

# Reporte práctica cinco: Método Monte-Carlo

José Anastacio Hernández Saldaña

Posgrado de Ingeniería de Sistemas

1186622

jose.hernandezsld@uanl.edu.mx

8 de septiembre de 2017

## Resumen

Este es un reporte sobre la práctica cinco con respecto al tema sobre el método Monte-Carlo que se realizó en la clase de Simulación de Sistemas, cómputo paralelo en R.

## 1. Tarea: Examinar el efecto del tamaño de la muestra en el método Monte-Carlo con la precisión y el tiempo de ejecución

El método Monte-Carlo es un método estadístico numérico usado para aproximar expresiones matemáticas complejas y computacionalmente costosas de manera eficiente. Consiste en usar un espacio delimitado donde se encuentra la función a evaluar, la idea principal del método es generar números aleatorios dentro de ese espacio y calcular la proporción de puntos que están dentro o fuera, por arriba o por debajo de la función y obtener así una aproximación del área cubierta. Esto puede ser utilizado para calcular integrales complejas, volúmenes complejos e incluso para hacer pronósticos.

### 1.1. Diseño del Experimento

Para el experimento se tomó como base el código de la página del curso, donde tenemos la función  $f(x) = \frac{1}{e^x + e^{-x}}$  de la cual queremos resolver  $\int_3^7 f(x)dx$  y utilizaremos el método Monte-Carlo para aproximar el valor de su integral, podemos ver esta gráfica en la figura 1.

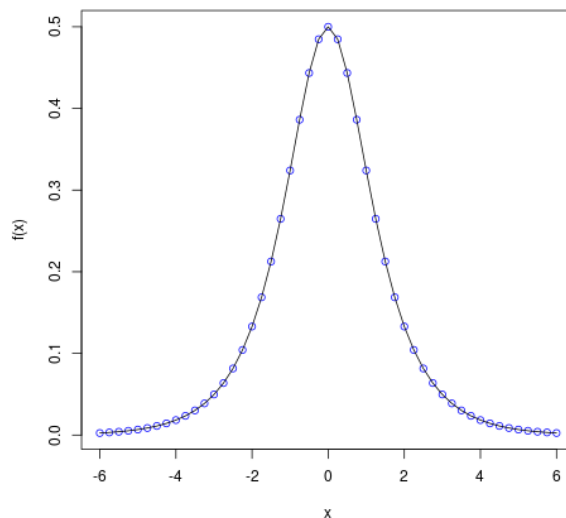


Figura 1: Gráfica de la función  $f(x) = \frac{1}{e^x + e^{-x}}$

La aproximación fue posible ya que la función  $g(x) = \int_{-\infty}^{\infty} \frac{2}{\pi} f(x) dx$  es una distribución de probabilidad válida, es decir, la integral de esa función es igual a uno; por esto podemos crear números aleatorios basados en ella. Utilizamos estos valores para calcular el área entre los límites de nuestra integral, usando la suma de los valores entre nuestros límites derecho e izquierdo tendremos el valor aproximado del área, esto se puede ver en la figura 2. Ahora solo queda hacer la transformación de nuestra distribución de probabilidad a nuestra función  $f(x)$ .

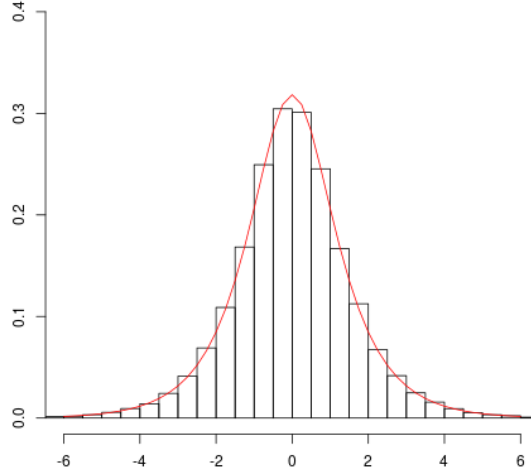


Figura 2: Gráfica de la función  $f(x) = \frac{2}{\pi(e^x + e^{-x})}$  y un área aproximada

Para comenzar la experimentación, se utilizó una computadora con las siguientes especificaciones, Procesador Intel Core i7-4790 CPU @ 3.6GHz  $\times$  8 y Memoria RAM de 24 GB. Utilizando solamente los cuatro núcleos físicos disponibles. Como el tamaño de la muestra nos da que tan delgadas son las barras que utilizaremos para calcular el área se varió el tamaño de la muestra  $m$  para poder identificar cómo afecta esto a la precisión del cálculo de la integral, dado que se trabajó una implementación paralela, se paralelizó un número fijo de réplicas  $r$ , para calcular el promedio de todas ellas y así obtener el valor aproximado. Este procedimiento es similar a incrementar el tamaño de la muestra en  $r$  dando el tamaño de la muestra por  $m \times r$ .

Los tamaños de  $m$  utilizados fueron 1,250, 2,500, 5,000, 10,000, 25,000, 50,000, 125,000 y el valor de  $r$  fue constante a 50, así que el total del tamaño de la muestra fueron 62,500, 125,000, 250,000, 500,000, 1,250,000, 2,500,000, 6,250,000. El porcentaje de separación entre el valor calculado  $vc$  y el valor real  $vr$  obtenido de la página de [www.wolframalpha.com](http://www.wolframalpha.com) se calculó a partir de la fórmula  $sep = \frac{vr - vc}{vr} \times 100$ . El tiempo se tomó en milisegundos y se hicieron 30 réplicas para cada muestra del experimento.

## 1.2. Resultados

Al revisar los resultados del experimento pudimos confirmar lo esperado con respecto al tamaño de la muestra y el porcentaje de separación, como podemos ver en la figura 3, ya que este disminuyó al aumentar el tamaño de la muestra, y no de una manera lineal, las variaciones disminuyeron conforme aumento el tamaño de la muestra ya que la aproximación fue más estable.

Con respecto al tiempo de ejecución, este aumentó con relación al tamaño de la muestra como se muestra en la figura 4, de una manera exponencial y las variaciones aumentaron; es decir, con muestras más grandes las variaciones de tiempo entre réplicas aumentaron, esto debido al cálculo de si el valor del elemento de la muestra estaba o no en la región a calcular, que fue lo que se paralelizó para este experimento.

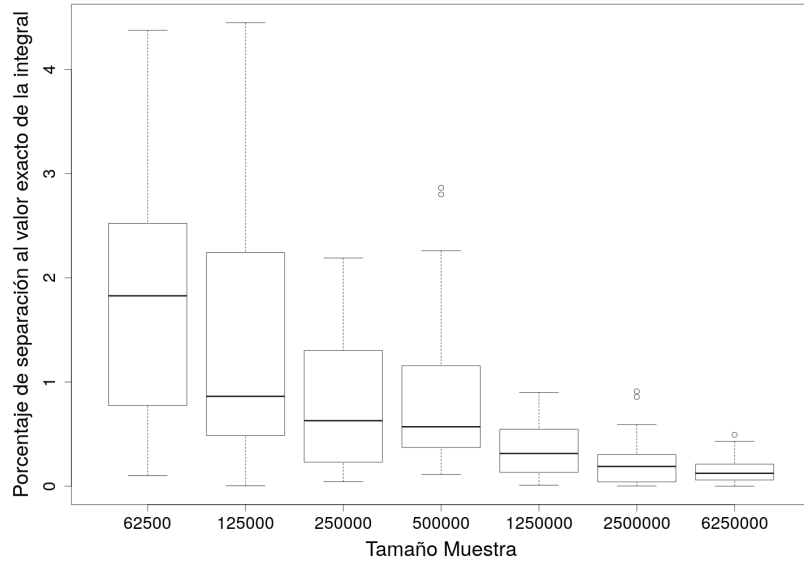


Figura 3: Variación del porcentaje de separación con respecto al tamaño de la muestra

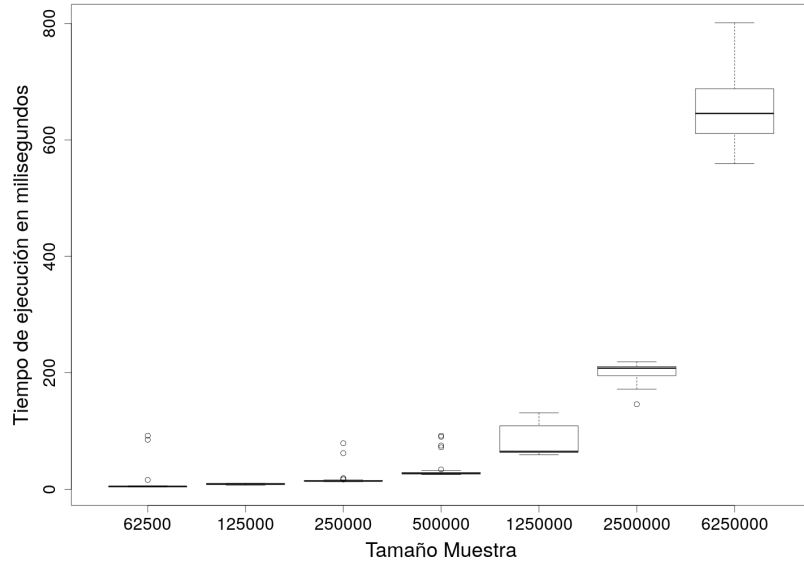


Figura 4: Variación del tiempo con respecto al tamaño de la muestra

## 2. Extra Uno: Aproximación del valor de $\pi$

El primer aporte extra para este reporte fue aproximar el valor de  $\pi$  por medio del método de Monte-Carlo, esta aproximación es muy común, pero no en una implementación paralela, así que se tomó como base el código sugerido en la implementación de Kurt, donde se tiene un área que va desde -1 a 1 tanto en el eje x como el eje y, y se quiere aproximar el área del círculo y de ahí obtener el valor aproximado de  $\pi$  como en la figura 5. También se desea conocer la separación al valor exacto de la aproximación y el tiempo que tarda en realizar la aproximación conforme se varía la muestra.

### 2.1. Diseño del Experimento

Para este experimento se utilizaron los mismos parámetros que el experimento anterior, se realizaron 30 réplicas y los mismos valores de muestra de 62,500, 125,000, 250,000, 500,000, 1,250,000,

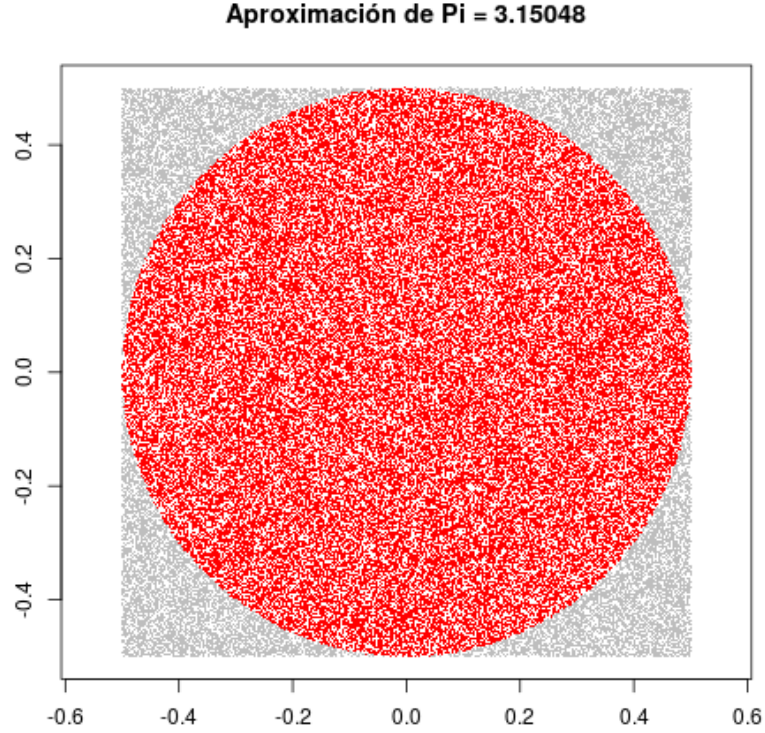


Figura 5: Ejemplo de una aproximación de  $\pi$

2;500,000 y 6;250,000 en el mismo equipo de cómputo. La paralelización se realizó en las muestras, se dividieron en 50 conjuntos de la misma cantidad de elementos que se realizaron en paralelo, después se promediaron los 50 valores para tener el valor aproximado con esa cantidad de muestras.

## 2.2. Resultados

Al igual que el experimento anterior, la separación del valor aproximado se redujo conforme aumentó el tamaño de la muestra como podemos ver en la figura 6. De igual manera, conforme se incrementó el tamaño de la muestra, el valor de la aproximación fue más estable y las réplicas tendían a estar cerca de la mediana.

Con respecto al tiempo, este se incrementó con relación al aumento del tamaño de la muestra como vemos en la figura 7, la variación de los valores de las réplicas también varió conforme aumentó el tamaño de la muestra por las variaciones de los valores que están dentro del círculo y los que no estuvieron adentro.

## 3. Conclusiones

Como se pudo apreciar en los experimentos, el método Monte-Carlos es muy eficiente con respecto a la aproximación de valores, ya que con tamaños de muestra que se puedan realizar en segundos podemos tener muy buenas aproximaciones y con la ayuda del cómputo paralelo podemos trabajar muestras más grandes y mejores aproximaciones.

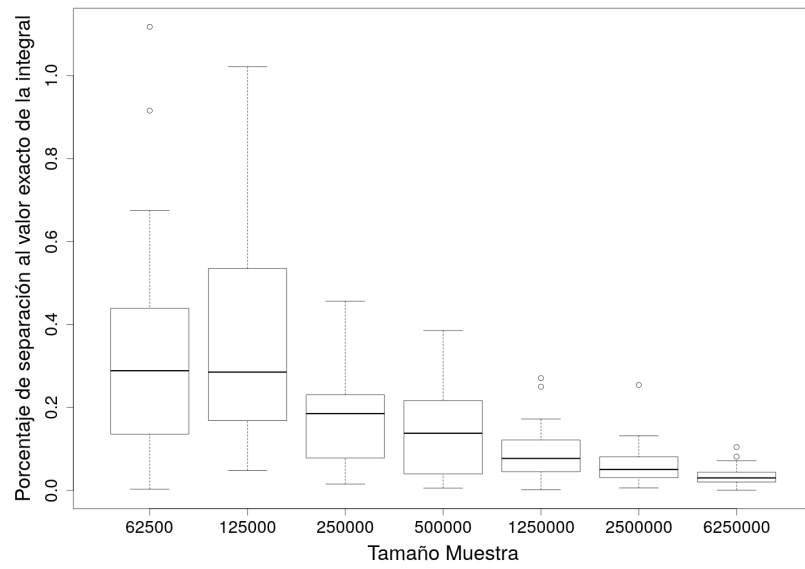


Figura 6: Variación del porcentaje de separación con respecto al tamaño de la muestra

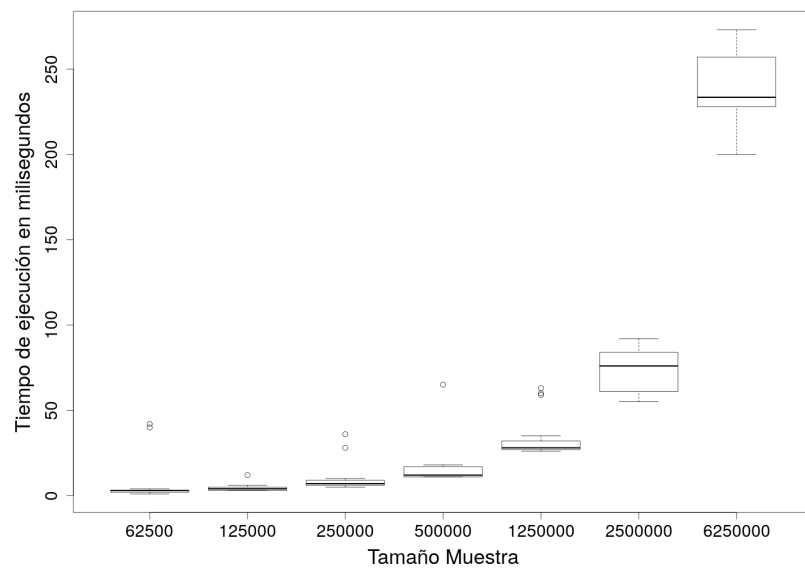


Figura 7: Variación del tiempo con respecto al tamaño de la muestra