Simulación a Eventos Discretos

Obligatorio 2

Informe

Andrés Baranzano 4.384.771-6

Emiliano Prigue 4.696.389-6

Ejercicio 8

Escribir un procedimiento para su biblioteca de simulación a eventos discretos, que permita obtener muestras de una distribución empírica. Testear la implementación con los datos de la figura 4.3 del libro del curso.

Se implementó un procedimiento un proyecto llamado Ejercicio_8 que permite obtener muestras de la distribución empírica obtenida a partir de los datos de la figura 4.3 del libro. Se obtienen muestras discretas y continuas, las cuales se imprimen en pantalla (con un resumen de las cantidades obtenidas en el caso de las muestras discretas). Este proceso se ejecuta 5 veces con 5 semillas distintas, esperando antes de comenzar una nueva iteración para poder visualizar los datos obtenidos en la ejecución anterior. Se realizó de esta form para poder comprobar que, aún cambiando la semilla, el muestreo obtenido es similar a los datos de la figura 4.3 del libro.

Ejercicio 9

Para el caso de estudio de transporte público:

A. **Considerar los arribos de pasajeros a su parada origen aleatorios, con distribución exponencial negativa con valor medio igual al inverso del valor indicado en la matriz origen-destino.**

Se actualizó la implementación del caso de estudio de transporte público realizado en el obligatorio anterior considerando los arribos de los pasajeros de la forma pedida. El proyecto actualizado se llama Ej9_pA, y considera solamente este cambio.

B. **Considerar los costos de las aristas del grafo de paradas aleatorios, con distribución normal con media igual al costo de la arista y desviación estándar igual a 1/5 del mismo.**

Se actualizó la implementación del caso de estudio de transporte público realizado en el obligatorio anterior considerando los costos de las aristas del grafo de la forma pedida. El proyecto actualizado se llama Ej9_pB, y considera solamente este cambio.

Ejercicio 10

Resolver el ejercicio 4.11.8 parte (a) del libro del curso.

TABLA 3-Distribución Chi Cuadrado χ^2. (Continuación)

v/p	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	0,95	0,975	0,99	0,995	0,9975	0,999
1	0,3573	0,2750	0,2059	0,1485	0,1015	0,0642	0,0358	0,0158	0,0039	0,0010	0,0002	0,0000	0,0000	0,0000
2	1,1957	1,0217	0,8616	0,7133	0,5754	0,4463	0,3250	0,2107	0,1026	0,0506	0,0201	0,0100	0,0050	0,0020
3	2,1095	1,8692	1,6416	1,4237	1,2125	1,0052	0,7978	0,5844	0,3518	0,2158	0,1148	0,0717	0,0449	0,0243
4	3,0469	2,7528	2,4701	2,1947	1,9226	1,6488	1,3665	1,0636	0,7107	0,4844	0,2971	0,2070	0,1449	0,0908
5	3,9959	3,6555	3,3251	2,9999	2,6746	2,3425	1,9938	1,6103	1,1455	0,8312	0,5543	0,4118	0,3075	0,2102
6	4,9519	4,5702	4,1973	3,8276	3,4546	3,0701	2,6613	2,2041	1,6354	1,2373	0,8721	0,6757	0,5266	0,3810
7	5,9125	5,4932	5,0816	4,6713	4,2549	3,8223	3,3583	2,8331	2,1673	1,6899	1,2390	0,9893	0,7945	0,5985
8	6,8766	6,4226	5,9753	5,5274	5,0706	4,5936	4,0782	3,4895	2,7326	2,1797	1,6465	1,3444	1,1042	0,8571
9	7,8434	7,3570	6,8763	6,3933	5,8988	5,3801	4,8165	4,1682	3,3251	2,7004	2,0879	1,7349	1,4501	1,1519
10	8,8124	8,2955	7,7832	7,2672	6,7372	6,1791	5,5701	4,8652	3,9403	3,2470	2,5582	2,1558	1,8274	1,4787

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{para } x < 0 \\ 1 - e^{-\lambda x} & \text{para } x \geq 0 \end{cases}$$

Lambda	(cantObservaciones)/(tiempoObservado)	0,02969697
TOTAL	Suma de los Oj = n	196

Oj : Cantidad de datos observados en la categoría j
Pj : Probabilidad teórica de la categoría j
Ej : Cantidad de datos esperados en la categoría j, Ej = nPj

Estadístico = Sumatoria(((Oj-Ej)^2)/Ej)

Datos originales

S	Oj	Aux	Pj	Ej	((Oj-Ej)^2)/Ej
0 - 10	60	0,25693347	0,25693347	50,3589603	1,845741977
10 - 20	42	0,44785213	0,19091866	37,4200578	0,56055152
20 - 30	23	0,5897174	0,14186527	27,8055925	0,830542245
30 - 40	18	0,69513273	0,10541533	20,6614051	0,342816817
40 - 50	14	0,77346334	0,0783306	15,3527986	0,119200678

50 - 60	10	0,83166819	0,05820485	11,4081508	0,173813319
60 - 70	7	0,87491827	0,04325008	8,47701499	0,257351588
70 - 80	5	0,90705595	0,03213768	6,2989861	0,267878809
80 - 90	8	0,93093639	0,02388044	4,68056574	2,35412649
90+	9	1	0,06906361	13,5364681	1,520303739

Estadístico = 8,27232718, **Cantidad de categorías** = 10, **Grados de libertad** = 9

Comprobando el estadístico en la tabla, vemos que acepta la hipótesis con una probabilidad menor a 0,55, por lo que intentaremos cambiar las categorías para ver si lo podemos mejorar.

Ajuste de categorías 1

S	Oj	Aux	Pj	Ej	((Oj-Ej)^2)/Ej
0 - 10	60	0,25693347	0,25693347	50,3589603	1,845741977
10 - 30	65	0,5897174	0,33278393	65,2256503	0,000780645
30 - 40	18	0,69513273	0,10541533	20,6614051	0,342816817
40 - 50	14	0,77346334	0,0783306	15,3527986	0,119200678
50 - 60	10	0,83166819	0,05820485	11,4081508	0,173813319
60 - 70	7	0,87491827	0,04325008	8,47701499	0,257351588
70 - 80	5	0,90705595	0,03213768	6,2989861	0,267878809
80+	17	1	0,09294405	18,2170339	0,081306949

Estadístico = 3,08889078, **Cantidad de categorías** = 8, **Grados de libertad** = 7

Comprobando el estadístico en la tabla, vemos que acepta la hipótesis con una probabilidad entre 0,85 y 0,90. Es una probabilidad bastante buena, pero seguiremos intentando mejorarla.

Ajuste de categorías 2

S	Oj	Aux	Pj	Ej	((Oj-Ej)^2)/Ej
0 - 30	125	0,5897174	0,5897174	115,584611	0,766966788

30 - 40	18	0,69513273	0,10541533	20,6614051	0,342816817
40 - 50	14	0,77346334	0,0783306	15,3527986	0,119200678
50 - 60	10	0,83166819	0,05820485	11,4081508	0,173813319
60 - 70	7	0,87491827	0,04325008	8,47701499	0,257351588
70 - 80	5	0,90705595	0,03213768	6,2989861	0,267878809
80+	17	1	0,09294405	18,2170339	0,081306949

Estadístico = 2,00933495, **Cantidad de categorías** = 7, **Grados de libertad** = 6

Comprobando el estadístico en la tabla, vemos que acepta la hipótesis con una probabilidad entre 0,90 y 0,95. Se mejoró la probabilidad ajustando un poco las categorías, veremos si podemos mejorarla.

Ajuste de categorías 3

S	Oj	Aux	Pj	Ej	((Oj-Ej)^2)/Ej
10 - 40	143	0,69513273	0,69513273	136,246016	0,334808351
40 - 50	14	0,77346334	0,0783306	15,3527986	0,119200678
50 - 60	10	0,83166819	0,05820485	11,4081508	0,173813319
60 - 70	7	0,87491827	0,04325008	8,47701499	0,257351588
70 - 80	5	0,90705595	0,03213768	6,2989861	0,267878809
80+	17	1	0,09294405	18,2170339	0,081306949

Estadístico = 1,23435969, **Cantidad de categorías** = 6, **Grados de libertad** = 5

Comprobando el estadístico en la tabla, vemos que acepta la hipótesis con una probabilidad entre 0,90 y 0,95. No se mejoró la probabilidad de que acepte la hipótesis y si seguimos juntando categorías tendríamos menos de 5 grados de libertad, lo que no es recomendable. Por lo que concluímos que tanto con el ajuste de categorías 2 como con el ajuste de categorías 3 realizados, la probabilidad de que acepte la hipótesis y ajuste a una distribución negativa exponencial está entre 0,90 y 0,95.

Ejercicio 11

Resolver el ejercicio 5.9.3 del libro del curso.

A. Se realizaron distintas ejecuciones, variando la media de la distribución de los tiempos de estadía, como el ejercicio sugería probar con valores entre 48 y 120 horas, nos pareció razonable utilizar los valores 48, 84 y 120 horas y los resultados fueron los siguientes:

Tiempo	Run-in	Largo Promedio	Espera Promedio
48	0	[]	[]
84	7500	[]	[]
120	12500	[]	[]

Al realizar una ejecución, se genera el archivo HistogramaColaEnTiempoXXX.txt donde “XXX” corresponde al tiempo de espera medio seleccionado. Analizando ese archivo, se obtienen los valores de run-in especificados anteriormente.

B. Al realizar una ejecución, se generan los archivos HistogramaTEsperaXXX.txt HistogramaLColaXXX.txt donde donde “XXX” corresponde al tiempo de espera medio seleccionado.

C. El proyecto Ejercicio_11 realiza 100 ejecuciones del sistema , y al finalizar, se muestran en pantalla los intervalos de confianza del largo de la cola promedio y del tiempo de espera promedio. Dichos valores son los reportados en la parte A, para los cuales se utilizó una distribución normal, ya que todas las ejecuciones son intependientes

Ejercicio 12

Planificar el análisis de resultados para el caso de estudio de transporte público, de la siguiente forma:

A. Identificar las medidas de salida apropiadas.

Una salida apropiada para nuestro programa sería, analizando el caso de estudio, una que mostrara la cantidad de flota total en el sistema (ómnibus operando en forma simultánea en el sistema) y el tiempo de espera medio de los pasajeros en para cada par [parada origen][parada destino] en el sistema, así como tiempo de espera medio global del sistema. Aunque también podría ser de interés mostrar estadísticas sobre la cantidad de asientos libres/ocupados por línea, lagro promedio de la cola de pasajeros para cada parada, como también la cantidad de ómnibus por línea.

B. Determinar qué histogramas deben construirse.

- Para el tiempo de espera medio de los pasajeros en cada par [parada origen][parada destino] se construirá un histograma de tipo **Observation** grabando el valor (ST. getSimTime() - Pasajero.getClock()) al momento de subir al Omnibus**.**
- Para el tiempo de espera medio global del sistema se hará de la misma forma al punto anterior.
- Para la cantidad de asientos libres/ocupados se utilizará un histograma de tipo **TimeWeighted** registrando la cantidad de asientos libres/ocupados parada a parada para cada Omnibus de la Linea.
- Para el largo promedio de la cola de pasajeros para cada parada se utilizará un histograma de tipo **TimeWeighted** registrando el largo de la cola de dicha parada

C. Identificar las variables de decisión y sugerir qué análisis deben realizarse en las diferentes replicaciones de la simulación.

Las variables de decisión en nuestra realidad son las frecuencias de cada línea, por lo que los análisis de los resultados obtenidos deberían estar orientados a llevar a conclusiones sobre dichas frecuencias.
Primero se podría hacer un análisis global del sistema, considerando el tiempo de espera global de los pasajeros del sistema y la cantidad de flota total del sistema, en donde si el tiempo es considerablemente bajo, se podría concluir que las frecuencias tienen un valor correcto. De lo contrario, si el tiempo no se considera aceptable, se podrá concluir que las frecuencias deberían aumentar por lo que el tamaño de la flota también lo hará.
Un problema que puede no estar resolviendo este análisis es que, a pesar de que el tiempo de espera global sea aceptado, el tiempo de espera para alguna [parada origen][parada destino] puede ser malo, por lo que habría que ver el tiempo de espera para cada [parada origen][parada destino] y luego ver a que líneas dichos pasajeros se podrían subir y de esa forma encontrar una solución posible a esa espera, ya sea aumentando alguna frecuencia en particular, o el conjunto de frecuencias de todas las líneas encontradas.
Ahora, para la solución comentada en el párrafo anterior, estaría bueno analizar estadísticas sobre la cantidad de asientos libres/ocupados para cada línea, ya que de esa forma, ayudaría decidir de cuáles líneas es mejor aumentar la frecuencia y de cuáles no. Es decir, si una línea se encuentra la mayoría del tiempo con asientos ocupados, podría ser recomendable aumentar la frecuencia de dicha línea en lugar de la de una que viaja la mayor parte del tiempo con asientos libres.
Todas estas decisiones afectarán en la cantidad de ómnibus por línea (y por lo tanto la flota total) en el sistema, por lo que sería de interés, luego de ir modificando frecuencias, estudiar la cantidad de flota por línea necesaria para los tiempos aceptados, ya que quizás tener dicha cantidad de flota es impracticable, por lo que habría que volver a realizar los análisis hasta encontrar una solución aceptada.

Ejercicio 13

Para el caso de estudio de transporte público:

A. **Preparar una especificación.**

Objetivos

Estudiar el impacto de modificaciones de las frecuencias en los tiempos de espera de los pasajeros y en el tamaño de la flota requerida (la cantidad de ómnibus operando en forma simultánea en el sistema).

Hipotesis de Trabajo

1. Cada línea utiliza la cantidad de ómnibus que sea necesaria para realizar los recorridos con la frecuencia establecida.
2. Un mismo ómnibus está siempre asignado a la misma línea.
3. Para líneas con recorridos de ida y vuelta, un ómnibus realiza estos dos recorridos sin ninguna pausa entre ellos y al regresar al punto de partida espera la hora de la próxima salida, de acuerdo a la frecuencia de la línea.
4. Para líneas con recorridos circulares, existe una parada inicial, que cumple el rol de la primera parada del recorrido de ida, a los efectos de las esperas para la próxima salida.
5. Un pasajero esperando en una parada sube al primer ómnibus que pasa por la misma, cuyo recorrido incluye a su parada destino, en la sección que resta por recorrer. Además un pasajero siempre podrá subir a dicho ómnibus
6. En recorridos circulares, los pasajeros pueden bajarse en cualquier parada luego de subirse al bus
7. Los ómnibus no se rompen.
8. Los ómnibus no se quedan sin combustible.
9. El costo de ir de una parada a otra no se ve afectada por tráfico, semáforos, etc.
10. La cantidad de pasajeros dentro de un mismo ómnibus es ilimitada.
11. La frecuencia no se ve afectada por ninguna variable.
12. La subida y bajada de pasajeros consume 0 unidades de tiempo.
13. Los arribos de los pasajeros se describen mediante una Distribucion Exponencial Negativa con Lambda definido para cada [origen][desitno].
14. Los costos para ir de una parada i a una parada j se describen mediante una Distribución Normal, con una media y una desviación estándar dada para cada recorrido i-j

Variables de Decision

Frecuencia de salida de los Ómnibus.

Respuestas

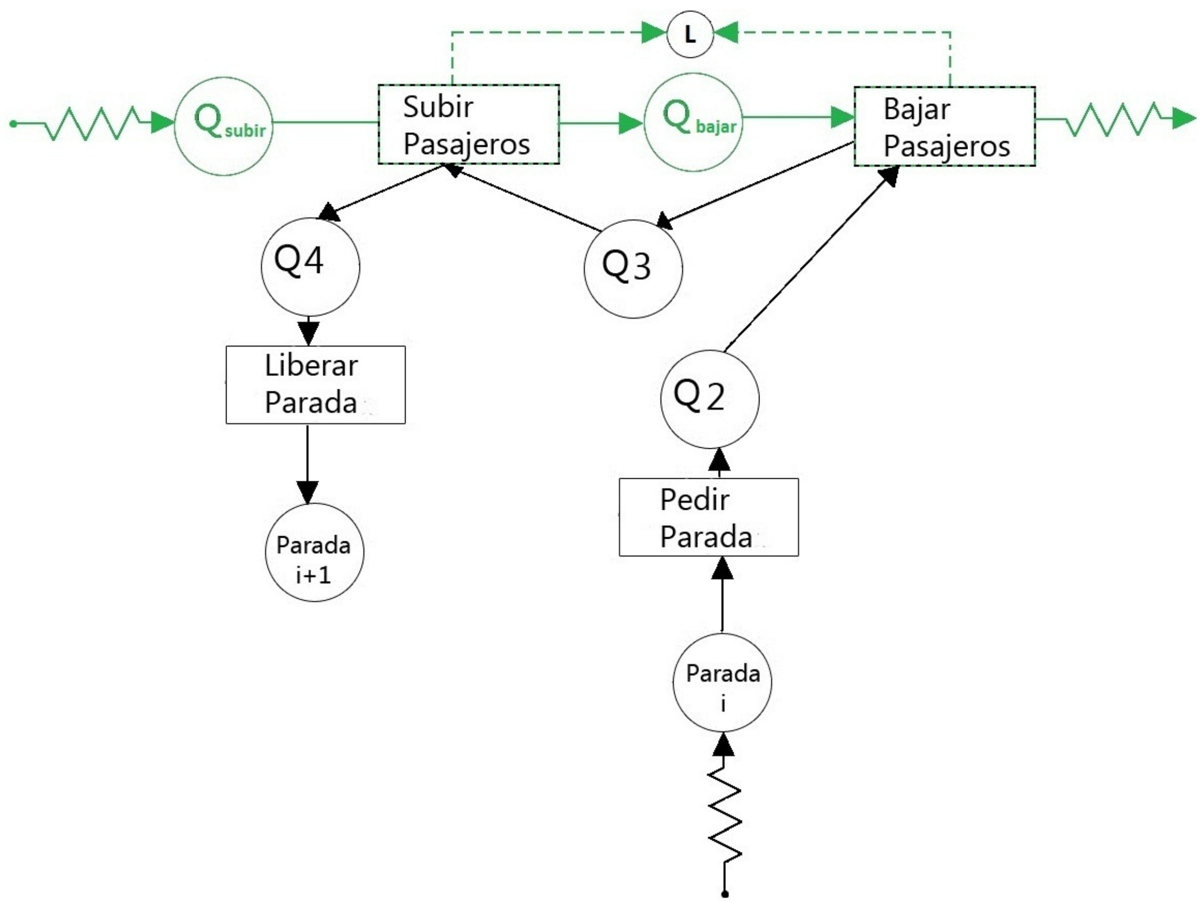
Tiempo de espera de pasajeros y tamaño de flota requerida para cada línea, dado un conjunto de frecuencias.

Duraciones de actividades

Los valores de duración de las actividades están descritos en minutos.

- Los arribos de los pasajeros se describen, para un pasajero con parada de origen i y parada de destino j, mediante una Distribucion Exponencial Negativa con Lambda 1/ValorIJ. Donde ValorIJ es el correspondiente al valor de frequencia leído del archivo matrizOD.txt.
- Los costos para ir de una parada i a una parada j se describen mediante una Distribución Normal (costoo) , con una media igual al valor i-j del archivo conexciones.txt (en metros) y una desviación estándar de media/5. Por lo que la duración corresponde a (costo / 13000) * 60
- El tiempo entre la salida de un ómnibus de una línea y la siguiente salida de un ómnibus de la misma línea es 1/frecuenciaLinea, donde frecuenciaLínea es la frecuencia obtenida del archivo recorridos.txt para la línea correspondiente.

Diagramas de actividades



B. Implementar un programa que ejecute lo especificado.

Se actualizó la implementación del Sistema de transporte público entregada en el laboratorio anterior, agregando las 2 modificaciones hechas en el ejercicio 9 y un histograma de tipo Observation para el tiempo de espera medio global.

C. Analizar los resultados a partir de varias replicaciones. Discutir según las hipótesis realizadas en la especificación.

Se realizaron 5 ejecuciones independientes de las cuales se recuperaron los datos observados en los archivos “LogCasoDeEstudioX.txt” y “tiempoEsperaGlobalX.txt” (siendo X el número de la ejecución) para poder llevar a cabo el análisis. Se puede observar que la media de los tiempos de espera son relativamente bajos, rondando los 13 minutos, lo cual para el mundo real no sería para nada problemático, por lo que en principio se podría decir que las frecuencias dadas en un principio podrían ser viables. Por lo que el paso siguiente sería mirar el tamaño de la flota, que en promedio está compuesta por 37 ómnibus aproximadamente, lo que no sería nada alto considerando que existen 84 paradas y las distribuciones de éstas en el mapa, así que ésta solución de frecuencias seguiría siendo viable.
Además, se observan una cantidad relativamente significativa de pasajeros con tiempos de espera de entre 30 y 60 minutos, y considerando que la media más baja utilizada para el tiempo de salida entre ómnibus (la cuál se repite en 5 líneas distintas de las 13), es de casi una hora, dichas esperas podrían llegar a entrar dentro de un rango razonable.
El problema, analizando más de cerca los datos, es que existieron 4 pasajeros que esperaron más de 90 minutos promedialmente, lo cual podría ser una falta grave en la realidad. Igualmente, sí esos 4 fueran los únicos casos, quizás no habría problema; sin embargo existen más de un valor con tiempos de espera altos (mayores a 1 hora), por lo cuál no es un problema despreciable. Considerando que el tiempo máximo de salida entre dos ómnibus es de 1 hora para las líneas con menos frecuencias, si un pasajero tiene un tiempo de espera mayor a dicho valor, implica que existió un ómnibus que pasó por su parada y que la parada de destino de dicho pasajero estaba en su recorrido y en el cual no pudo subirse. Esto quiere decir que el ómnibus estaba lleno, y que probablemente sea necesario aumentar la frecuencia de algunas líneas.

En conclusión, lo que habría que hacer es estudiar un poco más a fondo qué es lo que sucede en cada parada, cuál es la parada de origen y la parada de destino de los pasajeros con mayor tiempo de espera y ver de qué líneas habría que aumentar la frecuencia para solucionar éste problema (si es considerado un problema por el Departamento de Tránsito y Transporte). Para tomar esta decisión, podría ser útil el histograma del cuál se habló en el ejercicio 12 parte B, el que sugería un histograma (del tipo TimeWeighted) para la cantidad de asientos libres/ocupados por línea.