Práctica 3
Modelos de Simulación Dinámicos y Discretos

UNIVERSIDAD DE GRANADA

José Pineda Serrano
Álvaro Ruiz López

Simulación de Sistemas
4º A2 G. Ingeniería Informática

Índice

Además, se puede consultar el código fuente de la práctica en el siguiente enlace:

Código fuente de la práctica

1. Importancia del tipo de incremento del reloj.

En este apartado estudiaremos los efectos producidos por el uso de los diferentes tipos de incrementos de reloj. Para ello simularemos **un sistema de colas** formado por una única cola y un único servidor.

Los clientes llegan periódicamente al sistema siguiendo una distribución exponencial con media *tlleg* y el servidor realizará las peticiones de cada cliente en un tiempo que sigue una distribución exponencial con media *tserv*. Ambas medias de las distribuciones se expresarán en la misma unidad de tiempo que el incremento del reloj.

Los diferentes tipos de incrementos de reloj se dividirán entre:

- **Incremento Fijo.** El reloj avanzará en el tiempo una cantidad fija que podrá estar expresada en horas, medias horas, cuartos de hora, décimas de hora, minutos, segundos y décimas de segundo.
- **Incremento Variable.** El reloj avanzará en el tiempo la cantidad necesaria hasta llegar al siguiente suceso de la simulación.

Observaremos las ventajas y desventajas del uso de los diferentes tipos tanto en tiempos de ejecución, como en precisión de nuestros contadores estadísticos.

1.1. Efectos en los estadísticos.

La importancia de una buena elección en el tipo de incremento será vital para realizar una buena simulación. Una mala elección en el tipo de incremento nos llevará a perder precisión en nuestros valores estadísticos y, por tanto, resultar en un simulador inservible.

Para mostrar las diferencias ejecutaremos el simulador 100 veces con un número de clientes igual a 10000, que deberán ser atendidos por el servidor, con cada uno de los distintos tipos de incrementos.

En primer lugar obtenemos el número de clientes medio en cola durante cada simulación y observaremos si son coherentes los resultados.

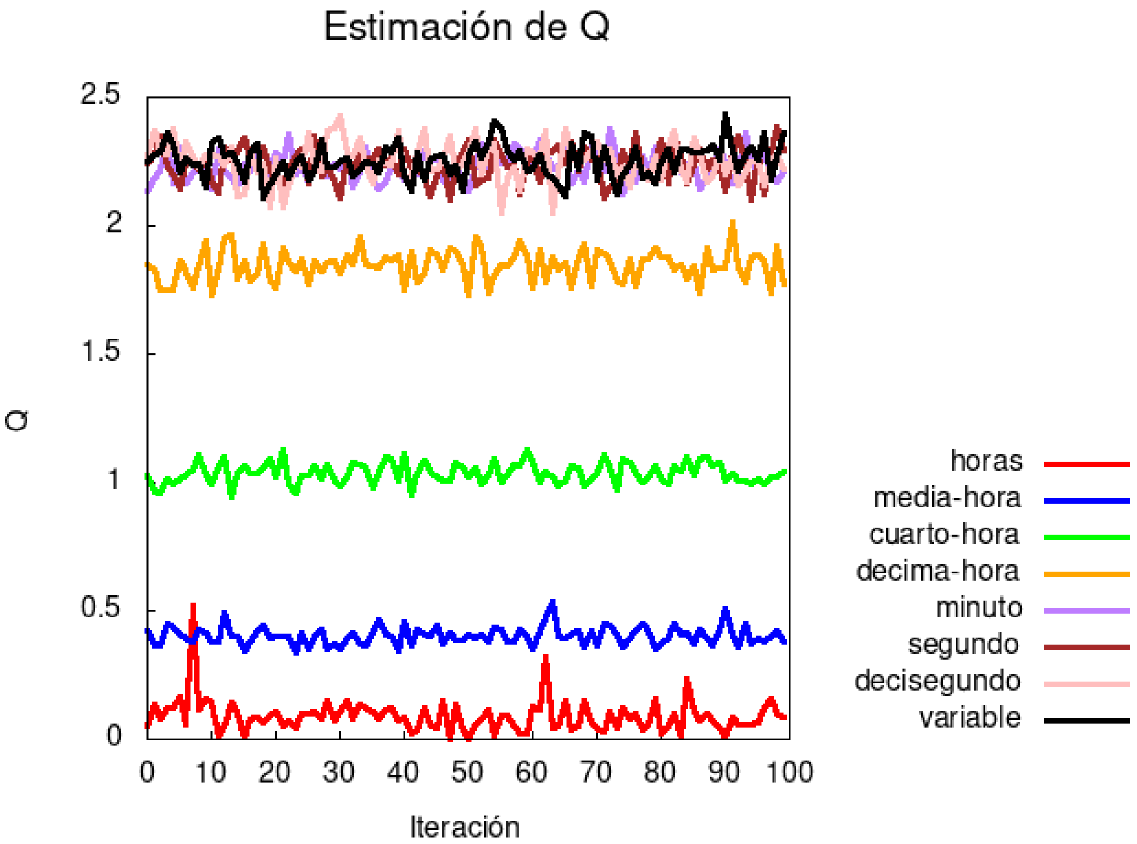

Figura 1. Número medio de clientes en cola.

Observamos que el número medio de clientes en cola aumenta progresivamente conforme se disminuye la unidad de tiempo del incremento. Esto se debe a que el sistema es incapaz de captar los cambios que se producen, es decir, los eventos suceden lo suficientemente "rápido" para que el sistema sea incapaz de determinar cuáles son los que han ocurrido antes o después.

Pongamos un ejemplo: en el caso de que un cliente llegue a las 5.05 pm y se marche a las 5.30pm, si el incremento de tiempo fuera de hora en hora, el sistema determinaría que el número de clientes en espera sería 0.

En el caso de los incrementos más pequeños tales como decisegundo, segundo y minuto, y el incremento variable obtienen unos resultados muy parecidos entre sí. Esto es debido a que los incrementos son lo suficientemente pequeños para captar los cambios ocurridos en nuestro sistema sin perder precisión. En el caso del incremento variable no tiene este tipo de problemas ya que el incremento siempre será el suficiente para alcanzar el siguiente suceso de la línea temporal.

Esta pérdida de precisión no es exclusiva del estadístico anterior sino que se ve reflejado en el porcentaje de tiempo de ocio del servidor.

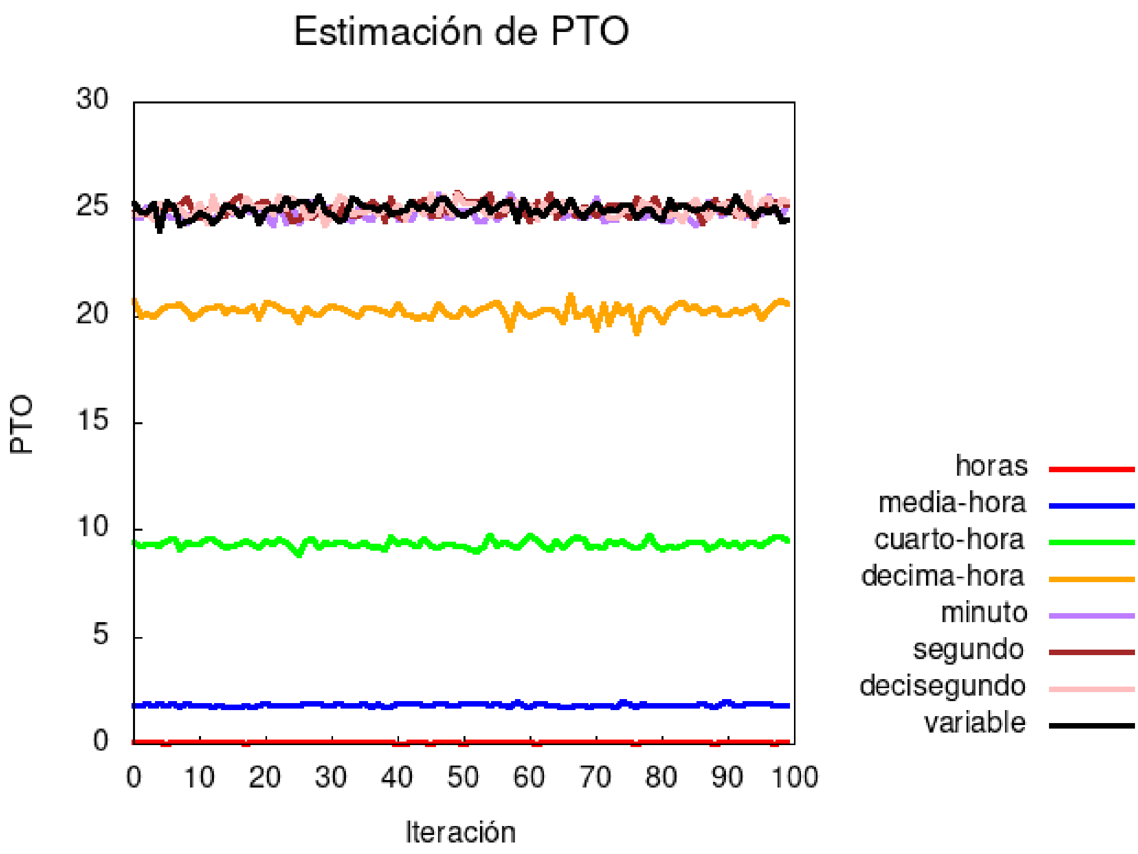

Figura 2: Porcentaje de tiempo de ocio del servidor .

Ocurre exactamente igual que en el caso anterior, los incrementos de tiempo más pequeños obtienen unos resultados muy parecidos al igual que el incremento variable, mientras que cada uno de los incrementos más grandes quedan alejados del porcentaje real. La razón es la misma que explicamos anteriormente.

Si bien hablamos de precisión constantemente es debido a que conocemos el valor teórico de ambas métricas:

$$Q(n) \rightarrow \frac{\rho^2}{1-\rho}$$

$$PTO(n) \rightarrow 100 * (1 - \rho)$$

Siendo $\rho = tserv/tlleg$ y siempre cumpliendo que $tserv < tlleg$. Con esto podemos calcular el valor de ambas, que corresponden 2,25 personas de media en cola y 25% del tiempo el servidor se encuentra en ocio.

Ahora podemos comparar cada uno de los valores de cada estadístico con su respectivo valor teórico, pudiendo obtener los tipos de incrementos más adecuados, es decir, aquellos que más se acerquen a dicho valor.

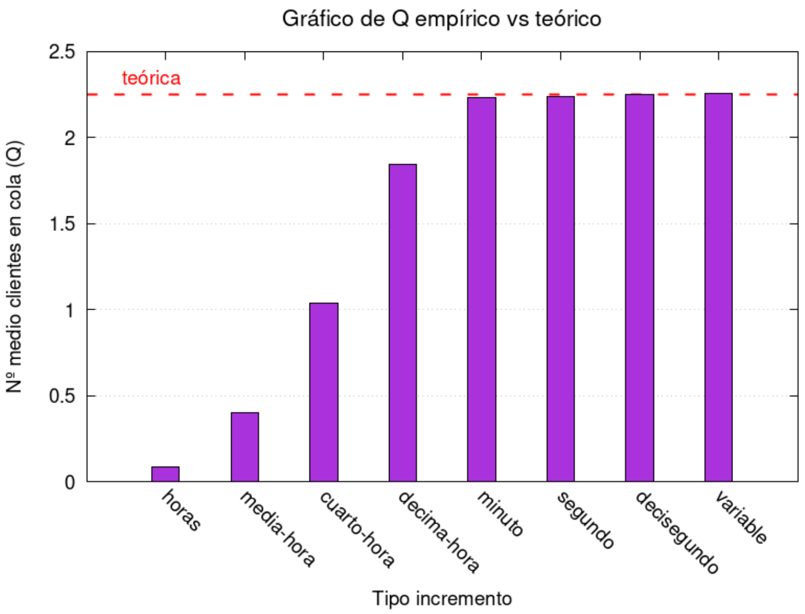

Figura 3: Comparación del estadístico número de personas medias en cola con el valor teórico .

Observamos en la figura que efectivamente los incrementos de minutos, segundos, decisegundos y variables son los más cercanos al valor teórico. Veamos si ocurre igual con el otro estadístico.

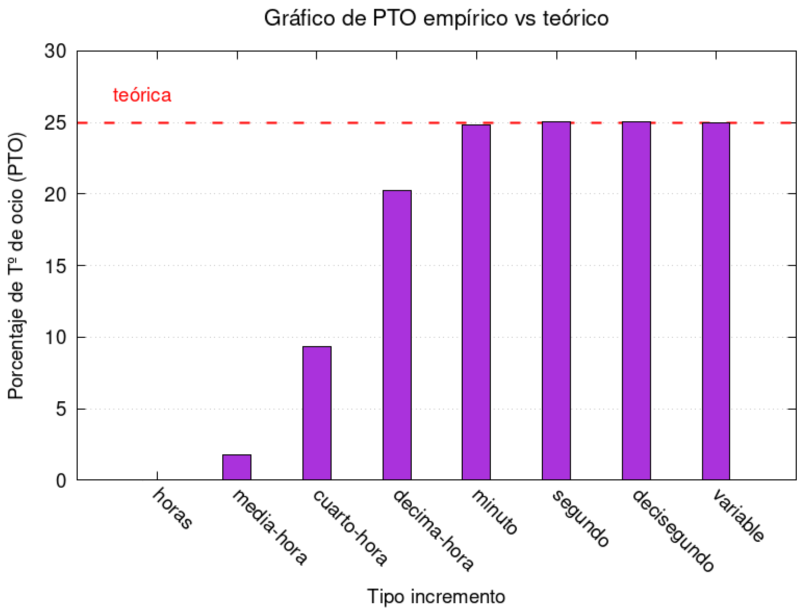

Figura 4: Comparación del porcentaje medio de tiempo de ocio con el valor teórico .

Ocurre exactamente igual con los incrementos más pequeños y el incremento variable.

Hasta ahora podríamos llegar a pensar que no existen diferencias entre los tipos de incrementos, al menos para aquellos que obtienen valores de los estadísticos muy cercanos al resultado teórico, pero esto no es del todo cierto.

Debemos seleccionar cuidadosamente el tipo de incremento que elijamos en nuestros programas ya que una mala elección puede darnos lugar a obtener conclusiones totalmente erróneas. Si bien es cierto que en este caso disponemos del valor teórico para comprobar que lo estamos haciendo correctamente, en un problema real podríamos no disponer de dicho valor.

Además de esto existen problemas más allá de las precisión de las métricas que influyen directamente en el comportamiento de nuestro programa tal y como veremos en el siguiente apartado.

1.2. Efectos en los tiempos de ejecución.

En este apartado veremos otro de los efectos relacionados con la elección del tipo de incremento. Pensemos en avanzar en el tiempo 100 años de hora en hora, deberíamos dar la friolera de 876.000 pasos, mientras que si elegimos de año en año daríamos únicamente 100 pasos. Esto es justo lo que ocurre con nuestro programa de simulación.

Los incrementos pequeños de tiempo obtendrán unos mayores tiempos de ejecución, mientras que los pasos más grandes obtendrán menores tiempos. Para poder mostrar este suceso ejecutaremos el programa con un número suficientemente grande de clientes, 100000 clientes, tomando los tiempos de cada iteración.

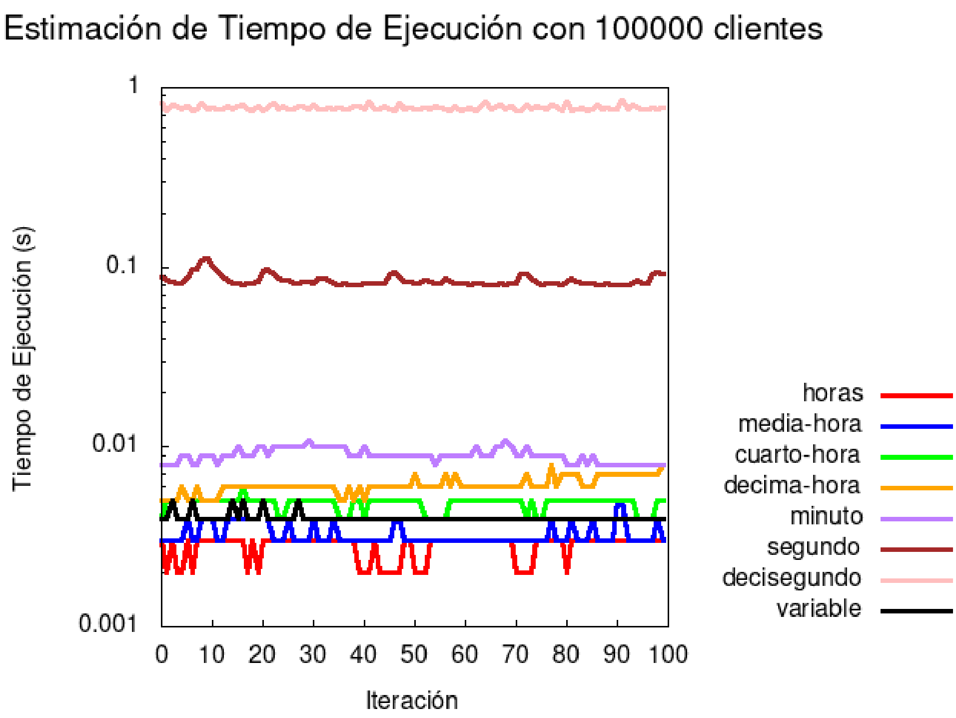

Figura 5: Tiempos de ejecución según el tipo de incremento .

Observamos que los incrementos mayores tienen los tiempos más reducidos y conforme vamos disminuyendo el tamaño de los incrementos obtenemos cada vez mayores tiempos.

Ahora nos encontramos en un dilema porque si queremos tener la precisión adecuada debemos elegir unos incrementos de tiempo lo suficientemente pequeños para poder captar adecuadamente los cambios de estado del sistema, pero esto a su vez hace que los tiempos de ejecución sean mayores. Esto hace también que debamos preocuparnos cada cuanto se producen cambios en el estado de nuestro sistema simulado para adecuar el tipo de incremento, y aún más importante estamos "desperdiciando" tiempo de simulación en franjas donde el estado no cambia en absoluto.

Aquí aparece **el incremento de tiempo variable.** Como ya hemos hablado en un inicio, el incremento de tiempo variable realiza el incremento necesario para coincidir con el evento/suceso más cercano en la línea temporal de la simulación.

Esto nos soluciona los dos problemas principales que se nos plantean:

- **Pérdida de precisión.** Al realizar incrementos de tiempo que sitúen al sistema en el siguiente suceso de la simulación, el programa es capaz de captar correctamente los cambios en el estado del sistema obteniendo así un cálculo de los estadísticos totalmente preciso.
- **Simulación de tiempo vacío.** Evitamos hacer incrementos de tiempo que nos lleven a tiempos en los que el sistema no cambia en absoluto, acelerando así la ejecución.

Por lo tanto, en este problema parece la mejor opción hacer uso del **incremento variable del tiempo**.

También cabe aclarar que la elección del tipo de incremento deberá basarse única y exclusivamente del tipo de sistema que queramos simular. En sistemas en el que los cambios de estado sean poco predecibles, como es nuestro caso, será conveniente hacer uso del incremento variable, en cambio en sistemas en los que los cambios ocurran periódicamente podría interesarnos el uso de incrementos fijos, debido a la simpleza de su implementación.

En conclusión, deberemos estudiar a conciencia las necesidades que nos plantea el sistema real a simular y hacer uso del tipo de incremento que más se ajusten a ellas.

2. Modelo simulación Discreto: Dos colas y un servidor.

En este apartado vamos a estudiar un modelo de simulación discreto en el que nos encontraremos **dos colas independientes** que son atendidas por un único servidor. Las entidades que llegan al sistema, los clientes, lo hacen con una distribución exponencial de media $tlleg_i$, servidos en un tiempo que sigue una distribución exponencial de media $tserv_i$, siendo *i* cada una de las colas del sistema.

Ambas colas tienen **una capacidad máxima** distinta, N_i. Cuando un cliente llega a la cola i y esta se encuentra llena, la llegada será **rechazada**. Una de las limitaciones del servidor de nuestro sistema es que es capaz de atender una única cola al mismo tiempo, por ello seguirá unas reglas preestablecidas sobre cómo tratar el servicio de ambas colas.

Centrándonos en los aspectos más técnicos de la simulación de dicho sistema se hará uso de un incremento de tipo variable en el reloj de la simulación.

Con respecto a la estructura que se utilizará durante la simulación para manejar los sucesos, haremos uso de una **lista de sucesos** para controlar la secuencialidad de los eventos de nuestro sistema, existen métodos para la inicialización de las variables necesarias de la simulación (contadores estadísticos, variables de control, etc), métodos de generación de informes y gestión/generación de los distintos sucesos de la simulación.

La implementación de la generación de informes está ligeramente modificada conforme a la estructura original que se propone en el guión de la práctica, sustituyendo la salida por consola, por una salida a un archivo de texto para ser tratado por las distintas macros que automatizan el tratamiento de los datos, facilitando así el desarrollo de la práctica.

El **objetivo principal** de esta práctica es evaluar el sistema con distintas configuraciones y distintos tipos reglas de decisión, siempre intentando buscar aquella regla que minimice el número de clientes perdidos. Estas reglas definirán la actuación del servidor para elegir cual de las colas va a atender en cada momento.

En primer lugar debemos estimar el número adecuado de simulaciones necesario para obtener resultados estables, es decir, que los resultados de nuestros estadísticos sean lo más invariantes posibles a partir de un número de simulaciones, actualmente desconocido.

2.1. Obtención del número de simulaciones.

Para el cálculo del número de simulaciones realizamos la ejecución de nuestro sistema de simulación con un tiempo de parada de 100 unidades de tiempo, incrementando el número de simulaciones en 10 hasta llegar a 1000 con las 4 distintas configuraciones haciendo uso de la regla de decisión 1 (cola más larga).

- **Configuración 1º.** *tlleg1=1.1, tlleg2=1.9, tserv1=0.5, tserv2=0.8, maxcola1=20, maxcola2=10.*
- **Configuración 2º.** *tlleg1=1.5, tlleg2=1.5, tserv1=0.9, tserv2=0.9, maxcola1=30, maxcola2=30.*
- **Configuración 3º.** *tlleg1=1.3, tlleg2=1.2, tserv1=0.3, tserv2=0.7, maxcola1=5, maxcola2=15.*
- **Configuración Propuesta.** *tlleg1=2.0, tlleg2=1.3, tserv1=0.7, tserv2=0.7, maxcola1=10, maxcola2=20.*

Con todas estas configuraciones intentamos abarcar alternativas lo suficientemente distintas para, a partir del comportamiento de nuestro estadístico (porcentaje medio de clientes perdidos), se suavizan los resultados.

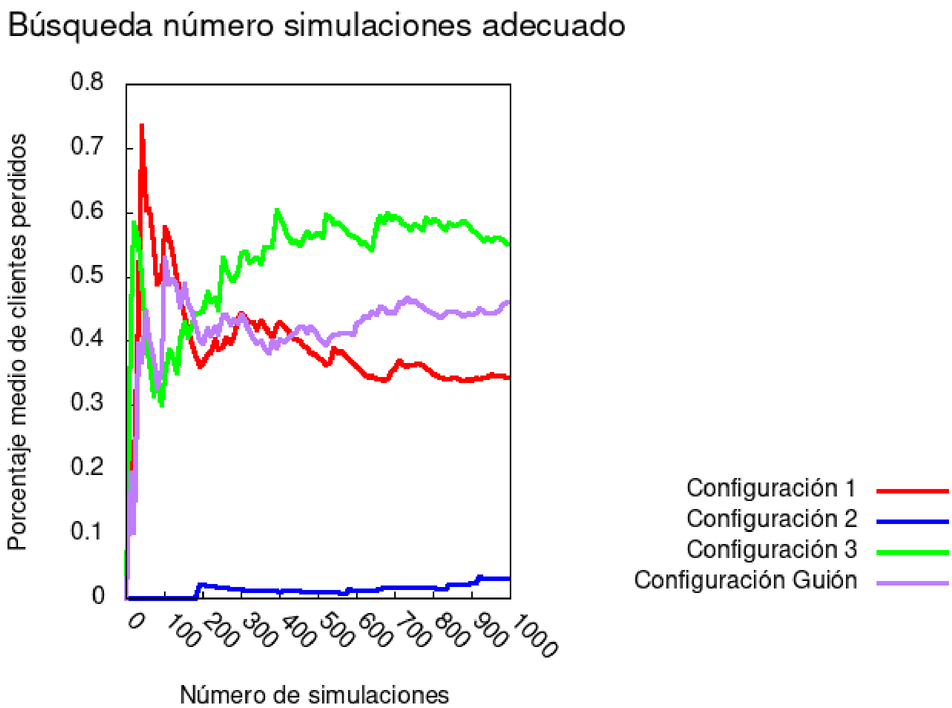

Figura 6: Porcentaje de clientes perdidos según el número de simulaciones y distintas configuraciones .

Observamos que el comportamiento comienza a estabilizarse a partir de las 250 simulaciones aproximadamente, por lo que elegimos 300 simulaciones. Este paso es importante ya que conseguimos delimitar el número de simulaciones y por lo tanto ahorrar tiempo de cómputo durante el estudio, asegurándonos que los resultados serán lo suficientemente estables para cualquier configuración.

2.2. Estudio de las reglas de decisión preestablecidas.

Una vez determinado el número adecuado de simulaciones necesarias pasamos a estudiar cuál de las reglas que se nos proporcionan es preferible para **minimizar** el porcentaje medio de clientes perdidos.

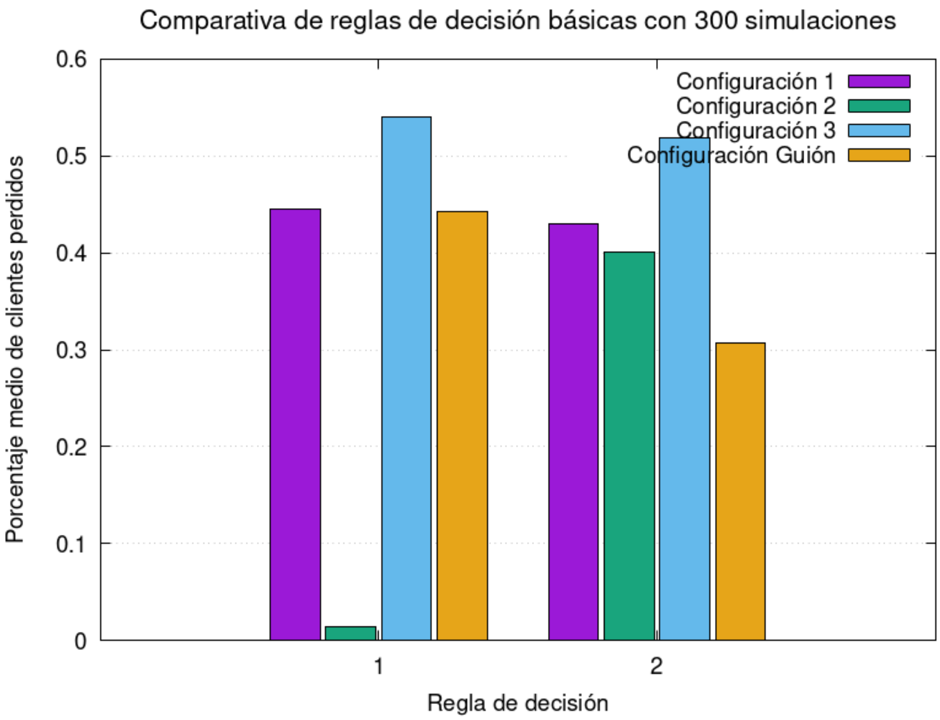

Figura 7: Porcentaje medio de clientes perdidos para las reglas preestablecidas según las diferentes configuraciones .

Observamos que en término medio la regla de decisión 2 (elección de la cola al azar) obtiene mejores resultados, salvo con la configuración 2, en la que la regla de decisión 1 destaca considerablemente.

Si comprobamos los valores de la configuración 2 detectamos porque ocurre este curioso fenómeno. La configuración 2 contiene igualdad de valores para cada uno de los tiempos de llegada/servicio y tamaño de ambas colas, es decir, que en igualdad de condiciones, es mucho más preferible que el servidor atienda la cola que más clientes tenga en espera.

2.3. Definición y estudio de nuevas reglas de decisión.

Las reglas que estudiaremos en este apartado son:

- **Regla 1:** Elección de la cola más larga.
- **Regla 2:** Elección al azar.
- **Regla 3:** Elección de la cola según el tiempo de espera mayor de los clientes que se encuentren al principio de la cola. Con esta regla, definida por los estudiantes, intentamos no depender tanto de la distribución de tiempos de servicio del servidor para cada una de las colas.
- **Regla 4:** Elección de la cola menor.
- **Regla 5:** Elección alternando entre colas.
- **Regla 6:** Elección alternando entre colas cada x_1 clientes en la cola 1 y x_2 clientes en la cola 2.
- **Regla 7:** Elección según la cola que más llena esté de acuerdo a su capacidad máxima.
- **Regla 8:** Elección de la cola dependiendo de la probabilidad de cada una, que será calculada en función de las tiempos entre llegadas de cada una.

En primer lugar mostraremos los resultados comparando las reglas básicas (1 y 2) con la definida por nosotros.

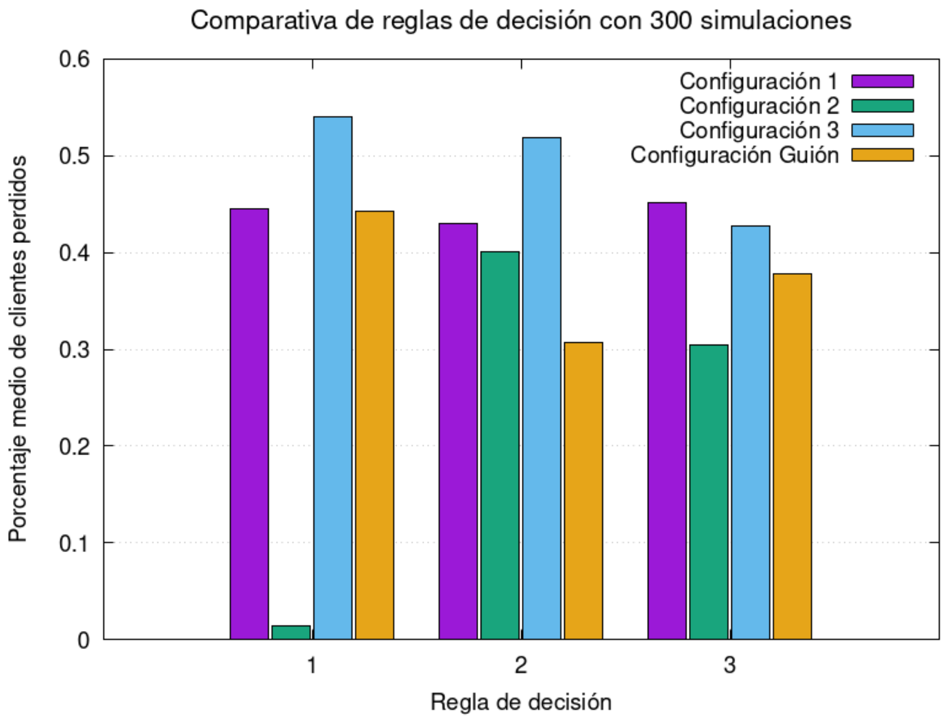

Figura 8: Porcentaje medio de clientes perdidos reglas básicas vs regla definida por alumno .

Se observa que la regla 3 destaca con la configuración 3, en la que los tiempos de servicio del servidor para cada una de las colas son lo suficiente distintos como para influir en el sistema. Como vemos destaca justo en el caso por el que se decidió implementar dicha regla.

Ahora observemos la comparativa de todas las reglas.

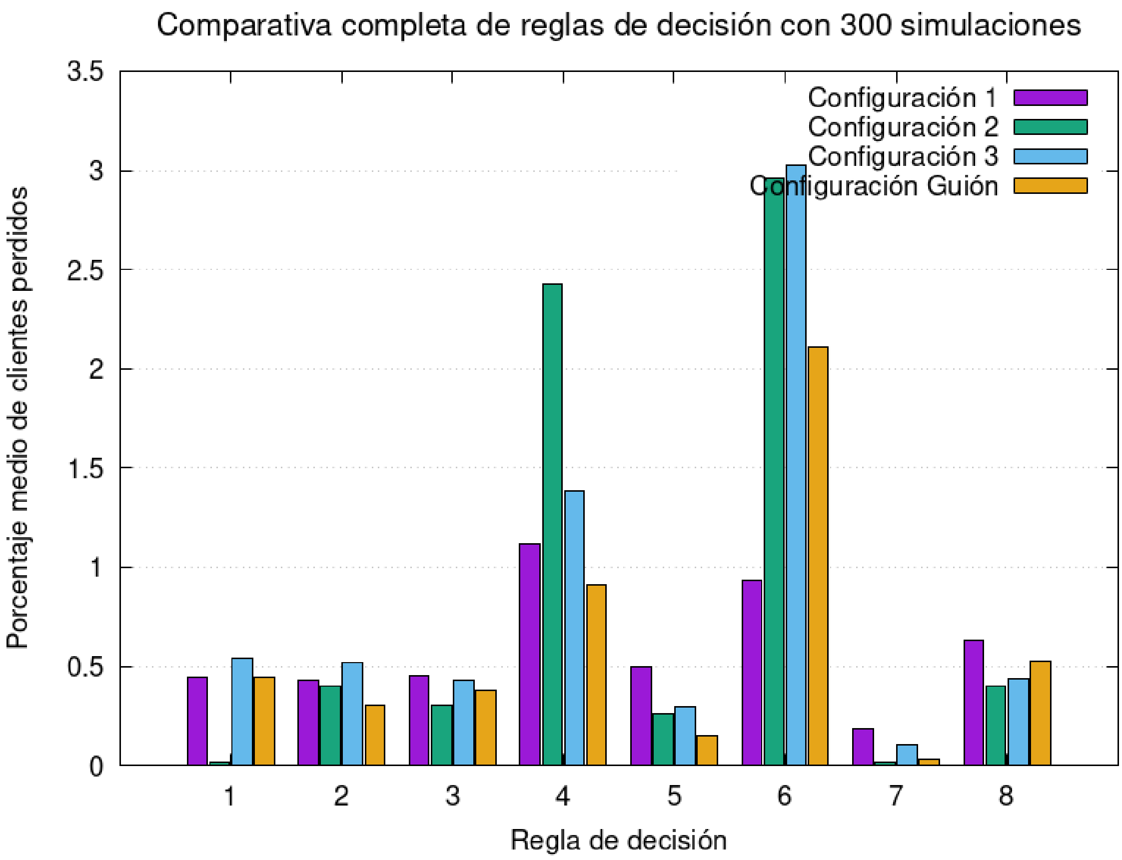

Figura 9: Porcentaje medio de clientes perdidos todas las reglas de decisión .

En esta gráfica observamos que la mejor de las reglas independientemente del tipo de configuración que utilizamos es la regla 7, elección según la cola que más llena esté de acuerdo a su capacidad máxima.

Esto realmente tiene bastante sentido ya que prioriza las colas más llenas, sin tener en cuenta el número de clientes, sino que el espacio que le queda disponible, liberando así las colas con menor capacidad restante y minimizando el número de clientes perdidos.

3. Análisis de Salidas y Experimentación.

En este capítulo nos centraremos en el análisis y comparación de las diferentes reglas de decisión del sistemas. Para ello nos basaremos en la simulación del sistema con cada una de las reglas de decisión que serán objetivo en este apartado.

En primer lugar realizaremos un estudio sobre las reglas de elección de la cola mediante una decisión aleatoria (Sistema A) y alternar entre colas (Sistema B). El estudio se centrará en la comparación del estadístico **porcentaje de llegadas rechazadas**, con un número de simulación que variará en el proceso.

1. **Simulación de cada sistema una única vez**. Observaremos cómo se comporta el sistema en una única simulación, repitiéndolo 100 veces, y quedándonos con el que obtenga un valor del estadístico menor (problema de minimización).
2. **Simulación de cada sistema cinco veces.** En este caso obtendremos un suavizado de las métricas a comparación con el caso anterior. También repetiremos este proceso 100 veces.
3. **Simulación de cada sistema 50, 100, 500, 1000, 2000 simulaciones.** Observaremos cómo influyen en los estadísticos el número de simulaciones.

Este mismo proceso será replicado para los sistemas con las reglas de decisión, elección de la cola más larga (Sistema C) y, elección de la cola más llena de acuerdo con su capacidad (Sistema D).

Los parámetros utilizados con ambos sistemas serán *tlleg1*=2, *tlleg2*=1.3, *tserv1*=0.7, *tserv2*=0.7, *maxcola1*=10, *maxcola2*=20 y *tparada*=480.

En último lugar realizaremos la comparación con dos métodos diferentes:

- **Método de los intervalos t-emparejados**. Este método se basa en la reducción del problema con dos sistemas y dos muestras a un problema con una sola muestra. Este método nos limita el número de sistemas a comparar al mismo tiempo , siendo como máximo 2 sistemas.
- **Selección del mejor de entre k sistemas.** Este método tiene una limitación diferente al caso anterior, el número de sistemas a comparar debe ser como mínimo 3 sistemas.

3.1. Estudio sistemas simulando N veces.

Para facilitar la comprensión de este apartado y realmente observar la diferencias generadas por el número de simulaciones utilizadas para obtener nuestros estadísticos, mostramos una única gráfica en la que se observa dichas diferencias.

En primer lugar comparamos los sistemas A y B:

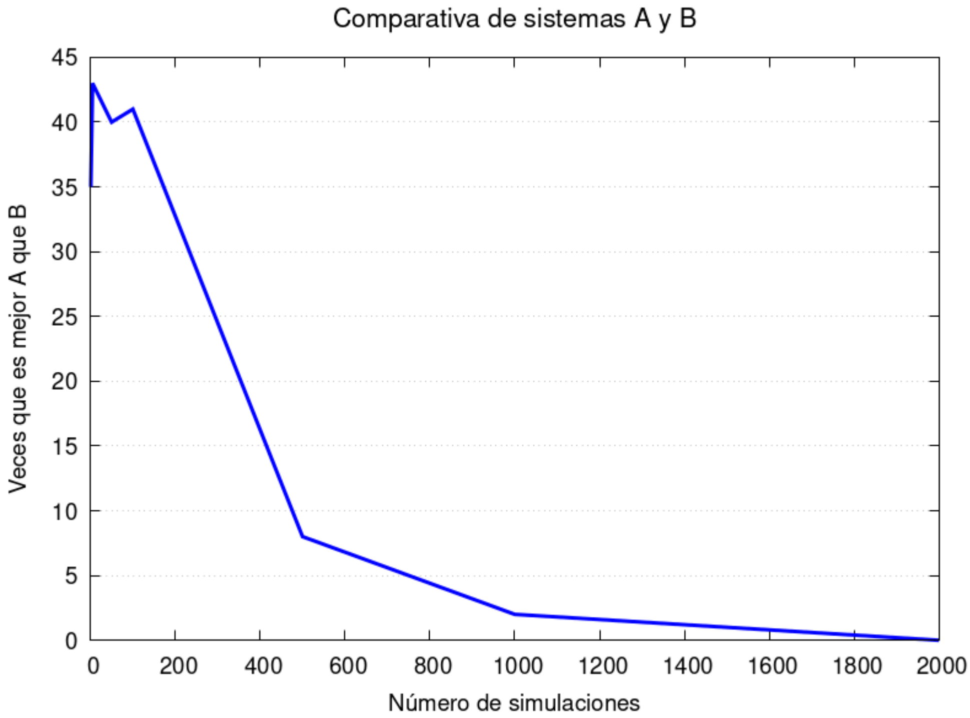

Figura 10: Número de veces que el sistema A es mejor que el B .

Observamos que en los valores más bajos del número de simulaciones no queda realmente claro cuál de los dos sistemas es el mejor de los dos, el sistema A obtiene un 35% de veces en el que es mejor que el sistema B con una única simulación. Conforme aumentamos el número de simulaciones realizadas para obtener el estadístico este valor termina convergiendo al 100% para alguno de los dos sistemas.

La velocidad a la que converja (número de simulaciones necesarias) dependerá de qué tan bueno sea un sistema frente a otro. En este caso podemos asegurar, al realizar 2000 simulaciones para el cálculo de cada iteración, que el sistema B es mejor que el A.

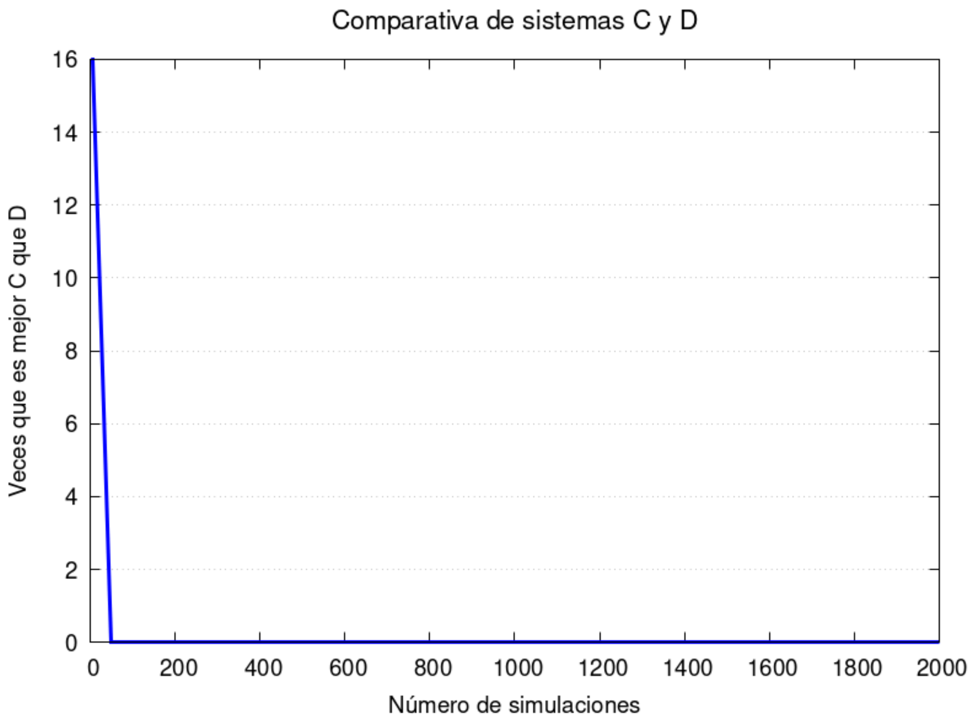

Figura 11: Número de veces que el sistema C es mejor que el D .

Si observamos el caso de la comparación del sistema C con el sistema D vemos como el número de simulaciones necesarias para que converja es mucho menor. Esto es debido a que el sistema D es mucho mejor que el sistema C.

3.1. Método de los intervalos t-emparejados.

Este método se basa en la construcción de un intervalo de confianza para la **diferencia de las dos medias.**

Como limitación de este método, además de la comentada en la introducción de este capítulo, necesita de un número de observaciones iguales para cada uno de los sistemas. Dado que para este apartado utilizamos los mismos resultados obtenidos anteriormente, tendremos un número de observaciones igual a 100 para cada uno de los sistemas.

Para el cálculo del intervalo de confianza hacemos uso de la libreria *boost-math* que contiene tanto los valores correspondientes a la distribución T-Student necesarios y un método que realiza el cálculo del intervalo. (Implementado en el archivo *calcula_intervalos.cpp*)

De nuevo compararemos el Sistema A con el Sistema B, mostrando los intervalos obtenidos con 1 y 2000 simulaciones.

Intervalos de confianza con 1 simulación.

Confianza (%)	Valor T	Ancho del Intervalo	Límite Inferior	Límite Superior
50.000	0.677	0.081	0.00852	0.16992
75.000	1.157	0.138	-0.04872	0.22716
90.000	1.660	0.198	-0.10872	0.28716
95.000	1.984	0.237	-0.14732	0.32576
99.000	2.626	0.313	-0.22388	0.40232
99.900	3.392	0.404	-0.31509	0.49353
99.990	4.055	0.483	-0.39418	0.57262
99.999	4.657	0.555	-0.46591	0.64435

Observamos que únicamente podemos tomar una conclusión con un 50% de confianza, en el que el sistema B es mejor que el A ya que el intervalo de confianza no contiene al cero y además el límite inferior es superior a 0. Esto nos llevaría a afirmar algo erróneo, que se demuestra con el intervalo calculado al 99.999% de confianza, donde sí se incluye al 0 y es el caso más extremo que consideramos, por lo que no podríamos afirmar que un sistema sea superior a otro en base a los resultados obtenidos con una única simulación.

Para excluir el 0 de los demás intervalos calculados se debe aumentar el número de simulaciones, además de reducir el propio intervalo.

Intervalos de confianza con 2000 simulaciones.

Confianza (%)	T Value	Ancho del Intervalo	Límite Inferior	Límite Superior
50.000	0.677	0.002096	0.08110	0.08530
75.000	1.157	0.003582	0.07962	0.08678
90.000	1.660	0.005140	0.07806	0.08834
95.000	1.984	0.006143	0.07706	0.08934
99.000	2.626	0.008131	0.07507	0.09133
99.900	3.392	0.010	0.07270	0.09370
99.990	4.055	0.013	0.07065	0.09575
99.999	4.657	0.014	0.06878	0.09762

Al aumentar el número de simulaciones observamos que ningún intervalo contiene al 0. En este caso podemos confirmar al 99.999% que el **sistema B es mejor que el sistema A**, ya que el intervalo de confianza para este valor es [0.06878, 0.09762], y no incluye al 0.

3.2. Selección del mejor entre k sistemas.

Para este apartado se han seguidos los siguientes pasos:

1. **Primera etapa**
   a. Simulamos cada uno de los sistemas durante 20 simulaciones obteniendo así sus varianzas y medias.
   b. Calculamos el número de iteraciones restantes para cada uno de los sistemas. El h_1 será igual a 2.583 que corresponde al valor de la tabla con $P*$ igual a 0.9 y k igual a 4. El $d*$ es igual a 0.1, lo que nos permitirá escoger el mejor sistema con probabilidad de al menos 90%, siempre y cuando la diferencia entre las medias de los sistemas sea al menos d*.

$$N_i = \max \left\{ n_0 + 1, \text{menor entero mayor o igual que } \frac{h_1^2 s_i^2 (n_0)}{d^{*2}} \right\},$$

   c. Almacenamos los valores en un archivo intermedio para la segunda etapa.
2. **Segunda etapa.**
   a. Simulamos cada uno de los sistemas con las simulaciones restantes y cargamos los resultados.
   b. Cálculo de la ponderación de cada una de las medias.

$$W_{i1} = \frac{n_0}{N_i} \left(1 + \left(1 - \frac{N_i}{n_0} \left[1 - \frac{(N_i - n_0) d^{*2}}{\left(h_1^2 s_i^2 (n_0) \right.} \right] \right)^{\frac{1}{2}} \right), W_{i2} = 1 - W_{i1}$$

   c. Cálculo de la media ponderada final.

Los resultados obtenidos de cada uno de los sistemas A, B, C y D.

ID	Media etapa 1	Var. etapa 1	Ni	Media etapa 2	W1	W2	Media Ponderada
A	0.411	0.387	259	0.612	0.092	0.908	0.593
B	0.380	0.393	263	0.555	0.091	0.909	0.539
C	1.290	1.562	1043	0.933	0.023	0.977	0.941
D	0.436	0.711	475	0.397	0.049	0.951	0.399

El modelo con mejor rendimiento será aquel que menor media muestral obtenga, que en este caso es el **sistema D**. En el caso de que hubiéramos sido más permisivos aumentando $d*$ el número de replicaciones se reduciría considerablemente.