1. Introducción

Lavadero de Coches

Un pequeño autoservicio de lavado en el que el coche que entra no puede hacerlo hasta que el otro haya salido completamente, tiene una capacidad de aparcamiento de 10 coches, incluyendo el que está siendo lavado. La empresa ha estimado que los coches llegan siguiendo una distribución de Poisson con una media de 20 coches/hora, el tiempo de servicio sigue una distribución exponencial de 12 minutos. La empresa abre durante 10 horas al día. ¿Cuál es la media de coches perdidos cada día debido a las limitaciones de espacio?

El problema original al que nos enfrentamos se basa en dar respuesta a la pregunta: ¿cuántos clientes se pierden en promedio cada día en el lavadero de coches dadas las condiciones planteadas? Algo que podemos solucionar de manera directa utilizando tanto modelos matemáticos como de simulación; pero que también da paso a otras interrogantes que enriquecen al grueso de nuestro proyecto como:

1. ¿Cómo se comporta el sistema si los clientes se impacientan y abandonan la cola?
2. ¿Será más eficiente el negocio si se tiene una cola con dos servidores, que dos colas por cada servidor?(Una pregunta que intuitivamente parece indicar que sí) ¿Qué sucederá si las colas se mantienen balanceadas?
3. ¿Existe alguna relación entre las variables que se encuentran bajo análisis? ¿Y si es así, de qué tipo?
4. ¿Qué comportamiento presentan dichas variables? ¿Siguen alguna distribución en particular?

Las respuestas a estás cuestiones constituyen los objetivos y metas de nuestro trabajo.

1.1. Variables que describen el problema

Posee como variables:

1. la capacidad de el autoservicio.
2. la distribución con la que llegan los coches a el autoservicio.
3. la distribución con la que son atendidos los coches.
4. el tiempo que abre la empresa durante un día.
5. la forma en que son atendidos los coches una vez en el autoservicio.
6. media de coches perdidos cada día debido a las limitaciones de espacio.

1.2. Sistema específico

El sistema tratado es un lavadero de autos, representado como un conjunto de máquinas de lavado (servidores) y un conjunto de coches.

Para modelarlo abstrajimos el sistema original como una lista de enteros (donde cada coche está representado por uno de estos) y variables booleanas (indican si alguno de los servidores se encuentra en funcionamiento o no). Encontramos este modelo suficiente, dado que contiene la información necesaria para dar respuesta a las interrogantes que enfrentamos, manteniendo solo los detalles indispensables.

Los estados del modelo están caracterizados por:

1. Cantidad de coches en la cola.
2. Cantidad de coches siendo atendidos.

Las entidades del modelo son:

1. Coches
2. Servidores (máquinas de lavado)

Cuyos atributos son:

1. identificadores para los coches.
2. identificadores para los servidores.
3. estado de funcionamiento de los servidores.

Los eventos principales son:

1. Arribo de un nuevo coche
2. Finalización de lavado de un coche
3. Partida de un coche por impaciencia

1.3. Variables de interés

Para dar soluciones a las objetivos del proyecto tomamos como variables de interés:

1. Tiempo de uso de los servidores
2. Cantidad de coches perdidos
3. Tiempo promedio de espera
4. Tiempo inter-arribos promedio

2. Modelo Matemático

2.1. Markov Chain

Primero hay que establecer una unidad de tiempo, por ejemplo, usar una Poisson con un rate de 20 llegadas por hora, sería equivalente a una cantidad $\lambda = 20/60/60$ llegadas por segundo.

El problema puede ser modelado como una cadena de Markov donde los estados representan la cantidad de clientes en el lavado en un momento dado, para poder usar este modelo, es necesario que las probabilidades de ir de un estado i a un estado j solamente dependa de el estado i. Además es necesario garantizar que en un momento dado de tiempo ocurre uno de varios sucesos con probabilidades tales que sumen 1, estos sucesos en el caso de el problema serían:

1. llega un cliente nuevo.
2. es atendido un cliente.
3. no sucede ninguno de los anteriores.

En particular es necesario garantizar que en un intervalo de tiempo no ocurre ninguna de las siguientes:

1. que lleguen dos o más clientes.
2. completen su servicio dos clientes.
3. o que llegue un cliente y es atendido otro.

Escogiendo como unidad de tiempo un segundo obtengo que:

1. $\lambda = 20/60/60$, rate con que llegan los clientes.
2. $\mu = 1/12/60$, rate con la que es atendido un cliente.

Con este λ se obtiene que si $X \sim Pois(\lambda)$:

- $P(X = 0) = 0{,}9944598$
- $P(X = 1) = 0{,}0055247$
- $P(X = 2) = 1{,}53466e - 05$ prácticamente despreciable, por lo que podemos asumir (1).

(2) no sucede porque el lavado atiende los clientes de uno en uno.

La probabilidad de que llegue un cliente en una unidad de tiempo se puede ver como la probabilidad de que una exponencial con un rate λ reporte un tiempo de llegada ≤ 1, o sea si : $Y \sim Exp(\lambda)$ entonces: $P(Y <= 1) = P(X = 1)$

La probabilidad de que suceda (3) es el producto de las probabilidades de ambos sucesos porque son independientes, que sería: $P(X = 1) * P(Y = 1) = 0{,}0055247 * 0{,}001386961$ (lo consideramos como despreciable por lo que se puede asumir (3))

Dado que puedo asumir (1), (2), (3), puedo asumir el modelo de Markov, tendría 11 estados : $0, 1, 2, ..., 10$.

Las transiciones serían (separando los estados $0, 10$):

$$p_{ji} = \begin{cases} \mu & \text{si } i = j - 1 \\ \lambda & \text{si } i = j + 1 \\ 1 - \mu - \lambda & \text{si } i = j \\ 0 & \text{en otro caso} \end{cases}$$

Esta cadena de Markov posee una sola clase recurrente que no es periódica, por lo que se puede afirmar (teorema) que a largo plazo las probabilidades de que la cadena esté en un estado específico convergerán a una distribución π, donde π_j representa la probabilidad de que la cadena se encuentre en el estado j.

Asumamos por ahora, que ya calculamos π, entonces la probabilidad de perder a un cliente debido a que el lavado está lleno, sería la probabilidad de que llegue un cliente cuando la cadena está en el estado 10 (llena), que sería:

$$\pi_{10} * \lambda$$

Esta relación es la que usaremos en la simulación para contar los clientes perdidos.

Para hallar la distribución π se puede usar la siguiente relación:

Notar que en la cadena debe haber un cierto balance entre las transiciones que van de el estado i al estado $i + 1$, y las que van de el estado $i + 1$ a el i, porque exceptuando la primera vez, cada vez que se va de el estado i al $i + 1$, se debe haber ido previamente de el $i + 1$ a el i.

Por lo que $\pi_i * p_{i,i+1} = \pi_{i+1} * p_{i+1,i}$, esto permite hallar π, notar que:

1. $\frac{p_{i,i+1}}{p_{i+1,i}}$ es constante $= \frac{\lambda}{\mu} = \rho$
2. $\sum_{i=0}^{10} \pi_i = 1$. porque π es una distribución.

Queda que:

$$\pi_0 = \frac{(1 - \rho)}{1 - \rho^{11}}$$

$$\pi_n = \frac{(1 - \rho)}{1 - \rho^{11}} * \rho^n$$

3. Detalles de Implementación

3.1. Simulación

Para la implementación del proyecto nos basamos en el Event Scheduling/Time Advance Algorithm (presente en el libro Discrete Event Simulation de Jerry Banks), el cual lo llevamos al contexto de nuestro problema manteniendo una lista de eventos futuros (cuyos posibles tipos fueron comentados previamente) y avanzando el reloj (clock) de evento más inminente al siguiente, mientras las simulaciones son ejecutadas y los futuros eventos son generados.

Como pseudocódigo:

```
while clock < duration:
  pop imminent event
  advance clock

  if arrival event:
      generate arrival
      generate impatience
      simulate arrival

  if departure event:
      simulate departure

  if server available:
      generate departure
      simulate car serving
```

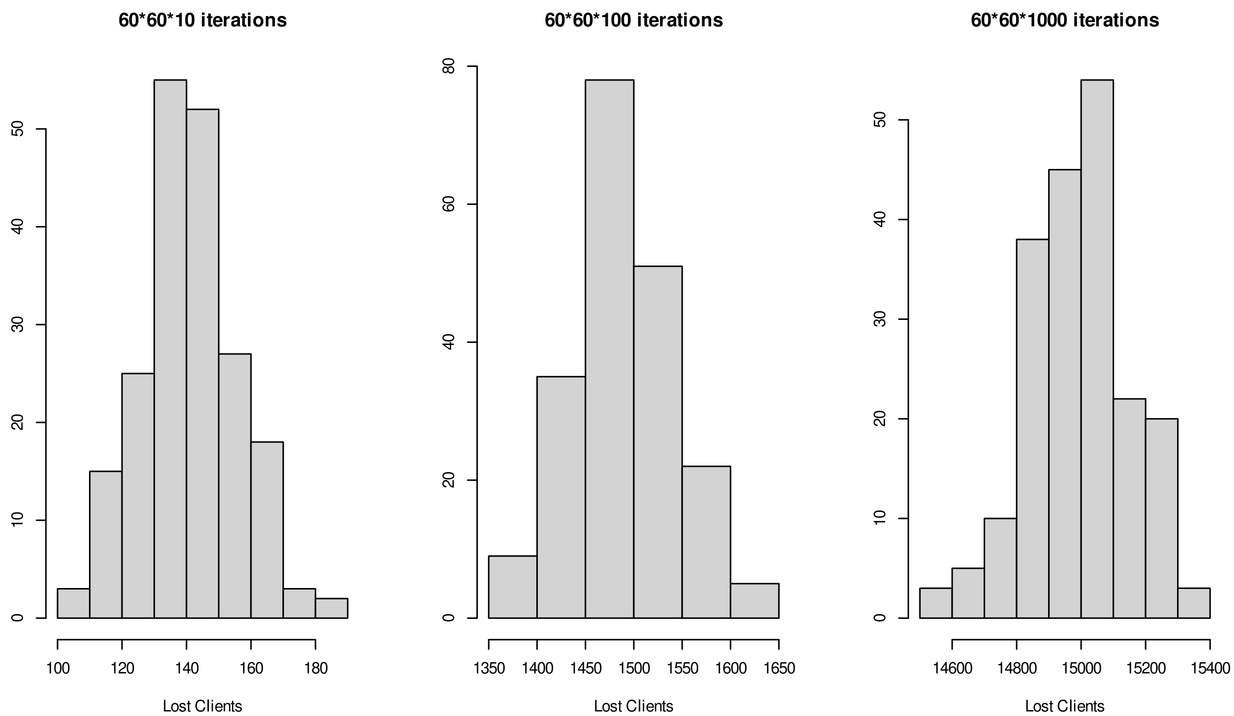
3.2. Markov Chain

En R es posible definir una cadena de Markov a partir de la matriz de transición, una vez realizado esto es posible realizar la simulación tantas iteraciones como se desee, en nuestro caso sería un total de $60 * 60 * 10$ que serían 10 horas en segundos, cada vez que se encuentre en el estado 10 (cola llena) podemos ver la probabilidad de que haya una llegada usando una variable aleatoria uniforme (como tirar una moneda biassed) para decidir si hay una llegada o no con probabilidad p.

4. Resultados y Experimentos

4.1. Markov Chain

A través de la simulación obtenemos que la media de clientes perdidos es alrededor de 140, por otro lado, con los valores de los parámetros del problema se obtiene que se deben perder alrededor de 150 clientes, obtenido a través de hallar la media de clientes que deben llegar a el lavado siguiendo la distribución de Poisson, que sería: $\mathbb{E}(x) = \lambda * 60 * 60 * 10$ multiplicado por la probabilidad de que cuando lleguen la cola esté llena π_{10}. Asumimos que la razón por la que hay esta diferencia de alrededor de 10 clientes, es porque la cantidad de iteraciones de la simulación no es suficiente, ya que la distribución π es cuando $n \to \infty$. Para confirmar lo anterior multiplicamos por 10 la cantidad de iteraciones dos veces, visualizamos los resultados en los siguientes histogramas:



Como conclusión vemos que la cantidad de clientes perdidos aumentará proporcionalmente con respecto a el tiempo en que el lavado se encuentre abierto. La función :

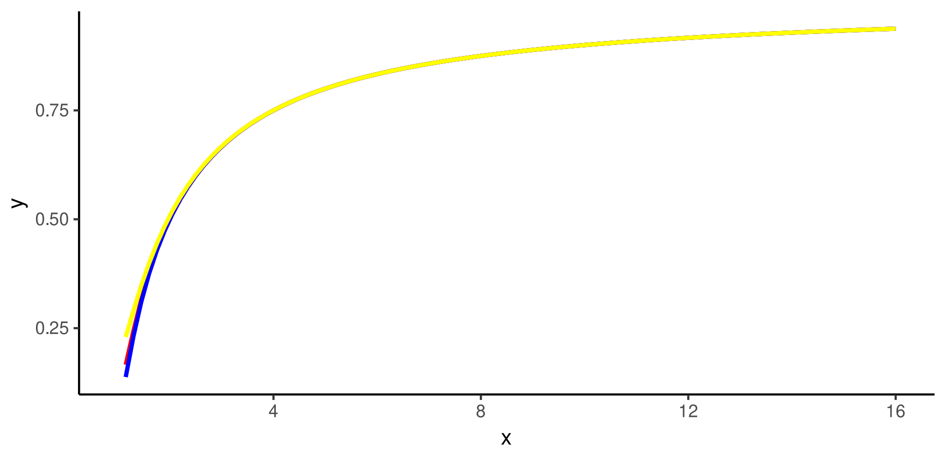
$$f(x) = \frac{(1-x)x^{10}}{1-x^{11}}$$

Es monótona creciente en $[0, \infty)$, $f(0) = 0, \lim_{x\to\infty} f(x) = 1$ (parece una CDF), y no está definida en 1. De esto se deduce que si $\frac{\lambda}{\mu}$ tiende a infinito la cantidad de clientes perdidos sería la máxima posible, que tiene sentido porque si la frecuencia con la que llegan clientes es mucho mayor comparada con la que son atendidos entonces habría congestión.

Aumentar la capacidad de la cola a m, resultaría en una nueva función:

$$f(x,m) = \frac{(1-x)x^m}{1-x^{m+1}}$$

Pero resulta que no hay mucha diferencia entre las funciones para diferentes valores de m:



Esta función en las condiciones del problema da como valor $0{,}75$ lo que significa que las $\frac{3}{4}$ partes del tiempo el lavadero está lleno, provocado porque la razón entre los rates es 4.

Como conclusión lo ideal sería disminuir el tiempo de atención a los clientes.

4.2. Problema original (Single queue, single server)

Al simular el problema original obtuvimos como respuesta que al día se perdían 141 coches en promedio; con un tiempo de espera medio de 8 minutos, una utilización del $99{,}5\,\%$ del servidor y un tiempo inter-arribos efectivo (tiempo en que los coches arriban a la cola y pueden pasar a esta) promedio de 10 minutos.

Estos buenos resultados del sistema (dada la utilización del servidor), nos llevaron al planteamiento de las siguientes preguntas o hipótesis:

Hipótesis:

1. Si los clientes se impacientan y se marchan, el comportamiento del sistema empeorará.

2. El sistema con otro servidor y una cola, tendrá mejor comportamiento que con otro servidor y dos colas; sin embargo no mejor que con dos colas balanceadas.

Clientes impacientes:

Una vez realizada la simulación para el sistema con clientes impacientes obtuvimos un resultado que refuta nuestra hipótesis; y es que el sistema tiene un mejor comportamiento que con clientes que no abandonan la cola:

	service_usage	lost_cars	mean_delay_time	effective_interarrival_mean
count	1000.000	1000.000	1000.000	1000.000
mean	35849.054	141.180	5143.117	610.675
std	223.141	15.606	672.968	73.610
min	34596.000	90.000	3567.412	446.167
25%	35777.000	131.000	4648.258	556.308
50%	35946.500	141.000	5082.186	605.593
75%	36000.000	152.000	5535.541	652.585
max	36000.000	194.000	8001.077	941.789

Figura 1: **Sistema Original:**

	service_usage	lost_cars	mean_delay_time	effective_interarrival_mean
count	1000.000	1000.000	1000.000	1000.000
mean	35814.129	142.971	1571.670	232.853
std	239.227	15.478	37.449	8.860
min	34518.000	94.000	1410.667	209.302
25%	35738.750	133.000	1550.506	226.941
50%	35901.000	143.000	1574.817	232.042
75%	35995.250	153.000	1597.219	238.201
max	36000.000	196.000	1667.873	275.713

Figura 2: **Sistema Impaciente:**

Como podemos observar, la cantidad de clientes perdidos al día y el tiempo de utilización del servidor en promedio son aproximadamente los mismos, mientras que para el sistema impaciente, se reducen en casi tres veces los tiempos de espera y los tiempos inter-arribos efectivos.

La explicación de este resultado, aparentemente contradictorio, está dada en que las pérdidas de clientes por impaciencias llevan a una reducción en las pérdidas por arribos al sistema lleno, manteniendo el total de clientes perdidos igual para ambos sistemas, y disminuyendo el tiempo inter-arribos efectivo. Además de esto, las salidas por impaciencia lleva a que los clientes que llegan a ser atendidos, pasen un tiempo menor en la cola (aquellos que pasan mucho se marchan de esta); esta última conclusión nos indica también, que si aumentáramos la media de la distribución normal truncada utilizada para modelar los tiempos de impaciencia, estos valores favorables para el sistema impaciente disminuirían.

Dos servidores, una o dos colas (¿balanceadas?)

De manera similar a la hipótesis anterior, los resultados obtenidos la refutaron; y es que las variables de interés obtuvieron valores casi idénticos para las tres simulaciones. Con lo que podemos afirmar que ninguna de los tres sistemas posee un mejor comportamiento.

	service_usage	lost_cars	mean_delay_time	effective_interarrival_mean
count	1000.000	1000.000	1000.000	1000.000
mean	35823.486	82.281	4666.470	304.994
std	235.726	17.906	550.264	26.171
min	34491.000	17.000	2625.674	242.156
25%	35717.750	70.000	4305.357	287.519
50%	35918.500	83.000	4648.019	302.914
75%	36000.000	94.250	5037.633	321.241
max	36000.000	137.000	6662.911	404.609

Figura 3: **Single queue:**

	service_usage	lost_cars	mean_delay_time	effective_interarrival_mean
count	1000.000	1000.000	1000.000	1000.000
mean	35825.602	83.401	4692.467	305.610
std	240.372	16.969	547.755	26.212
min	34630.000	31.000	2903.600	240.926
25%	35732.750	73.000	4337.625	286.968
50%	35925.000	84.000	4664.399	303.894
75%	36000.000	94.000	5024.674	320.908
max	36000.000	135.000	6389.549	395.233

Figura 4: **Two queues:**

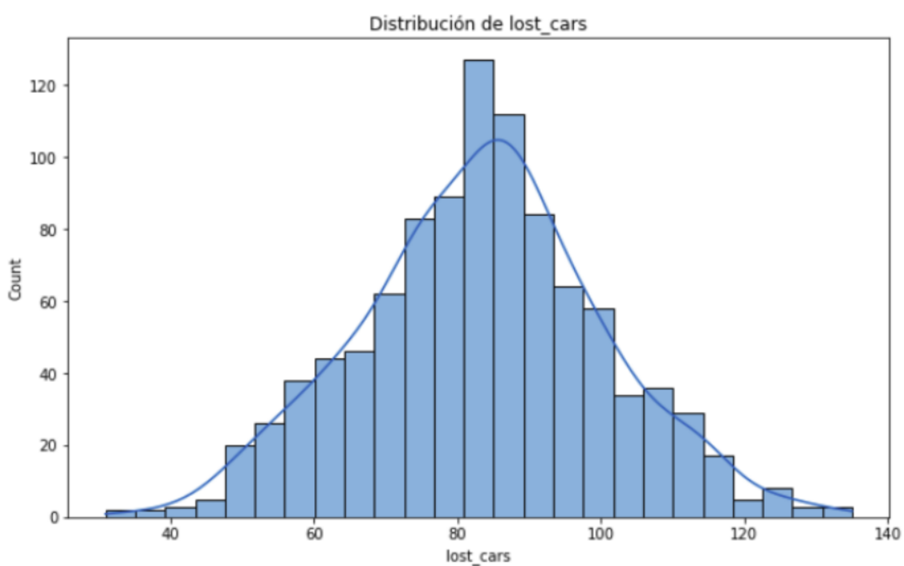
	service_usage	lost_cars	mean_delay_time	effective_interarrival_mean
count	1000.000	1000.000	1000.000	1000.000
mean	35834.834	82.200	4697.930	305.798
std	234.118	17.348	556.483	27.088
min	34263.000	20.000	2506.524	238.566
25%	35758.750	71.000	4318.056	287.188
50%	35937.000	82.000	4677.274	302.660
75%	36000.000	94.000	5072.871	321.908
max	36000.000	141.000	6770.830	394.462

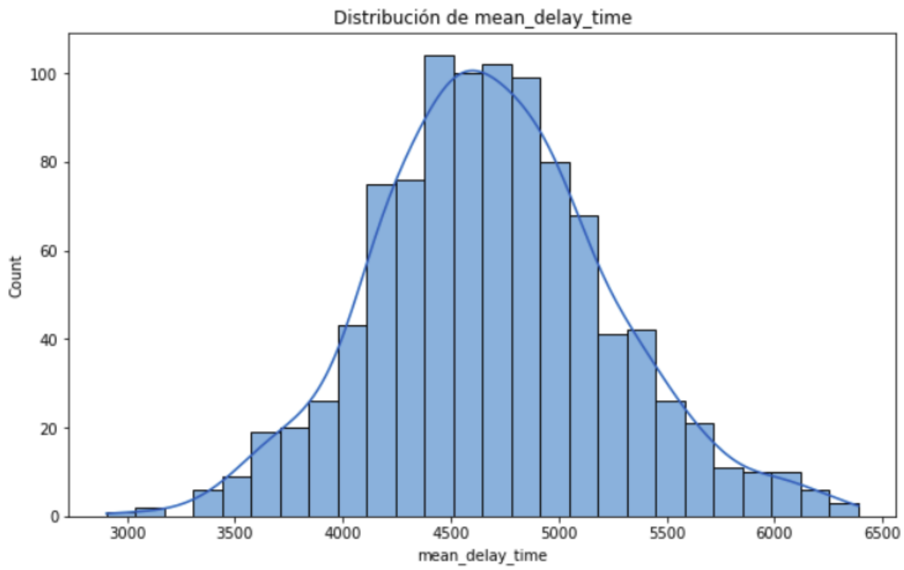
Figura 5: **Balanced queues:**

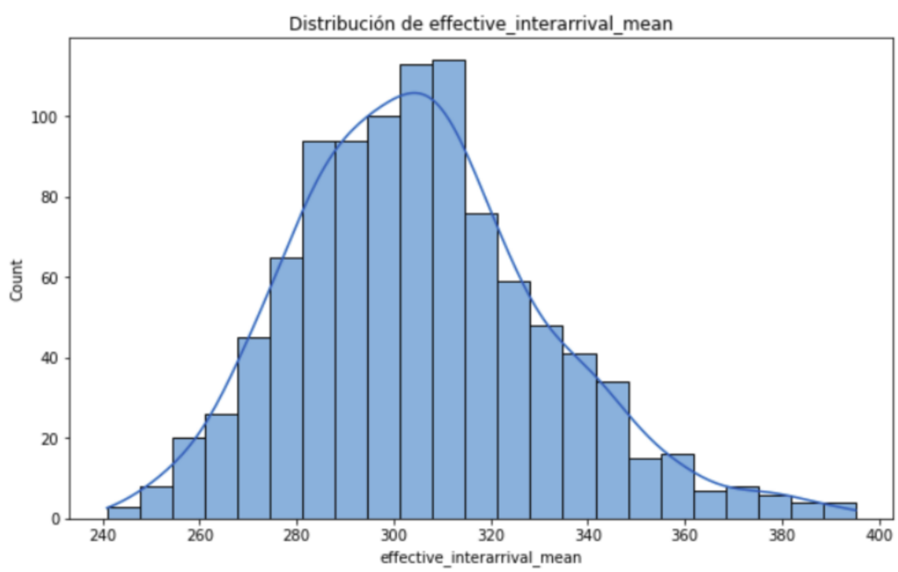
4.3. Análisis de parada

En el caso de las cadenas de Markov se realizaron las simulaciones un total de 200 veces, dado que para realizar la simulación con tiempos de duración más grandes la complejidad computacional es suficiente como para que no sea posible realizar la simulación. La distribución de los clientes perdidos se puede observar en los histogramas anteriores mostrados.

En el otro caso decidimos realizar las simulaciones un total de 1000 veces, esta cantidad es computacionalmente posible y además como se puede ver en los siguientes gráficos las variables en cuestión tienden a seguir una distribución normal, (aunque en los test realizados no obtenemos eso).




Distribución de mean_delay_time
Count
mean_delay_time
0
20
40
60
80
100
3000
3500
4000
4500
5000
5500
6000
6500


Distribución de effective_interarrival_mean
Count
effective_interarrival_mean
0
20
40
60
80
100
240
260
280
300
320
340
360
380
400