|  |  |
| --- | --- |
|  | **Pontificia Universidad CatÓlica de Chile**  **Escuela de IngenierÍa**  **DEPARTAMENTO DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS**  **PROFESOR: PEDRO GAZMURI S.**  **ICS 3723 – SIMULACIÓN** |

.

Tarea N°3

Julio Adriazola Soto

14-05-2013

Contenido

[1 Resumen Ejecutivo 4](#_Toc356247195)

[2 Archivos involucrados 5](#_Toc356247196)

[2.1 Hoja “Datos” 5](#_Toc356247197)

[2.2 Hoja “Pregunta 1” 5](#_Toc356247198)

[2.3 Hoja “Pregunta 2” 5](#_Toc356247199)

[2.4 Hoja “Pregunta 3” 7](#_Toc356247200)

[2.5 Hoja “Pregunta 4” 7](#_Toc356247201)

[3 Análisis de largo plazo 7](#_Toc356247202)

[3.1 Estimación del periodo transiente 7](#_Toc356247203)

[3.2 Estimación media y varianza para las medidas de desempeño 9](#_Toc356247204)

[3.2.1 Disponibilidad por camión 10](#_Toc356247205)

[3.2.2 Tiempo promedio de permanencia de un camión en el taller. 11](#_Toc356247206)

[3.2.3 Tiempo promedio de permanencia de un camión en el estacionamiento 13](#_Toc356247207)

[3.2.4 Comparación medida de desempeño B con y sin transiente 14](#_Toc356247208)

[3.3 Intervalos de confianza y error relativo 15](#_Toc356247209)

[3.3.1 Intervalos de confianza al 95% 16](#_Toc356247210)

[3.3.2 Error relativo 16](#_Toc356247211)

[3.3.3 Réplicas adicionales para error relativo menor al 1% 17](#_Toc356247212)

[3.4 Batch means para la medida B 18](#_Toc356247213)

[3.5 Conclusiones 19](#_Toc356247214)

[4 Anexos 20](#_Toc356247215)

Índice de ilustraciones

[Ilustración 1: Esquema general del problema 4](#_Toc356247177)

[Ilustración 2: Medida de desempeño B para 1956 observaciones 8](#_Toc356247178)

[Ilustración 3: Medida de desempeño B para 400 observaciones 9](#_Toc356247179)

[Ilustración 4: Estabilización de la media para el valor de la medida de desempeño A 10](#_Toc356247180)

[Ilustración 5: Estabilización de la varianza para el valor de la medida de desempeño A 11](#_Toc356247181)

[Ilustración 6: Estabilización de la media para el valor de la medida de desempeño B 12](#_Toc356247182)

[Ilustración 7: Estabilización de la varianza para el valor de la medida de desempeño B 12](#_Toc356247183)

[Ilustración 8: Estabilización de la media para el valor de la medida de desempeño C 13](#_Toc356247184)

[Ilustración 9: Estabilización de la media para el valor de la medida de desempeño C 14](#_Toc356247185)

[Ilustración 10: Estabilización de la media para la medida de desempeño B con y sin transiente 15](#_Toc356247186)

[Ilustración 11: Estabilización de la varianza para la medida de desempeño B con y sin transiente 15](#_Toc356247187)

Índice de tablas

[Tabla 1: Media y Varianza para la medida de desempeño A al considerar n réplicas 11](#_Toc356247188)

[Tabla 2: Media y Varianza para la medida de desempeño B al considerar n réplicas 13](#_Toc356247189)

[Tabla 3: Media y Varianza para la medida de desempeño C al considerar n réplicas 14](#_Toc356247190)

[Tabla 4: Media, varianza e intervalo de confianza para las medidas de desempeño con 20 réplicas 16](#_Toc356247191)

[Tabla 5: Errores relativos para las medidas de desempeño con 20 réplicas 17](#_Toc356247192)

[Tabla 6: Cantidad de réplicas adicionales para conseguir error relativo menor al 1% 18](#_Toc356247193)

[Tabla 7: Promedios calculados para los distintos grupos 19](#_Toc356247194)

# Resumen Ejecutivo

En este documento, se quiere retomar el problema visto para la entrega uno. Pequeñas modificaciones se le han realizado al modelo, como la incorporación de 4 camiones adicionales, y reparaciones por tipo (Eléctricas o mecánicas), lo que implicará que existirá una tendencia a encontrar más camiones en taller, y, por otro lado, los camiones estarán más tiempo estacionados.

En base a los resultados arrojados por Arena, se buscará realizar un análisis de largo plazo para las principales medidas de desempeño, estas son:

1. Disponibilidad por camión.
2. Tiempo promedio de permanencia de un camión en el taller.
3. Tiempo promedio de permanencia de un camión en el estacionamiento.

Para ello, se ha tomado un archivo de output que cuenta con 100 réplicas independientes del modelo, con un largo de 10.000 horas cada uno. Este se puede revisar en el archivo “tiempos\_en\_taller.txt”. En adelante, nos referiremos a estas medidas según su letra.

El análisis no estará puesto en el modelo en sí, por lo que consideraremos que este representa la realidad del problema y todas las distribuciones se ajustan bien.

A continuación presentamos un esquema general del problema:

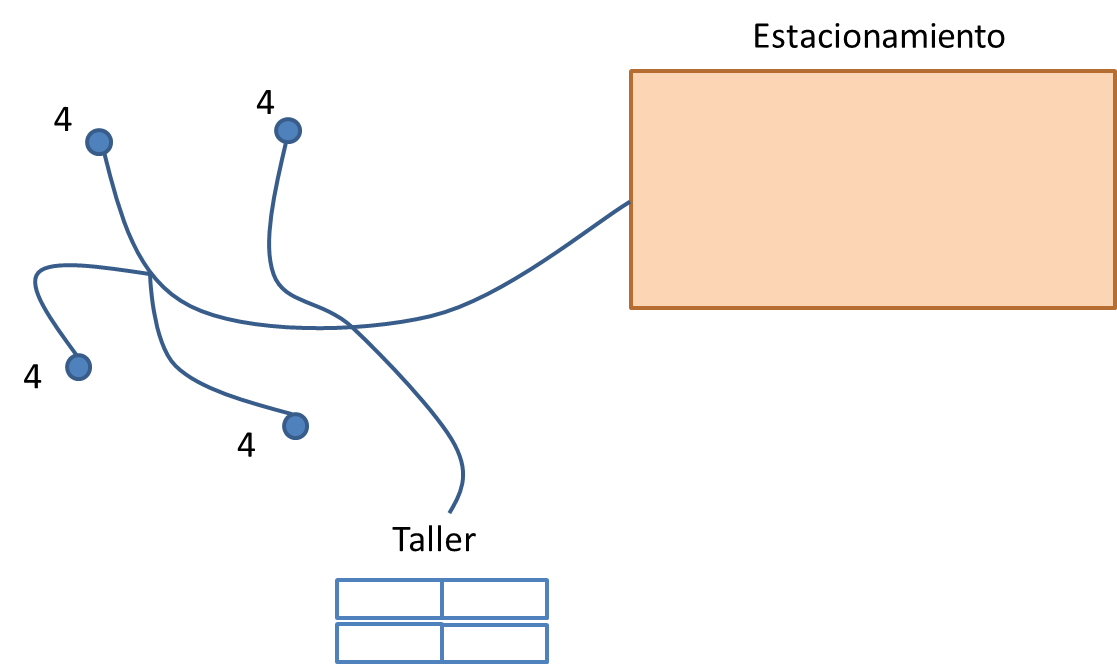
 

Ilustración 1: Esquema general del problema

# Archivos involucrados

A partir del output generado en el archivo “tiempos\_en\_taller.txt”, se ha creado una hoja de cálculo para facilitar la manipulación de los datos, la cual se encuentra en el archivo “Simulación.xlsx”. Cabe destacar que se han quitado la mayoría de las fórmulas del archivo para que, al cargar, este no demore tanto. Si se desea ahondar más sobre cómo se obtuvieron ciertos valores, se generaron comentarios en algunas cabeceras con la fórmula para alguna de las celdas a modo de ejemplo, estas últimas están en español debido a la versión de Excel usada.

## Hoja “Datos”

Esta hoja nos muestra principalmente los datos obtenidos, se han agregado algunas columnas adicionales como el número de la observación en una réplica (Columna **A:A**) y el tiempo de permanencia en el taller acumulado (Columna **C:C**), cuyo valor usaremos posteriormente para calcular la medida de desempeño A.

## Hoja “Pregunta 1”

En esta se encuentran los datos usados para calcular el periodo transiente. En las columnas **A:B** encontramos el número de observaciones por réplica para cada una de estas. En las columnas **F:G**, y en base al mínimo número de observaciones obtenido (Celda **B103**), se calculó el promedio para cada una de estas observaciones. Además se presenta un gráfico, el cual usaremos posteriormente para hacer la aproximación visual del periodo transiente. Por último, en las celdas **I1:J2** se muestra el número de “corte” para el periodo transiente y el equivalente –en promedio- en horas.

## Hoja “Pregunta 2”

Para cada una de las réplicas, se ha calculado el índice del primer y último dato con respecto a la hoja “Datos” (Columnas **C:C** y **D:D** respectivamente). En base a estos índices, se ha calculado:

* Tiempo en que se registró la última observación, en la columna **E:E**. Se consideró este tiempo y no el tiempo total de la simulación (conocido), porque esto complicaría el cálculo, ya que habría qué preguntarse qué estuvo realizando cada camión en este “delta”.
* Tiempo acumulado de permanencia en el taller sobre todos los camiones, para la última observación, en la columna **F:F**.
* Medida de desempeño A para cada réplica, en la columna **G:G**.
* Media muestral para la medida de desempeño A de las primeras n réplicas, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **H:H**.
* Varianza muestral para la medida de desempeño A de las primeras n réplicas, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **I:I**.
* Medida de desempeño B para cada réplica, en la columna **J:J**.
* Media muestral para la medida de desempeño B de las primeras n réplicas, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **K:K**.
* Varianza muestral para la medida de desempeño B de las primeras n réplicas, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **L:L**.
* Medida de desempeño C para cada réplica, en la columna **M:M**.
* Media muestral para la medida de desempeño C de las primeras n réplicas, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **N:N**.
* Varianza muestral para la medida de desempeño C de las primeras n réplicas, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **O:O**.
* Valor considerado para el largo del periodo transiente, en número de observaciones, en la celda **P1**.
* Medida de desempeño B para cada réplica sin considerar periodo transiente, en la columna **Q:Q**. Estos valores en realidad serán utilizados en la pregunta 3.
* Media muestral para la medida de desempeño B de las primeras n réplicas sin considerar periodo transiente, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **R:R**.
* Varianza muestral para la medida de desempeño B de las primeras n réplicas sin considerar periodo transiente, donde n está dado por el valor de la columna A:A correspondiente. Este valor está en la columna **S:S**.
* Tabla con los valores de las medias y varianzas muestrales para las medidas de desempeño A, B y C para 20, 30, 40, 50, 60, 70, 80, 90 y 100 réplicas. Ésta se encuentra en **U2:AD11**.
* Desde la fila 103 se muestran datos que son la copia de columnas de arriba, solo con la finalidad de realizar gráficos.

## Hoja “Pregunta 3”

Se muestran en las columnas **A:B** los valores para una Student con probabilidad 0.975, con valores de grado de libertad entre 1 y 4707. Además se muestra una tabla con los valores de mínimo y máximo del intervalo de confianza, promedio, error relativo y el “verdadero” error relativo para las 3 medidas de desempeño más la medida de desempeño B sin considerar periodo transiente, esto en las celdas **E3:L6**. Por último, se muestran los cálculos realizados para calcular cuántas réplicas adicionales se necesitan para obtener un error relativo del 1%, esto en las celdas **F9:I1912**.

## Hoja “Pregunta 4”

En las columnas **B:I** simplemente se muestra una copia de los valores obtenidos para la réplica 1, mientras que en la columna **A:A** se muestra dentro de qué grupo caerá cada observación, considerando grupos de 50 elementos. En la tabla **L1:M42** se muestran los promedios para la medida de desempeño B para los primeros 41 grupos. Por último, en las celdas **L44:O48** se muestran la media y varianza muestrales para estos valores, el mínimo y máximo para el intervalo de confianza, y el error relativo y “verdadero” error relativo del intervalo.

# Análisis de largo plazo

## Estimación del periodo transiente

Uno de los pasos más importantes al hacer un análisis de largo plazo, es estimar el periodo transiente. Al calcular las medidas de desempeño, buscaremos que estas se acerquen lo más posible a su valor en estado de régimen, y, al considerar el periodo transiente, nos alejamos de nuestra aproximación.

Primero, calcularemos el periodo transiente tomando como base el valor de la medida de desempeño B y para ello usaremos una aproximación visual. Posteriormente, compararemos algunas de las medidas, en el caso en que se considera el periodo transiente, versus el que lo elimina.

Se necesita sacar una media muestral para cada observación sobre las 100 réplicas realizadas, independiente de si ésta corresponde a algún camión en particular. Para ello se enumeraron las observaciones para cada réplica, y se obtuvo la mínima cantidad de observaciones totales. Esto arroja 1956 observaciones para la réplica número 93.

Entonces, para cada una de estas observaciones calculamos el promedio. Por ejemplo, si se desea obtener el valor promedio de la observación número j, se debe hacer lo siguiente:

Sea el valor del tiempo de permanencia en el taller de la observación j de la réplica i, entonces:

Ahora, realizamos un gráfico con estos valores para estudiar el comportamiento en el largo plazo, el cual mostramos a continuación:

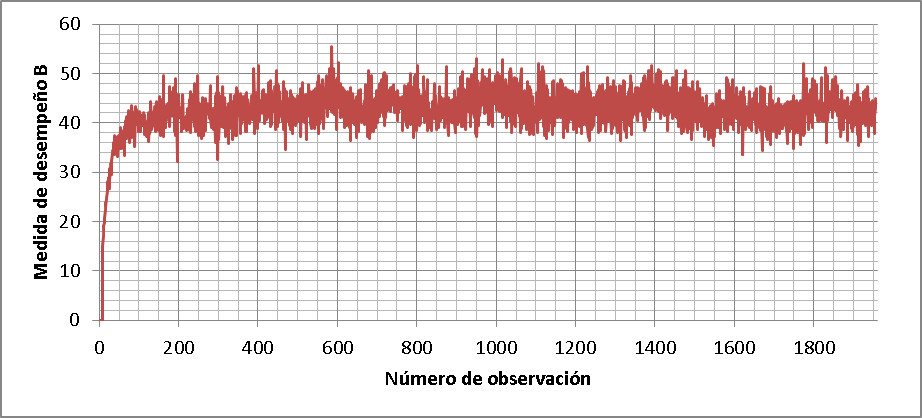


Ilustración 2: Medida de desempeño B para 1956 observaciones

Logramos apreciar que después de las 400 observaciones existe un comportamiento menos variable en torno a un valor que aproximadamente parece ser 42. En la hoja “Pregunta 1” se puede ver en extensión este gráfico, con lo cual se logra apreciar con mayor precisión cómo se va estabilizando el valor de la medida de desempeño. A continuación se muestra un acercamiento hasta las 400 primeras observaciones para poder encontrar un valor más preciso para el periodo transiente:

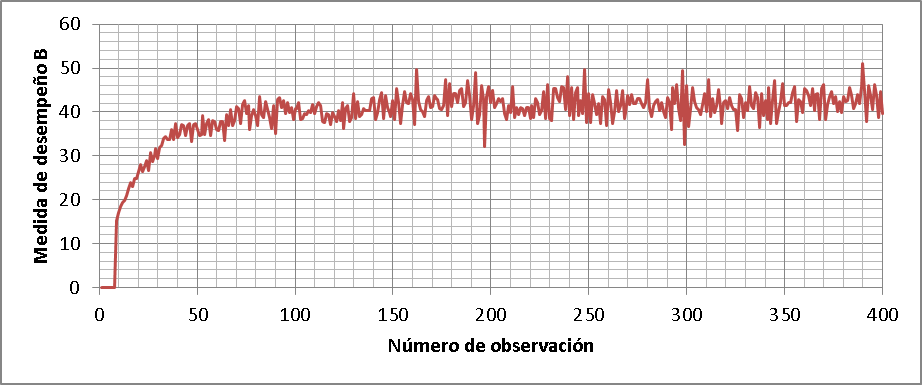


Ilustración 3: Medida de desempeño B para 400 observaciones

Pareciera que en la observación número 200 ya se consigue un comportamiento más deseable, sin embargo entre las observaciones 200 y 230 apreciamos una oscilación mucho menor en relación a los siguientes valores, por lo que podríamos inferir que aún se está en el transiente.

Para el valor 300 se empieza a ver un comportamiento más uniforme, sin embargo, solo con la finalidad de asegurar que no se está involucrando el periodo transiente, asumiremos como 320 el valor de este, es decir, desde la observación número 321 consideraremos que ya se está en periodo de régimen.

Si traducimos a tiempo estas 320 réplicas, corresponde, en promedio, a 1547.75 horas, es decir, se botará alrededor de un 15.47% de cada réplica si es que no se considera el periodo transiente.

## Estimación media y varianza para las medidas de desempeño

Para lograr sacar conclusiones, es importante conocer el valor promedio de las medidas de desempeño, así como también su varianza.

A continuación se realizará el cálculo de estos valores, pero primero explicaremos cuál es el procedimiento adquirido y/o las ecuaciones que rigen los valores que se presentarán.

### Disponibilidad por camión

Para un camión en particular, el valor de la disponibilidad de este se debiera calcular como:

Esto, debido a que el tiempo en que está operativo el camión corresponde precisamente a la diferencia entre el tiempo total y el tiempo que estuvo en el taller. Puesto que tenemos 24 camiones, es una buena aproximación calcular el valor de la medida de desempeño A para una réplica como sigue:

Donde corresponde al tiempo en que fue registrada la última observación (TNOW), y corresponde al tiempo de permanencia en el taller de la observación i en una réplica.

A partir de esto, se obtuvieron los valores para las 100 réplicas, los cuales se muestran con más detalle en la hoja de cálculo. Además, para la réplica n se calculó la media y varianza muestral considerando los valores de la medida de desempeño de las n primeras réplicas. Con esto generamos los siguientes gráficos:

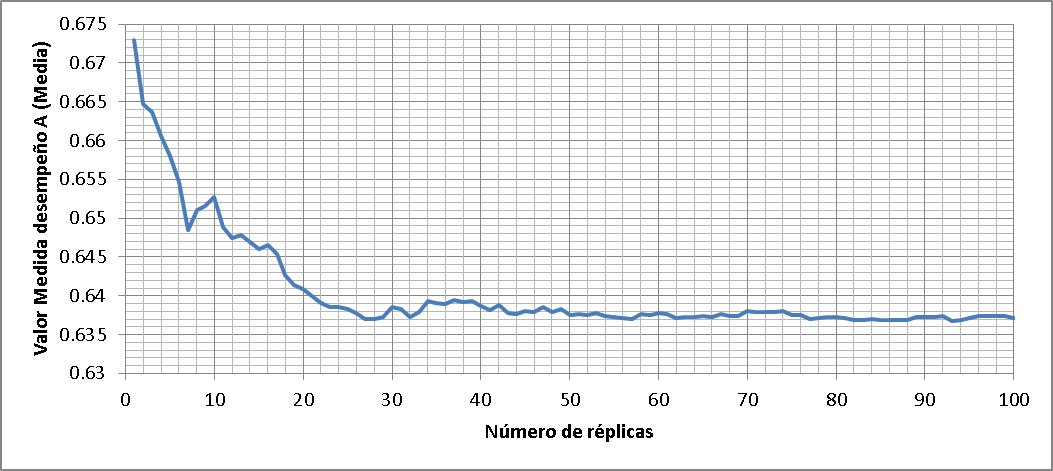


Ilustración 4: Estabilización de la media para el valor de la medida de desempeño A

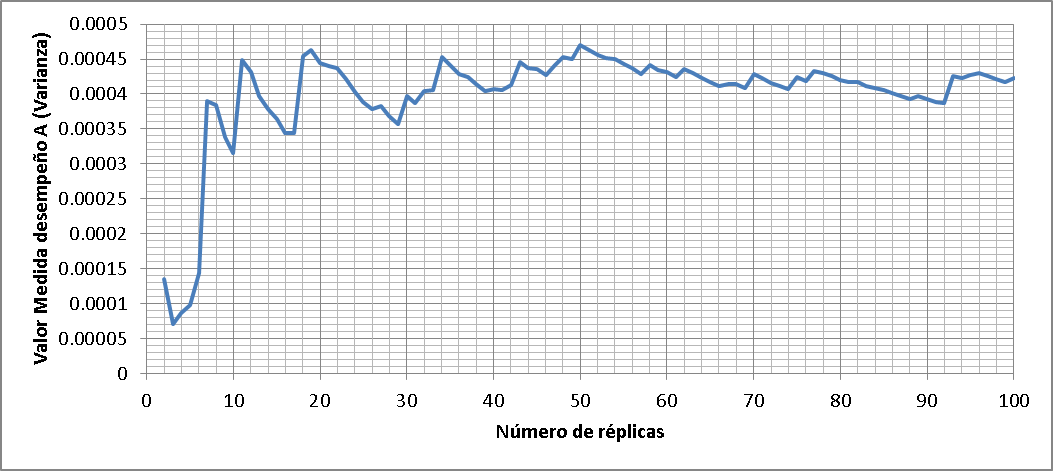


Ilustración 5: Estabilización de la varianza para el valor de la medida de desempeño A

Se aprecia que a medida que se van considerando cada vez más réplicas, los valores de la media y varianza muestral para la medida de desempeño A comienzan a estabilizarse.

A continuación se presentan los valores al considerar entre 20 y 100 réplicas, con saltos de 10:



Tabla 1: Media y Varianza para la medida de desempeño A al considerar n réplicas

Los valores a los cuales se estabilizan son a una media de 0.6371 y a una varianza de 0.00042.

### Tiempo promedio de permanencia de un camión en el taller.

El cálculo de esta medida se vuelve mucho más simple, dado que implica simplemente sacar el promedio sobre todas las observaciones de una réplica para el valor de el tiempo de permanencia en el taller.

El valor para las 100 réplicas se puede ver con mayor detalle en la hoja de cálculo, donde también se encuentran las medias y varianzas para las n primeras réplicas, con lo que generamos los siguientes gráficos:

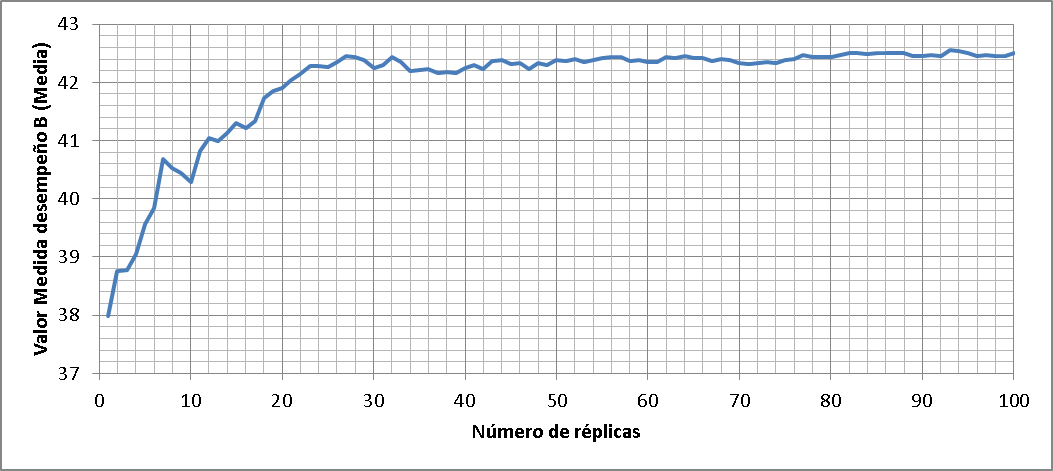


Ilustración 6: Estabilización de la media para el valor de la medida de desempeño B

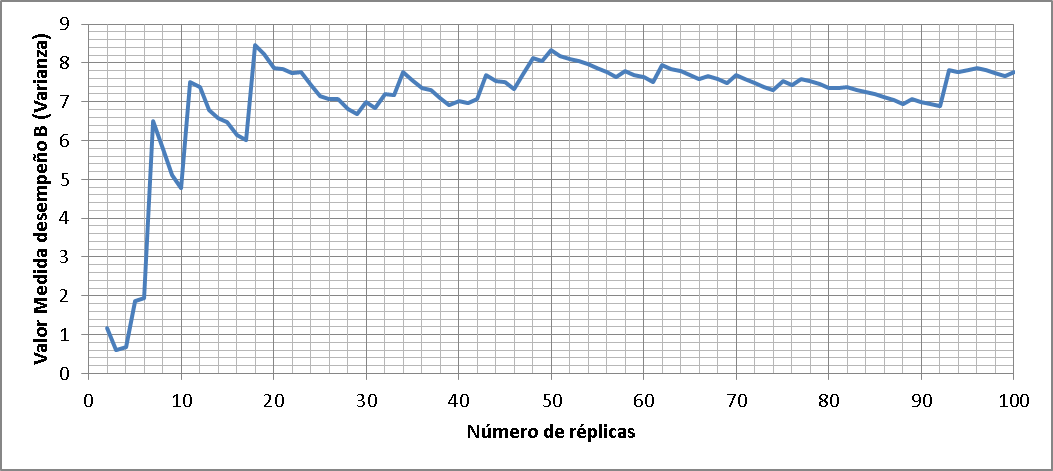


Ilustración 7: Estabilización de la varianza para el valor de la medida de desempeño B

Nuevamente notamos que ambas medidas se estabilizan mientras mayor sea el número de réplicas. A continuación presentamos los valores cuando consideramos n réplicas:



Tabla 2: Media y Varianza para la medida de desempeño B al considerar n réplicas

Por lo tanto, los valores a los cuales tienden son, aproximadamente 42.50 y 7.77 para la media y varianza respectivamente.

### Tiempo promedio de permanencia de un camión en el estacionamiento

Al igual que para la medida de desempeño B, el cálculo se realiza sacando el promedio sobre todas las observaciones de una réplica para el valor del tiempo en el estacionamiento. En la hoja de cálculo se muestran los 100 valores, y también las medias y varianzas respectivas. Al graficar los valores, nos queda lo siguiente:

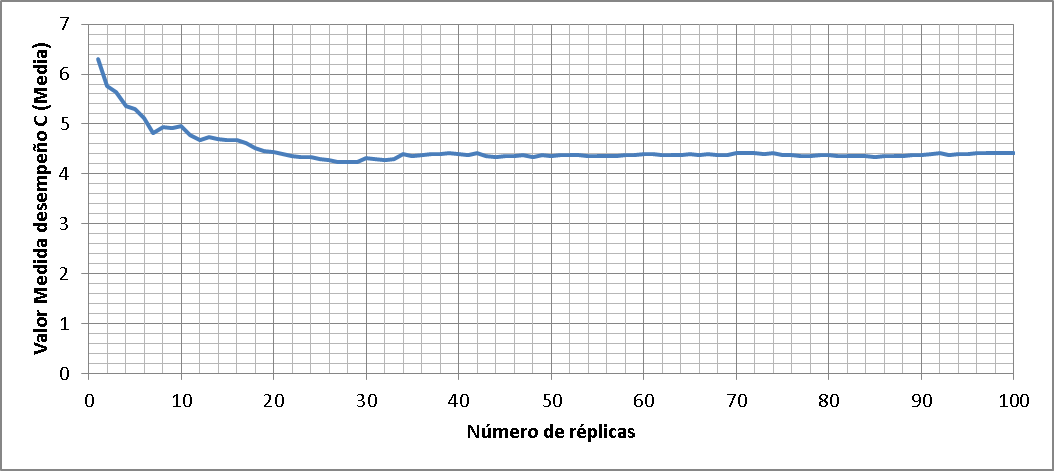


Ilustración 8: Estabilización de la media para el valor de la medida de desempeño C

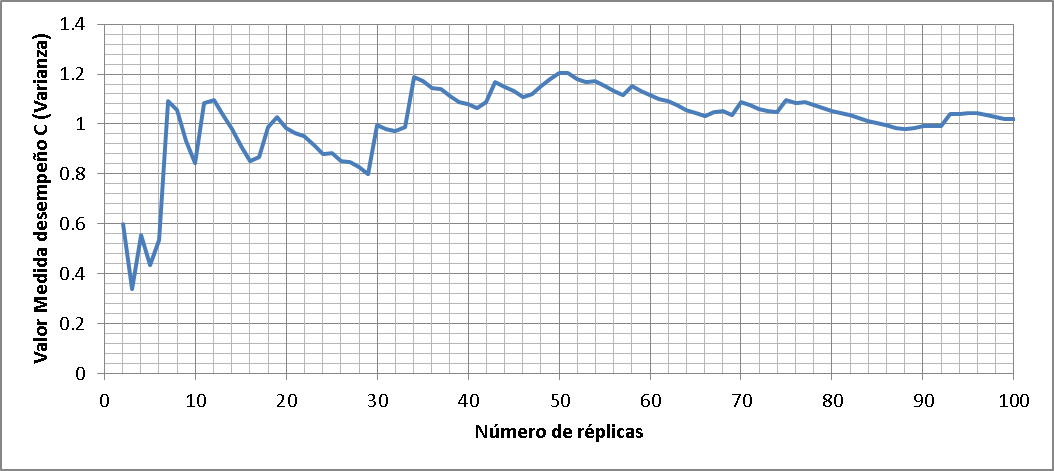


Ilustración 9: Estabilización de la media para el valor de la medida de desempeño C

A continuación, revisamos los valores más precisos para la media y varianza cambiando el número de réplicas consideradas. Notamos que los valores se van acercando a 4.41 y 1.021, para la media y varianza respectivamente.



Tabla 3: Media y Varianza para la medida de desempeño C al considerar n réplicas

### Comparación medida de desempeño B con y sin transiente

A continuación se muestran los gráficos en los cuales se compara la medida de desempeño B con y sin transiente:

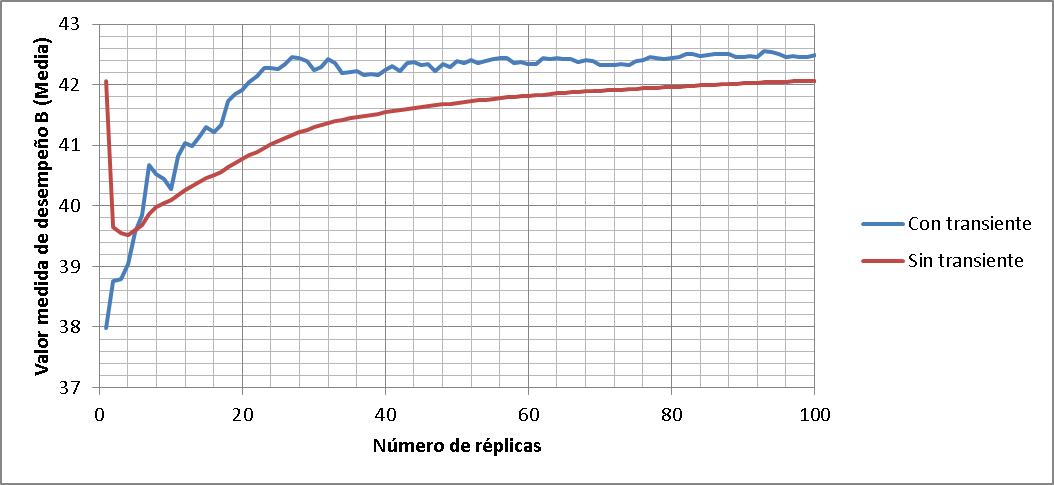


Ilustración 10: Estabilización de la media para la medida de desempeño B con y sin transiente

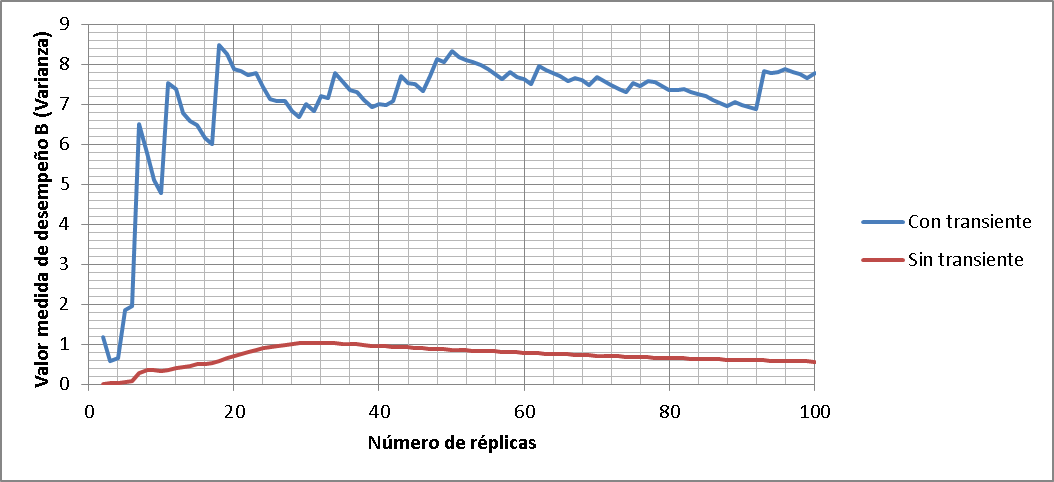


Ilustración 11: Estabilización de la varianza para la medida de desempeño B con y sin transiente

Como se puede apreciar, la estabilización se logra de manera mucho más “suave” al eliminar el periodo transiente. Mientras que las medias tienden a un valor parecido en ambos casos, notamos una disminución significativa de la varianza al eliminar el transiente.

## Intervalos de confianza y error relativo

Nos interesa saber qué tan confiable es nuestra estimación de las medidas de desempeño, para esto estimaremos el intervalo de confianza, que mientras más pequeño sea, mejor será, y también el error relativo que se está cometiendo. Para este punto se considerará una medida de desempeño B’, que corresponde a la medida de desempeño B sin considerar el periodo transiente, que, como ya se mencionó con anterioridad, significa eliminar las primeras 320 observaciones de cada réplica.

### Intervalos de confianza al 95%

En base a los valores ya obtenidos para las 4 medidas de desempeño, estableceremos el intervalo de confianza considerando 20 réplicas. En la siguiente tabla se muestran los valores de la media y varianza muestral, y valores mínimo y máximo para el intervalo:



Tabla 4: Media, varianza e intervalo de confianza para las medidas de desempeño con 20 réplicas

Cabe destacar que los valores para el intervalo se han calculado como:

Logramos notar que para el caso de la medida de desempeño B’, en la cual no se considera el periodo transiente, el intervalo de confianza es más pequeño, y, por lo tanto, más preciso.

### Error relativo

Para esto, usamos los mismos valores de media y varianza mostrados con anterioridad. El cálculo del error relativo tiene dos formas:

* Error relativo:
* Error relativo “Verdadero”:

Al hacer los cálculos, nos queda la siguiente tabla:



Tabla 5: Errores relativos para las medidas de desempeño con 20 réplicas

Se puede apreciar que la medida de desempeño C es la que nos genera un mayor error, a pesar de que su varianza es poca con respecto al resto. El problema se genera debido a que tiene mucha varianza con respecto a su propia media. Por otro lado, la medida de desempeño B’ nuevamente es la que “gana”, tiene un error relativo muy inferior a la del resto, en particular, la medida B que mide lo mismo pero considerando el transiente, tiene un error 3 veces mayor que su versión “optimizada”.

### Réplicas adicionales para error relativo menor al 1%

Nos interesa disminuir el error relativo para las 3 medidas de interés a menos del 1%. En el caso de la medida B’, como podemos apreciar, su error relativo en ambos casos es menor al 1%.

Entonces, nos interesa saber cuántas réplicas adicionales necesitamos para obtener dicho error relativo. Para esto consideraremos que la media y varianza no cambian mucho a partir de la réplica número 21, por lo que nos quedamos con los valores ya mencionados, y calculamos el error como:

Donde i es el número de réplicas totales que genera dicho error. Para el caso de error relativo “normal”, deberemos buscar el mínimo valor de i talque el error es menor que 0.01, en el caso de considerar el error relativo “verdadero”, deberá ser talque el error es menor que 0.00990099 (). El proceso nos llevará a evaluar uno por uno los valores de i hasta llegar al adecuado, y deberemos tomar el máximo i entre las 3 medidas. El proceso debiera iterar volviendo a calcular la varianza y media para este nuevo i y ver si se logra obtener un error relativo menor al deseado, sino, se debiera nuevamente calcular el número de réplicas adicionales y así sucesivamente hasta alcanzar el error. Sin embargo, en esta ocasión solo daremos la cantidad de réplicas adicionales para la primera iteración.

El proceso se puede revisar con más detalle en la hoja de cálculo. A continuación se presentan la cantidad de réplicas necesarias para alcanzar el error relativo para cada medida:



Tabla 6: Cantidad de réplicas adicionales para conseguir error relativo menor al 1%

Si consideramos el error relativo “normal”, entonces deberemos efectuar 1901 réplicas adicionales, mientras que si consideramos el error relativo “real”, deberemos realizar 1940 réplicas adicionales. Al quitar el periodo transiente, estas cantidades debieran disminuir considerablemente, de hecho, para la medida de desempeño B’ no es necesario realizar más réplicas para obtener el error deseado.

Suponiendo que el costo de generar las 20 réplicas ya es alto, no sería viable tener dicha precisión sin eliminar el periodo transiente.

## Batch means para la medida B

Para realizar Batch Means para la medida de desempeño B consideramos solo la primera réplica, y haremos grupos de 50 observaciones. La réplica contiene 2066 observaciones, por lo que haremos 41 grupos, y el último, que queda incompleto, lo desechamos. Para cada grupo calculamos la media muestral para la medida involucrada (Tiempo de permanencia en el taller). Esto nos arroja los siguientes valores:



Tabla 7: Promedios calculados para los distintos grupos

Con estos valores, calculamos el promedio y varianza que son 38.0015 y 87.3830 respectivamente. El intervalo lo calculamos de la misma forma que lo hicimos anteriormente, es decir, considerando la Student.

Al realizar los cálculos, nos queda que el intervalo 35.051 – 40.952, con un error relativo del 7.77%, y un error relativo “verdadero” del 8.42%, un poco más del doble calculado para la medida de desempeño B con anterioridad.

## Conclusiones

Hemos calculado el periodo transiente y hemos realizado la prueba de qué pasa si recalculamos la medida de desempeño B eliminando dicho periodo (la cual llamamos B’) considerando 20 réplicas. También, en una segunda instancia, calculamos la misma medida utilizando el método Batch Means.

Al comparar los 3 casos, gana por lejos la medida de desempeño B’, mientras le sigue la medida B normal, ambas considerando 20 réplicas. Para el caso Batch Means, y como era de esperarse, el resultado arroja un error mucho mayor, puesto que la cantidad de datos que nos ofrece una sola réplica no nos es suficiente para sacar conclusiones fiables.

Ante la imposibilidad de generar una mayor cantidad de réplicas debido a un alto costo -por ejemplo- el método de Batch Means de todas formas nos ayudará a estimar los valores deseados, y estos serán más precisos en la medida que la réplica considerada sea más larga. Sin embargo, ante la posibilidad de generar una cantidad mayor de réplicas, es recomendable, primero, calcular el periodo transiente, y a partir de esto y su eliminación, calcular las medidas de desempeño.

# Anexos

Todos los archivos mencionados en este documento, pueden ser visitados en el siguiente link: https://github.com/julioadriazola/Simulacion/tree/master/Entrega%203