

Evaluación Final de Simulación de Sistemas

45.288 Brasca, Juan Alejandro
45.020 Stancato, Lucila
45.002 Modernell, Damián
45.418 Negro, Conrado Luis

a) Modelar el intervalo de tiempo entre arribos, a partir de los datos medidos del archivo *llegadasregistro* y modelar el intervalo de tiempo de atención de la estación E3 con los datos del archivo *e3registro*. Para ello hacer estadística descriptiva, justificando el número de intervalo de clases y realizar los test pertinentes.

Tiempo entre arribos al sistema

Mediante mediciones del horario de arribo de clientes al sistema, computamos los intervalos de tiempo entre arribos, y los representamos en el histograma de la figura 1. Está dividido en 7 intervalos de clase determinados a partir de la fórmula de Sturges que observamos en la ecuación 1.

$$k = 1 + 3.3 \text{Log}(n) \quad (1)$$

donde k es el número de intervalos y n es el número de muestras.

Como vemos en la figura 1, el tiempo entre llegadas de clientes parece ser una variable aleatoria con distribución exponencial. Para determinar si esta suposición es cierta, realizamos el test de bondad de ajuste χ^2 . Las hipótesis del test son:

- $H_0: \chi_0^2 < \chi_{n-1, \alpha}$ (Las llegadas de clientes al sistema están exponencialmente distribuidas)
- $H_1: \chi_0^2 \geq \chi_{n-1, \alpha}$ (Las llegadas de clientes al sistema NO están exponencialmente distribuidas)

A partir de las mediciones y de los valores esperados para cada intervalo de clase, computamos el estadístico $\chi_0^2 = 10.709$. Como el valor crítico con un nivel de significación de 5% y con 5 grados de libertad resulta $\chi_{0.05, 5}^2 =$

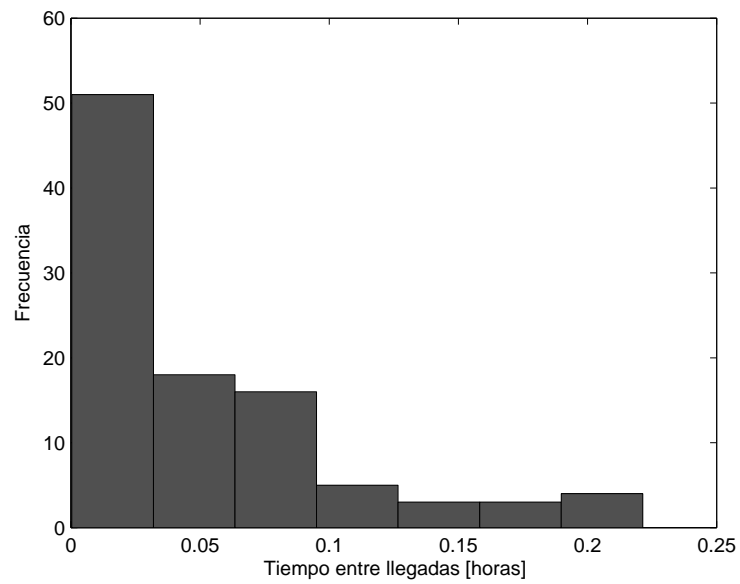


Figure 1: Histograma de intervalos de tiempo entre arribos de clientes al sistema. Se puede observar que parece provenir de una variable aleatoria con distribución exponencial.

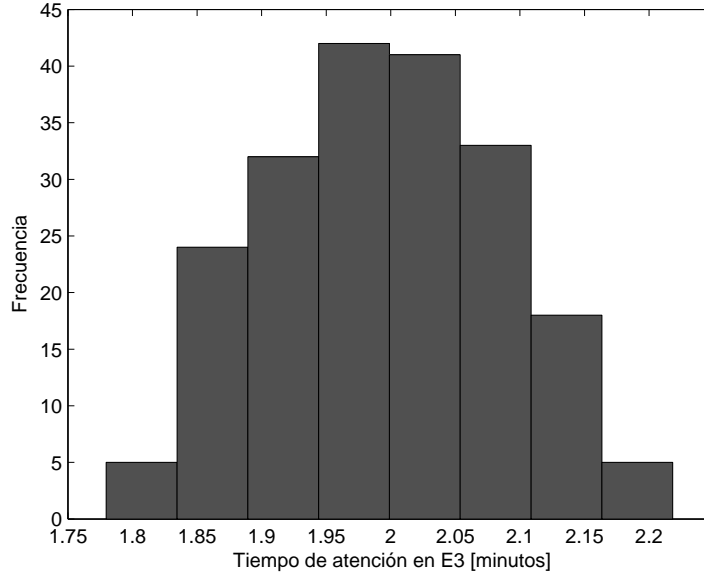


Figure 2: Histograma de la medición de los tiempos de servicio del servidor *E3*. Observamos que los datos parecen provenir de una variable aleatoria con distribución normal.

11.07, rechazamos H_1 en favor de H_0 .

También realizamos el test de Kolmogorov-Smirnov sobre las mediciones de los tiempos de arribos con las mismas hipótesis que el test χ^2 . Obtenemos el estadístico para dichos datos $D = 0,1238$, con un valor crítico para un nivel de significación $\alpha = 5\%$, que es $D_{5,0.05} = 0.4834$. Luego al ser $D < D_{t,\alpha}$, se rechaza H_1 en favor de H_0 .

Habiendo superado ambos tests, podemos suponer que los arribos de clientes al sistema están exponencialmente distribuidos, con una tasa media de 18.825.

Tasa de servicio de la estación *E3*

Analizamos los tiempos de servicio de la estación *E3* medidos del sistema, y volcamos los datos a un histograma (ver figura 2) con 8 intervalos de clase también computados a partir de la fórmula de Sturges.

En la figura 2, observamos que los datos parecen provenir de una variable aleatoria con distribución normal. Para poder determinar si esto es cierto, realizamos el test estadístico χ^2 con las siguientes hipótesis:

- H_0 : $\chi_0^2 < \chi_{n-1,\alpha}$ (El tiempo de servicio de la estación *E3* es una variable aleatoria con distribución normal)

- $H_1: \chi_0^2 \geq \chi_{n-1, \alpha}$ (El tiempo de servicio de la estación $E3$ NO es una variable aleatoria con distribución normal)

Estimamos la media y la varianza de la distribución resultando:

- $\mu = 1.995$
- $\sigma^2 = 0.008332$

El valor del estadístico resulta $\chi_0^2 = 4,7492$, y como el valor crítico es $\chi_{5,0.05}^2 = 11.07$ rechazamos H_1 en favor de H_0 .

b) Indicar las variables de estado del sistema y el espacio de estados.

Considerando el gráfico de la figura 3, definimos el espacio de estados S :

$$S = \{(x_R, x_{E_1}, x_{OFT}, x_{PSF}, x_{E_2}, x_{E_3}, x_{C_1}, x_{C_2}, x_{C_3}, y_R, y_{E_1}, y_{OFT}, y_{PSF}, y_{E_2}, y_{E_3}, y_{C_1}, y_{C_2}, y_{C_3}) \\ x_R, x_{E_1}, x_{OFT}, x_{PSF}, x_{E_2}, x_{E_3}, x_{C_1}, x_{C_2}, x_{C_3} = 1, 2, 3, \dots \\ \wedge y_R, y_{E_1}, y_{OFT}, y_{PSF}, y_{E_2}, y_{E_3}, y_{C_1}, y_{C_2}, y_{C_3} = 0, 1\}$$

Donde:

x_R es la longitud de la cola de recepción.
 x_{E_1} es la longitud de la cola del servidor E1.
 x_{E_2} es la longitud del servidor E2.
 x_{E_3} es la longitud del servidor E3.
 x_{OFT} es la longitud de la cola de oftalmología.
 x_{PSF} es la longitud de la cola del estudio Psico-Físico.
 x_{C_1} es la longitud de la cola de la caja C1.
 x_{C_2} es la longitud de la cola de la caja C2.
 x_{C_3} es la longitud de la cola de la caja C3.

y_R Ocupación de la recepción.
 y_{E_1} Ocupación de la estación E1.
 y_{E_2} Ocupación de la estación E2.
 y_{E_3} Ocupación de la estación E3.
 y_{OFT} Ocupación del consultorio oftalmológico
 y_{PSF} Ocupación del consultorio Psico-Físico.
 y_{C_1} Ocupación de la caja 1.
 y_{C_2} Ocupación de la caja 2.
 y_{C_3} Ocupación de la caja 3.

Sabiendo el valor de cada una de estas variables, se puede determinar el estado del sistema completo.

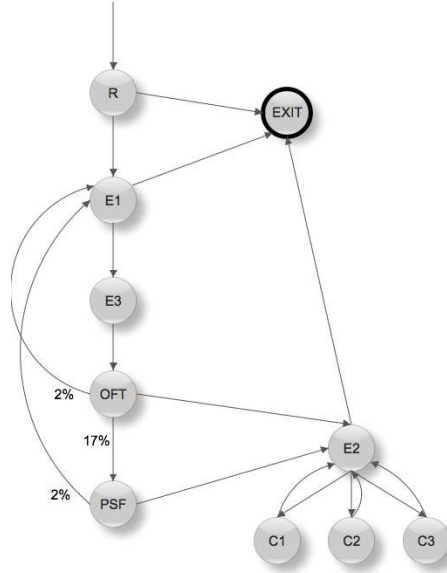


Figure 3: Modelo del sistema de renovación del registro.

c) Indicar los tipos de eventos y el espacio de eventos.

El espacio de eventos E del sistema lo definimos como:

$$E = \{A, D, r_{e1}, e1_{e3}, e3_{oft}, oft_{psf}, oft_{e2}, psf_{e2}, oft_{e1}, psf_{e1}, e2_{c1}, e2_{c2}, e2_{c3}, c1_{e2}, c2_{e2}, c3_{e3}\}$$

Donde:

A es la entrada de una persona a R .

D es la partida de una persona del sistema (ya sea desde R , E_1 o E_2).

r_{e1} es la partida de una persona desde R hacia E_1

$e1_{e3}$ es la partida de una parsona de E_1 hacia E_3

$e3_{oft}$ es la partida de una persona de E_3 hacia OFT

oft_{psf} es la partida de una persona de OFT hacia PSF .

oft_{e2} , psf_{e2} son las partidas de una persona desde OFT o desde PSF hacia E_2 .

oft_{e1} , psf_{e1} son las partidas desde OFT o de PSF hacia E_1

$e2_{c1}$, $e2_{c2}$ y $e2_{c3}$ son las partidas desde E_2 hacia C_1 , C_2 y C_3 .

$c1_{e2}$, $c2_{e2}$ y $c3_{e2}$ son las partidas desde C_1 , C_2 y C_3 hacia E_2 .

d) Realizar una simulación indicando las condiciones iniciales, el tipo de generador usado y la semilla. Mostrar las gráficas de las colas en cada estación.

Para generar números pseudoaleatorios, usamos el generador de L'Ecuyer con semillas 50 y 23 (recordando que lo importante es que las semillas no sean cero). El generador de L'Ecuyer genera una secuencia de números pseudoaleatorios con distribución uniforme entre 0 y 1. A partir de esa secuencia, generamos secuencias de números pseudoaleatorios con distribución exponencial. Para generar una secuencia de números pseudoaleatorios con distribución normal usamos el generador de Box-Muller (usando el generador de L'Ecuyer para generar la secuencia que tiene distribución uniforme). Con estos generadores, podemos obtener realizaciones de las variables necesarias para simular el modelo.

Para hacer la simulación consideramos que el porcentaje de personas con más de 70 años es de 7%. Esto hace que la probabilidad de que un cliente haga el examen de aptitud psicofísica pase de un 10% (que renuevan el registro profesional) a un 17%. Consideramos también que un 20% de los clientes vienen solamente a buscar el instructivo para hacer el trámite y se van, y solamente el 80% hace el trámite de renovación. Además suponemos que un cliente tarda entre 0.5 y 2 minutos en llenar el formulario que le entregan en la estación E_1 , y lo simulamos con una variable pseudo-aleatoria uniformemente distribuida. Para las cajas C_1 , C_2 y C_3 , asumimos que los clientes se encolan en la caja que tenga menos personas, siempre optando primero por C_1 , luego por C_2 y como última opción por C_3 .

Con estas consideraciones, realizamos una simulación que termina a las 16:56 hs y atiende a 128 clientes.

En las figuras 4 a 12 podemos observar los estados de los servidores durante la simulación. Cuando el estado es 0, la estación se encuentra desocupada, mientras que cuando vale 1 la estación se encuentra atendiendo a un cliente.

En las figuras 13 a 16 se puede observar la longitud de cada cola en función del tiempo. Como las estaciones E_1 , PSF , C_1 , C_2 y C_3 nunca tienen clientes en sus colas, no los graficamos. E_1 no tiene clientes en su cola porque el tiempo medio de atención es de 3 segundos, y los clientes llegan a un ritmo menor desde la recepción que atiende clientes con tiempo medio de entre 5 y 30 segundos. La estación PSF nunca tiene cola ya que recibe solo al 17% de los clientes que ingresan al sistema, y tiene tiempo de atenderlos sin acumular clientes en la misma. Por último, al tener tres cajas, la longitud de las colas esta dividida entre las tres, y el tiempo de atención de cada caja es menor que el tiempo de arribo de clientes a este sector.

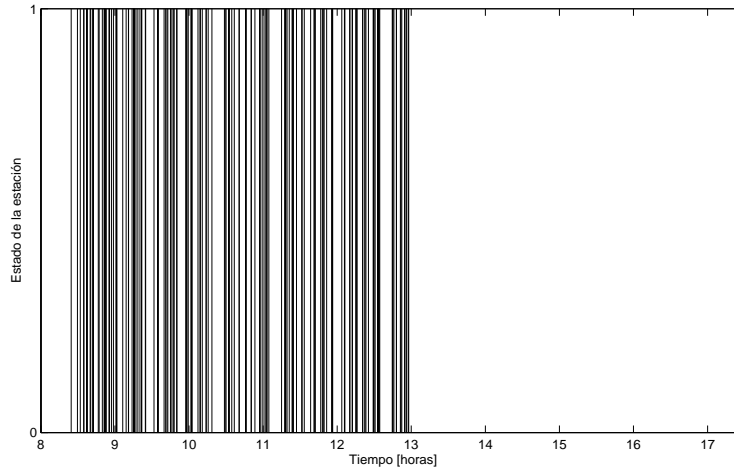


Figure 4: Estado de la recepción en función del tiempo. El primer cliente llega a las 8hs y las puertas se cierran a las 13hs.

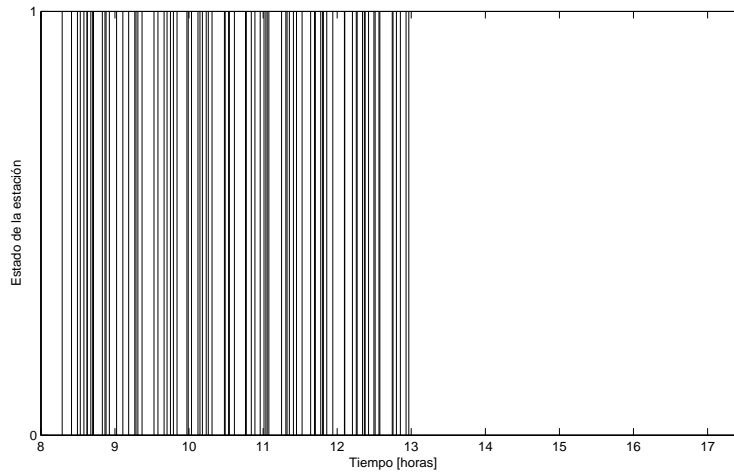


Figure 5: Estado de la estación E_1 en función del tiempo. Se puede observar que aproximadamente a las 13hs, cuando no llegan más clientes, la estación E_1 cesa sus actividades.

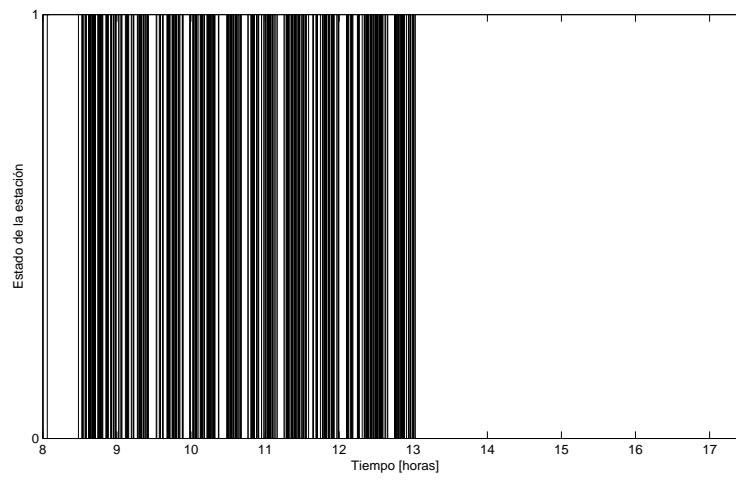


Figure 6: Estado de la estación E_3 en función del tiempo.

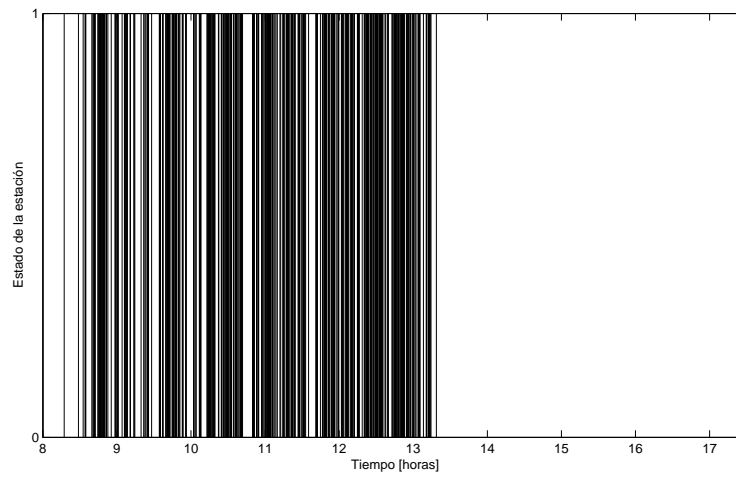


Figure 7: Estado del consultorio OFT en función del tiempo.

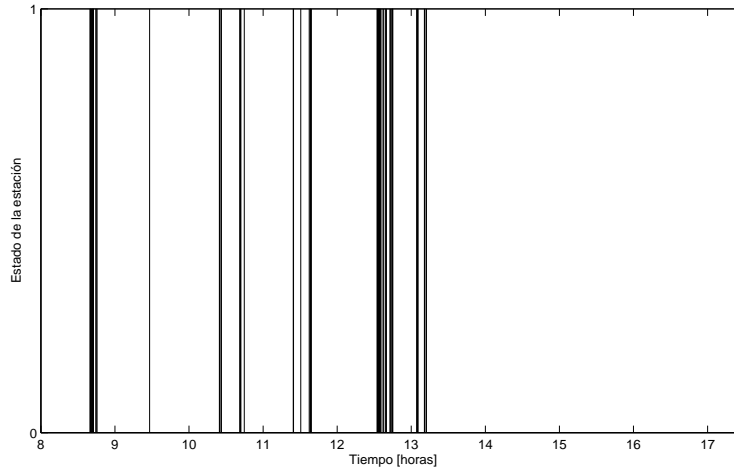


Figure 8: Estado del consultorio PSF en función del tiempo. Esta estación está visiblemente menos ocupado que las demás ya que sólo el 17% de los clientes renuevan un registro profesional o son mayores de 70 años.

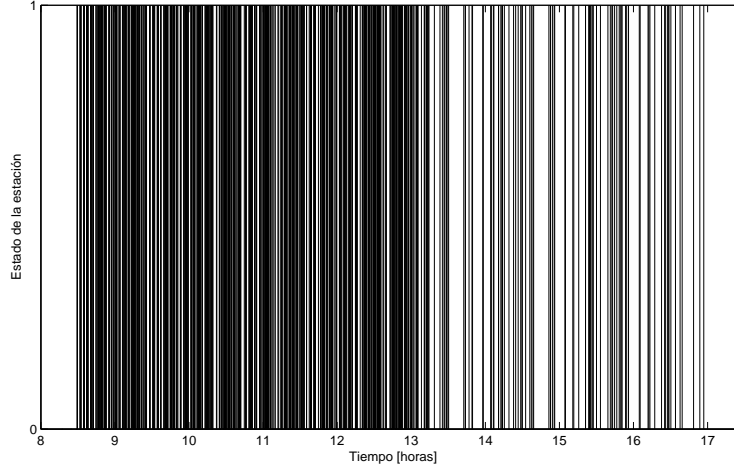


Figure 9: Estado de la estación E_2 en función del tiempo. Esta es la estación más ocupada del sistema, ya que tiene un alto tiempo de atención en comparación con los otros, y además, recibe dos veces a cada cliente (cuando llegan de OFT o PSF , y cuando llegan desde las cajas C_1 , C_2 o C_3).

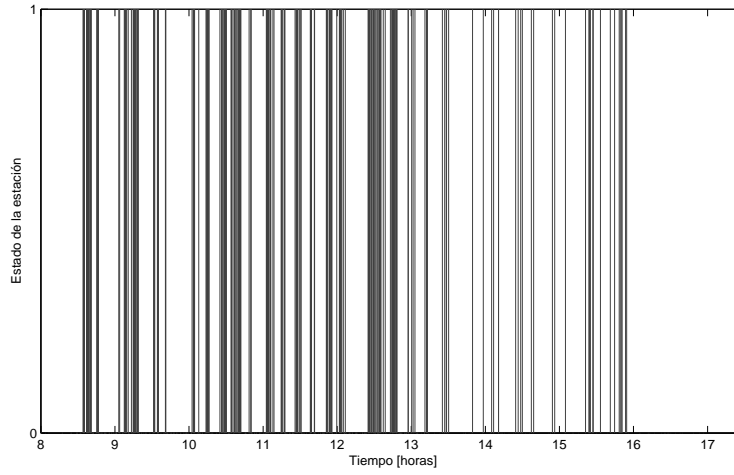


Figure 10: Estado de la caja C_1 en función del tiempo. Esta es la caja más usada de las tres ya que los clientes la tienen como primer opción a elegir cuando todas las cajas tienen una fila de la misma longitud. Si se encontrara ocupada, probarían con la caja C_2 , y luego con C_3 .

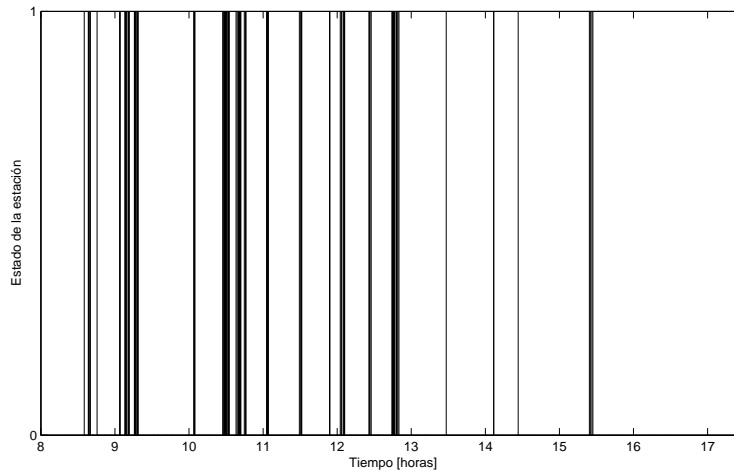


Figure 11: Estado de la caja C_2 en función del tiempo. Esta es la segunda caja más usada, ya que es elegida sólo si la caja C_1 tiene una fila más larga que la caja C_2 .

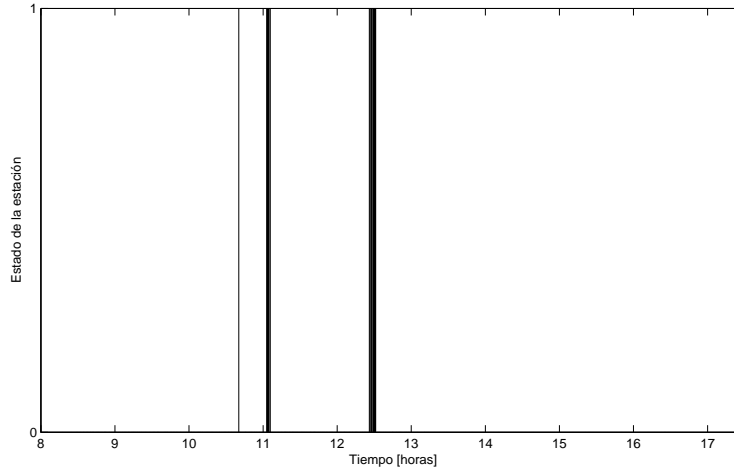


Figure 12: Estado de la caja C_3 en función del tiempo. Se puede observar el poco uso que los clientes hacen de esta caja, ya que C_1 y C_2 cumplen con casi todo el trabajo.

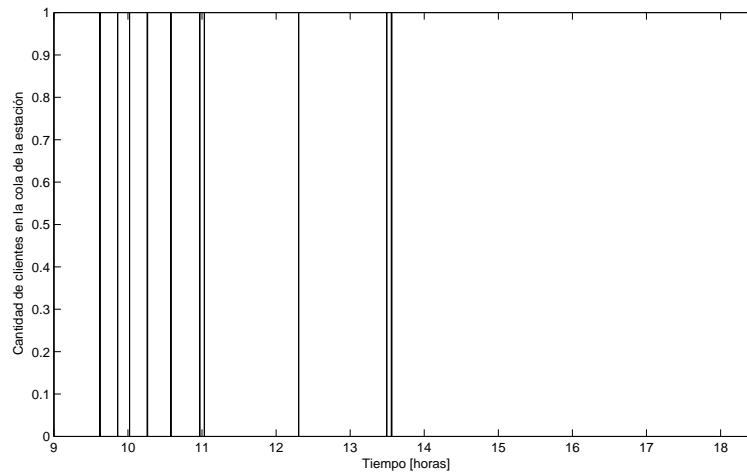


Figure 13: Cantidad de clientes en la cola de la recepción en función del tiempo.

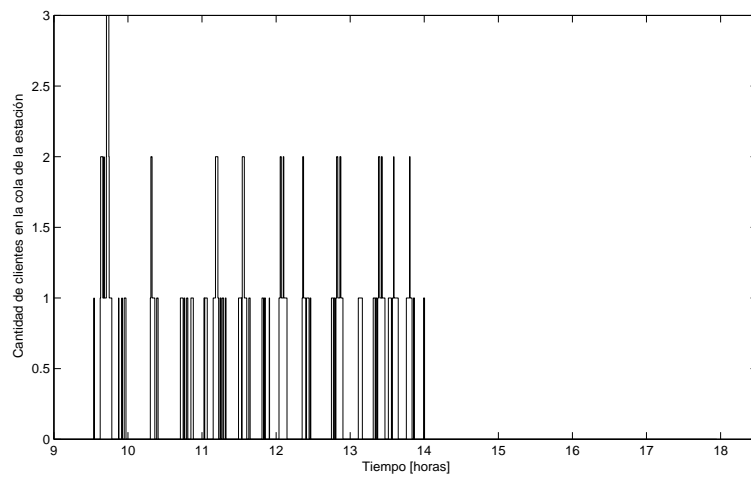


Figure 14: Cantidad de clientes en la cola de la estación E_3 en función del tiempo.

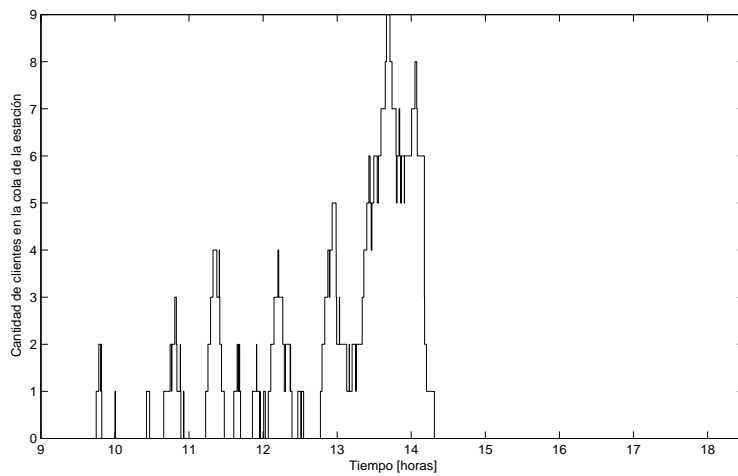


Figure 15: Cantidad de clientes en la cola del consultorio OFT en función del tiempo.

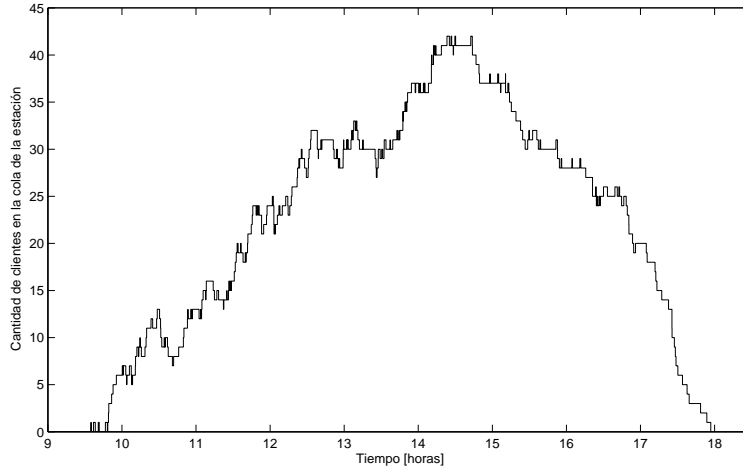


Figure 16: Cantidad de clientes en la cola de la estación E_2 en función del tiempo. Como esta es la estación más ocupada del sistema, la cola que se genera hacia el final es la más larga del sistema, llegando hasta 42 clientes.

e) Realizar 10 simulaciones independientes. Hallar el tiempo medio por cliente en el sistema, suponiendo que las tres cajas C1, C2 y C3 están operativas mostrar los resultados convenientemente en una tabla. Computar con estos resultados el estimador del tiempo medio por cliente y su error.

Realizamos 10 simulaciones con diferentes semillas. En cada simulación se mide el tiempo medio \bar{t} de un cliente en el sistema, la cantidad N de clientes, y la hora t_{final} a la que finaliza la simulación. Los resultados se muestran en la tabla 1.

A partir de las simulaciones realizadas computamos la media muestral del tiempo medio de un cliente en el sistema, y obtenemos $\bar{t} = 1 : 53$ hs, con un desvío muestral de $\sigma = 30.3369$. El error de la estimación lo calculamos haciendo $\sigma_t = \sigma/\sqrt{(n)}$, lo que resulta en $\sigma_t = 9.5934$ minutos, es decir, un error del 8.5%.

Tabla 1: Simulaciones con distintas semillas

| Semilla 1 | Semilla 2 | \bar{t} [hs] | N | t_{final} [hs] |
|-----------|-----------|----------------|-----|------------------|
| 50 | 23 | 1:45 | 128 | 16:56 |
| 150 | 263 | 1:50 | 125 | 16:39 |
| 18 | 639 | 1:17 | 119 | 15:39 |
| 9018 | 78639 | 2:40 | 122 | 17:39 |
| 19018 | 278639 | 1:13 | 129 | 16:10 |
| 111 | 223 | 2:16 | 127 | 17:03 |
| 41114 | 82238 | 2:37 | 123 | 17:59 |
| 4199114 | 9 | 1:35 | 132 | 16:08 |
| 9 | 4199114 | 1:58 | 141 | 17:25 |
| 90 | 70 | 1:39 | 119 | 16:24 |

f) Si la probabilidad de que la caja C3 esté operativa es p , computar el tiempo medio por cliente en el sistema en función de p . Graficar.

Dada p , la probabilidad de que la caja C_3 esté operativa, realizamos distintas simulaciones para estimar el tiempo medio de un cliente en el sistema. Los resultados de las simulaciones se muestran en la figura 17. Vemos que si la probabilidad de que la caja C_3 esté abierta, es menor que 0.9, entonces el tiempo medio de un cliente en el sistema se incrementa en 22 minutos. Si bien se ve en las secciones anteriores, que la caja 3 es la que menos se usa, hay una diferencia entre tener un sistema con 2 cajas o tener un sistema con 3 cajas.

g) Computar la probabilidad de que el tiempo medio de espera en la cola de OFT sea mayor a 5 minutos

Para computar la probabilidad de que el tiempo medio de espera en la cola OFT sea mayor a 5 minutos, almacenamos el tiempo de espera de cada cliente en la cola OFT. Luego contamos la cantidad de clientes que tienen un tiempo de espera mayor a 5 minutos, y lo dividimos por la cantidad de clientes que esperaron. Con las condiciones iniciales de la sección d , obtenemos que 40 clientes esperaron más de 5 minutos, sobre un total de 85. Lo que indica que la probabilidad de que el tiempo de espera sea mayor a 5 minutos es de 0.5294.

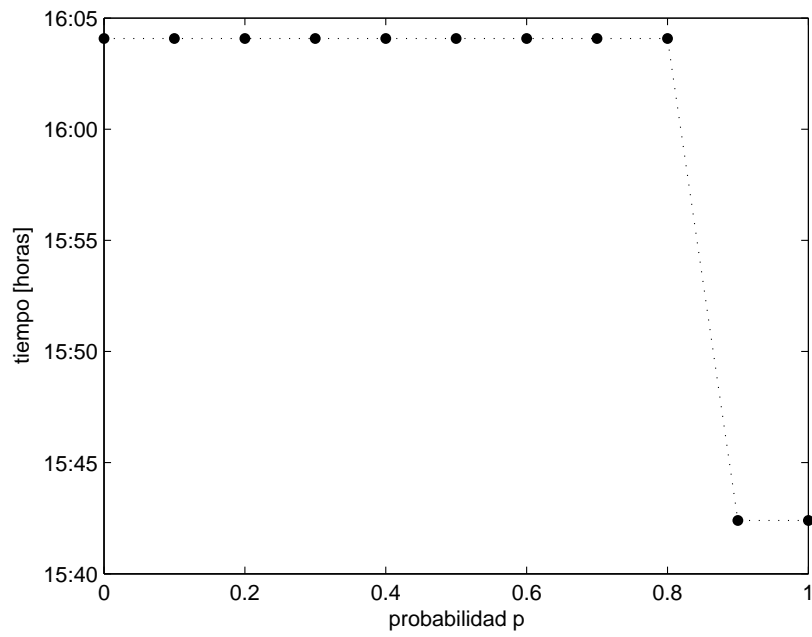


Figure 17: Tiempo medio de un cliente en el sistema, en función de la probabilidad p de que la caja C_3 esté operativa.

References

- [1] http://es.wikipedia.org/wiki/Regla_de_Sturgess
- [2] http://www.cyta.com.ar/biblioteca/bddoc/bdlibros/guia_estadistica/modulo_8.htm
- [3] Diaz, Alejandro: Notas de Clase Curso 2009, Simulación por Eventos Discretos