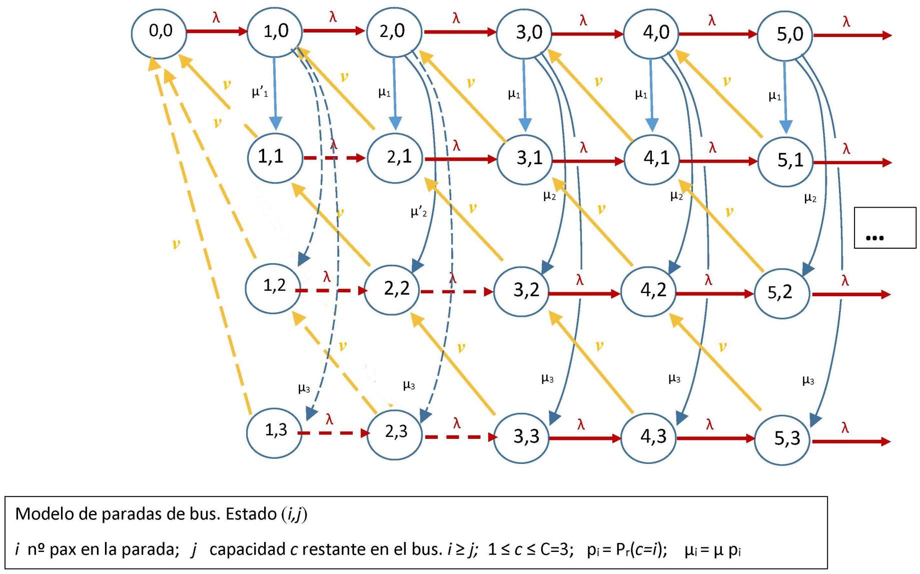

Figura 1: Cadena para modelizar las colas de espera en una parada de autobús

Las transiciones se producen reguladas por las fuentes de sucesos mostradas en la figura 2

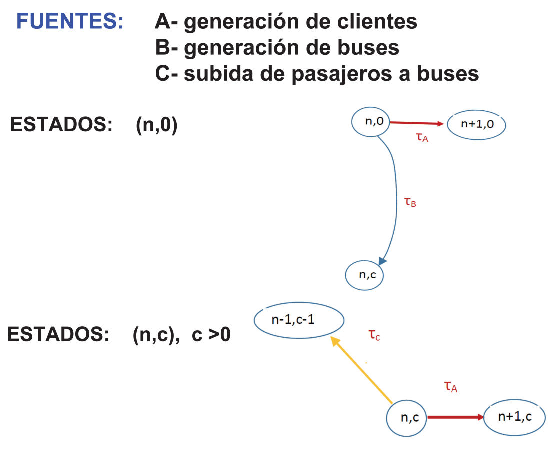

Figura 2: Fuentes asociadas a las transiciones de la cadena en la figura 1

2. Descripción de la traza de simulación y su tratamiento

El detalle del código a desarrollar para hacer la simulación aparece en el Apéndice 1. A destacar en él está qué debe registrarse en el fichero de traza de simulación.

Observar que se obtendrá un conjunto de registros en los que figurará en cada uno de ellos una transición $i \to j$:

- Estado $i = (n_i, c_i)$ (si $c_i = 0$, entonces se supone que no hay bus en parada)
- Instante t_i en que aparece el estado i
- Estado $j = (n_j, c_j)$ (si $c_j = 0$, entonces se supone que no hay bus en parada)
- Instante t_j en que aparece el estado j

El tiempo de permanencia en el estado i será pues, $t_j - t_i$.

Habrá tantos registros en la traza como transiciones se hayan simulado. Supondremos siempre simulaciones efectuadas dentro del horizonte finito $[0, T]$ en las que resultaran un número total p de pasajeros generados y un número total q de autobuses que han llegado a la parada. Hay que observar que, puesto que se limita el horizonte de tiempo, la traza ya no recogerá transiciones correspondientes a subidas en autobús, que afecten a los últimos pasajeros que llegaron.

Evidentemente, para facilitar el tratamiento posterior de la traza, pueden añadirse más elementos en cada registro si se considera oportuno.

2.1. Magnitudes a estimar mediante la simulación

Mediante la simulación se generará pues, un fichero de traza con una estructura de registros como la indicada anteriormente en 2. Dos magnitudes cuya estimación resulta natural en este sistema son:

1. La ocupación del sistema sobre un intervalo de tiempo $[0, T]$. Indicadores: ocupación temporal media L (longitud de cola), varianza de la ocupación.

2. El tiempo de permanencia en la parada por pasajero. Consideraremos el tiempos de permanencia w:

   - Por w entenderemos el tiempo entre la llegada del pasajero y el instante en que éste sube al autobús.

   Estas magnitudes no aparecen directamente en la traza, sino que deben obtenerse para cada pasajero de forma indirecta de la información que sí registra la traza.

   Se obtendrá una muestra de tiempos de permanencia en la parada $w_1, w_2, ..., w_{p'}$, $p' < p$, dado que los últimos en llegar subirán fuera ya de $[0, T]$. Deberá procesarse para obtener $\bar{w}$ i s_w^2 como estimadores de la media W y de la varianza $Var[w]$.

2.1.1. Cálculo de la ocupación media o longitud de cola L y su varianza

Simplemente recorrer la traza:
- Inicialización: $M = 0$;
- Para cada registro hacer: $M = M + n_i(t_j - t_i)$

Al final la ocupación media en # de pasajeros es estimada mediante $L_T = M/T$, siendo T el período de simulación.

Puede también estimarse la varianza de la ocupación

$$Var[N] = \frac{1}{T} \int_0^T (N(t) - L_T)^2 dt = \frac{1}{T} \int_0^T N^2(t) dt - L_T^2 \tag{1}$$

Simplemente recorrer la traza:
- Inicializar $M = 0$;
- Para cada registro hacer: $M = M + n_i^2(t_j - t_i)$

Al final la varianza del # de pasajeros es estimada mediante

$$s_N^2 = \frac{M}{T} - L_T^2 \tag{2}$$

2.1.2. Cálculo del tiempo de espera W

Se supondrá comportamiento FIFO de los pasajeros. Se trata también de ir recorriendo la traza viendo qué subida al autobús se corresponde con la llegada a la parada del correspondiente pasajero.

Para poder estimar el tiempo de permanencia en la parada w_ν, del pasajero ν, se necesitará de un vector o espacio de trabajo p[ ] y de dos apuntadores ini, fi. El apuntador ini indica la posición más a la izquierda de p en la que hay información en curso, mientras que fi indica la posición final. En una posición determinada de p[ ] se recoge el instante correspondiente a la llegada de un pasajero a la parada o el instante de subida de un pasajero al autobús.

- Inicialización: fi = ini =0; $n_c = 0$
- <u>Mientras</u> no fin de traza <u>hacer:</u>
    - leer registro (i, t_i, j, t_j)
    - <u>Si</u> $n_i < n_j$ (llegada) <u>entonces</u>: { fi =fi+1; p(fi)= t_j ; $n_c = n_c + 1$; }(llegada de pasajero a la parada)
    - <u>Si</u> $n_i > n_j$ (subida) <u>entonces</u>: { ini = ini+1; $w_{ini} = t_j -$ p(ini); (subida de pasajero al bus)
      Añadir w_{ini} a la muestra de tiempos de espera }
      (en este caso se genera el tiempo de permanencia del pasajero *ini*)
    - <u>FinSi</u>
- <u>FinMientras</u>

A posteriori habrá de tratarse la muestra (exhaustiva) de tiempos de permanència de cada cliente para ver estadísticas bàsicas.

3. Realización de la práctica

Para cada estudiante se especificarán

1. Las distribuciones de probabilidad para las fuentes B, C
2. K = capacidad máxima de cada autobús.
   La capacidad con la que venga un autobús se tomará como una v.a. γ siguiendo una exponencial truncada a K pasajeros y redondeada por arriba, de esperanza $K(e^2\text{-}3)/(2e^2\text{-}2) \sim K/3$ y función de densidad de probabilidad:

$$f_\gamma(c) = \frac{2}{K(e^2-1)} \exp\left(2(1-\frac{c}{K})\right), \quad 0 \le c \le K \tag{3}$$

3. Los factores de carga ρ bajo los que se quiere hacer funcionar el sistema. Se tomará en todos los casos $\rho = 0{,}3, 0{,}53333, 0{,}75, 0{,}9$

Se deberán efectuar 4 simulaciones, una para cada factor de carga ρ, tomándose en cada una de ellas un proceso poissoniano de llegada de pasajeros a la parada (fuente A), con tasa de llegada λ pax/min, de forma que se ajusten a los factores de carga establecidos, es decir, si el tiempo medio entre dos llegadas consecutivas de autobús es $1/\mu$, entonces

$$\lambda = \rho \cdot \mu \cdot E[\gamma] \tag{4}$$

Se tomará un intervalo de tiempo de $T = 300min$ (5 horas)

1. **(1p)** Implementar un método para generar muestras de alguna de l as variables aletorias de acuerdo con l as distri-buciones para l as f uentes B, C y l a variable aleatoria γ.
   Generar muestras mediante el método implementado y efectuar un test de bondad de ajuste para comprobar que, efectivamente obedece a la distribución deseada, reportando histogramas de las muestras
2. **(4.5p)** Implementar en el l enguaje que se prefiera el código de l a figura 3.
3. **(1.5p)** (apartado c) de la preentrega) Establecer una evaluación preliminar del simulador calculando L:

   - Mediante el modelo $M/M^{[X]}/1/K$ que hayáis establecido en la preentrega, resolviendo las probabilidades de e.e. mediante las ecuaciones de equilibrio (deberéis resolverlas)
   Contrastad el valor de L calculado con el que proporcionaría vuestro simulador si los autobuses siempre llegasen con capacidad cte. X=3

   (Una vez hecho se efectuará con garantías la simulación con llegadas de buses con capacidad aleatoria según los datos de vuestra práctica.)
4. **(1.5p)** Efectuar varias simulaciones para cada uno de l os f actores de carga especificados partiendo de diferentes semillas y r eportar estimaciones de l as magnitudes especificadas en l a sección 2.1. Reportad, al menos para dos de l os f actores de carga, histogramas de l as muestras de tiempos de espera.
5. **(1.5)** Para cada uno de l os f actores de carga r eportad el i ntervalo de confianza de W

PODRÁ ESCOGERSE CUALQUIER PROCEDIMIENTO DE SIMULACIÓN DE LOS EXPLICADOS EN CLASE (Event-Scheduling, Gillespie, Ciclos, ...). En el presente documento únicamente se describe el pseudocódigo para el metodo de Gillespie. Para cualquiera de los métodos empleados aplicarán los anteriores apartatdos 1 a 5

3.1. Entrega de la práctica

La entrega del trabajo se hará empaquetando en un .zip los ficheros que se especificarán y subiéndolos a ATENEA. En el .zip deberá haber:

1. Un informe que no deberá tener más de 20 páginas en el cuerpo principal, aunque puede contener apéndices si se considera necesario. Las páginas y figuras deberán estar numeradas y en la portada deberán figurar los nombres de cada uno de los autores y el conjunto de los parámetros con los que se haya efectuado el ejercicio.
2. Los ficheros que implementen los códigos desarrollados, tanto para la simulación como para el tratamiento de las trazas.
3. Las trazas de simulación obtenidas.
4. Una descripción sucinta de como deben ejecutarse los códigos desarrollados (puede figurar en el informe o aparte).

Apéndice 1. Código

$T_{ck} = \theta_A = \theta_B = \theta_C = 0, \;\; i = (0,0);$

Mientras $T_{ck} < T$ **hacer:**

- Generar $u_1, u_2 \sim unif[0,1]$
- Si $i = (n,0)$ ->$R_{\tau_C} = 1$, $h_{\tau_C} = 0$, **(C desactivada)**;
- Resolver: $u_1 = \dfrac{R_{\tau_A}(\theta_A + \Delta t) R_{\tau_B}(\theta_B + \Delta t) R_{\tau_C}(\theta_C + \Delta t)}{R_{\tau_A}(\theta_A) R_{\tau_B}(\theta_B) R_{\tau_C}(\theta_C)} \rightarrow \Delta t$
- $\lambda_A = h_{\tau_A}(\theta_A + \Delta t)$, $\lambda_B = \ldots$, $\lambda_C = \ldots$; $\Lambda = \lambda_A + \lambda_B + \lambda_C$
- $\pi_j = \lambda_j / \Lambda$, $j = A, B, C$
- Mediante u_2 seleccionar transición A, B, C.
- **Si** $\rightarrow$**(n,0): desactivar C;**
- Si A: $i = (n,\cdot) \rightarrow i = (n+1,\cdot)$; $\theta_A = 0$; $\theta_B = \theta_B + \Delta t$; ***Si* C:** $\theta_C = \theta_C + \Delta t$;
- Si $i = (n,0) \& B$: $i = (n,0) \rightarrow i = (n,c)$; $\theta_A = \theta_A + \Delta t$; $\theta_B = 0$; **Activar C**
- Si $i = (n,c) \& C$: $i = (n,c) \rightarrow i = (n-1, c-1)$; $\theta_A = \theta_A + \Delta t$; $\theta_B = \theta_B + \Delta t$; $\theta_C = 0$;
- Si $i = (n,c) \& B$: $\theta_C = \theta_C + \Delta t$; $\theta_A = \theta_A + \Delta t$; $\theta_B = 0$;

$T_{ck} = T_{ck} + \Delta t$

Fin Mientras

Figura 3: Código de simulación en detalle

Apéndice 2. Tiempo Δt de transición por método de la inversa

Si $\phi(t) = \prod_{j=1}^{N} \varphi_j(t) \rightarrow \phi'(t) = \phi(t) \displaystyle\sum_{j=1}^{N} \frac{\varphi'_j(t)}{\varphi_j(t)}$, $R_{s|\{\theta\}}(\Delta t) = \prod_{j=1}^{N} \dfrac{R_{\tau_j}(\theta_j + \Delta t)}{R_{\tau_j}(\theta_j)}$,

$$(\varphi'_j \equiv) \quad \frac{d}{d\Delta t}\left[\frac{R_{\tau_j}(\theta_j + \Delta t)}{R_{\tau_j}(\theta_j)}\right] = \frac{-f_{\tau_j}(\theta_j + \Delta t)}{R_{\tau_j}(\theta_j)} = -f_{s_j|\theta_j}(\Delta t)$$

$$\left(\frac{\varphi'_j}{\varphi_j} \equiv\right) \quad \frac{-f_{s_j|\theta_j}(\Delta t)}{\left[\dfrac{R_{\tau_j}(\theta_j + \Delta t)}{R_{\tau_j}(\theta_j)}\right]} = - \frac{f_{\tau_j}(\theta_j + \Delta t)}{R_{\tau_j}(\theta_j + \Delta t)} = -h_{\tau_j}(\theta_j + \Delta t)$$

$$R'_{s|\{\theta\}}(\Delta t) = -R_{s|\{\theta\}}(\Delta t) \sum_{j=1}^{N} h_{\tau_j}(\theta_j + \Delta t)$$

Para resolver $u = R_{s|\{\theta\}}(\Delta t)$

1. Partir de $\Delta t^{(0} = -\frac{1}{H^{(0}} \ell n\, u$, siendo $H^{(0} = \displaystyle\sum_{j=1}^{N} h_{\tau_j}(\theta_j)$; $i = 0$
2. Iteración del método de la tangente:

   $u = R_{s|\{\theta\}}(\Delta t^{(i)}) + R'_{s|\{\theta\}}(\Delta t^{(i)})(\Delta t^{(i+1} - \Delta t^{(i)})$

   $$\Delta t^{(i+1} = \Delta t^{(i} + \frac{u - R_{s|\{\theta\}}(\Delta t^{(i)})}{R'_{s|\{\theta\}}(\Delta t^{(i)})}$$

   Parar cuando $|R_{s|\{\theta\}}(\Delta t^{(i)}) - u| < \epsilon \cdot u$

☐ Se aconseja no hacer más de ~ 4 iteraciones para que computacionalmente no resulte muy gravoso.