Práctica 5: Método Monte Carlo

Gabriela Sánchez Y.

El método de Monte Carlo es un esquema dirigido hacia la estimacón de parámetros estocásticos o determinísticos con base en el muestreo aleatorio. La presente práctica muestra ejemplos de aplicaciones del mismo.

Cálculo de integrales

Sea

$$f(x) = \frac{1}{e^x + e^{-x}}.$$

El primer ejemplo consiste en calcular la integral definida (1)

$$\int_3^7 f(x)dx. \tag{1}$$

Ya que

$$\frac{2}{\pi}\int_{\infty}^{\infty} f(x)dx = 1,$$

la función

$$g(x) = \frac{2}{\pi}f(x),$$

es una función de distribución, de manera que se pueden generar números pseudoaleatorios que sigan esta distribución.

El experimento consiste en generar una determinada cantidad de números con la función $g(x)$, sumando todos los que se encuentren dentro del rango en el cual se desea calcular la integral, repitiendo el experimento varias veces. El promedio de los resultados es un estimado de la integral

$$\int_3^7 g(x)dx,$$

por lo que para determinar el valor de (1) sólo resta multiplicar por el factor $\frac{\pi}{2}$.

Una pregunta que surge de manera natural es, ¿de qué tamaño debe ser la muestra para obtener una buena precisión?. Para responder esta cuestión se examina el efecto del tamaño de muestra en la precisión del estimado, comparando con Wolfram Alpha, por un lado y por otro lado en el tiempo de ejecución.

Se toma como valor real el obtenido mediante Wolfram Alpha el cual es $V_r = 0{,}04883411$. Para determinar la precisión del estimado V_a, se calcula el error relativo que determina el porcentaje de error el la precisión y se define como:

$$\frac{V_r - V_a}{V_r} * 100.$$

Se repite el experimento 500 veces para cada variación en el tamaño de la muestra, donde éstos varían de acuerdo al siguiente vector:

```
tamMuestra <- c(1000, 5000, 10000, 50000, 100000, 500000,
                1000000)
```

En la figura 1 se grafican los errores relativos correspondientes a cada muestra. Los resultados indican que el aumento en el tamaño de la muestra disminuye el error relativo, esto es, mejora la precisión en el estimado.

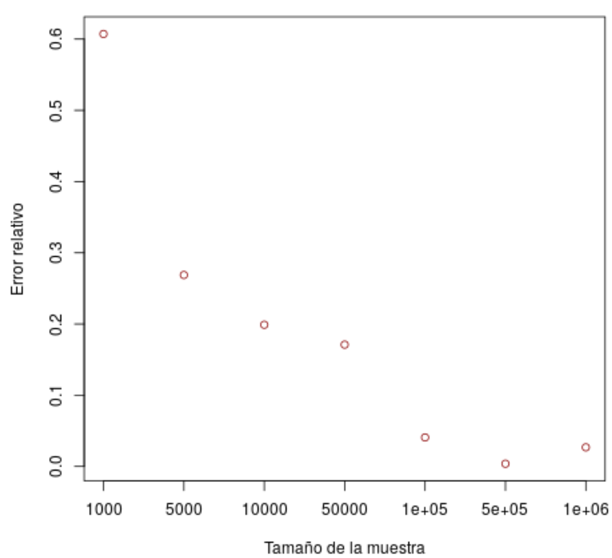

Figura 1: Errores relativos para los diferentes tamaños de la muestra.

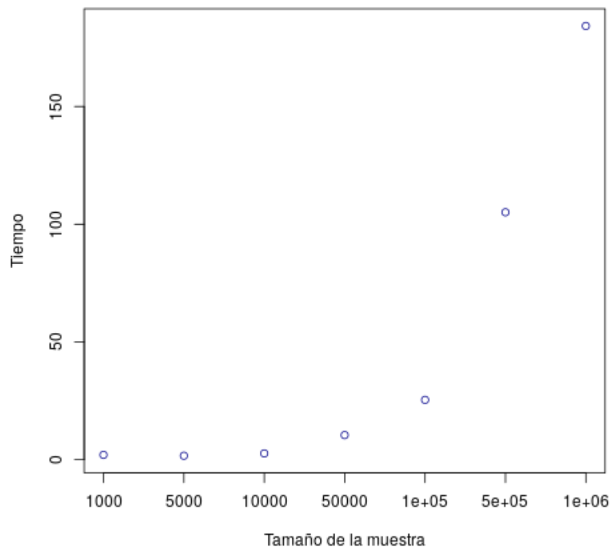

Figura 2: Tiempo de ejecución del experimento para cada tamaño de muestra.

La figura 2 muestra el tiempo total que tardó cada experimento para los diferentes tamaños de la muestra, obsérvese que, como era de esperarse, mientras más grande sea la muestra, más tardardado será el cálculo.

Estimación del valor de π

Recuérdese que el área de un círculo de radio r es

$$A_c = \pi r^2.$$

Si se dibuja un cuadrado que contenga este círculo, entonces su área está dada mediante

$$A_s = 4r^2.$$

Conociendo esto, para obtener π lo único que debe hacerse es dividir ambas áreas de donde se tiene el valor $\frac{\pi}{4}$, que al multiplicar por cuatro da el valor buscado.

El método de Monte Carlo en este caso consiste en generar números que se encuentren en el cuadrado de lado $2r$, guardando aquellos que pertenezcan al

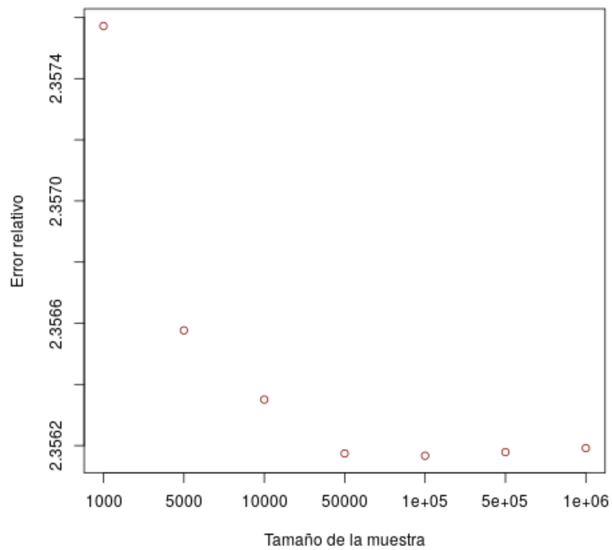

Figura 3: Errores relativos en el cálculo de π

cículo unitario circunscrito en el cuadrado. El promedio de éstos, al multiplicarse por cuatro, es el valor de π.

De la misma manera que en el ejemplo anterior, se repite el experimento variando las muestras, haciendo 500 réplicas en cada caso. Se calculan los tiempos de ejecución y los errores relativos.

En las figuras 3 y 4 se muestran los resultados en los errores relativos y tiempos de ejecución, respectivamente, para las simulaciones con distintos tamaños en la muestra. Los resultados obtenidos son los esperados; mientras la muestra incrementa su tamaño, el tiempo de ejecución crece y el error relativo disminuye.

Pronósticos

El ejemplo utiliza el método Monte Carlo para pronósticar, en base los casos de Zika reportados en el estado de Oaxaca por semana desde la primera semana epidemiológica del 2016 hasta la semana epidemiológica 34 del mismo año, los casos en semanas posteriores.

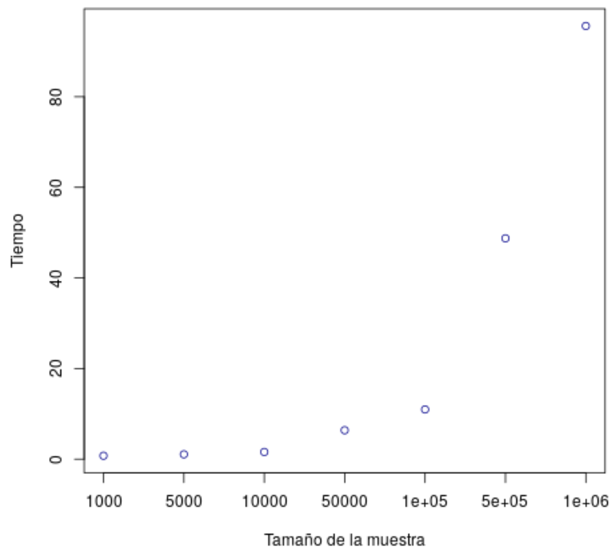

Figura 4: Tiempo de ejecución de la simulación para el cálculo de π

La figura 5 gráfica los datos dados. Para hacer un pronóstico son necesarios datos que se comporten como los datos de casos de Zika. Para entender mejor el comportamiento de nuestros datos observamos la figura 6. Se podría suponer que siguen una distribución exponencial o bien, una distribución normal.

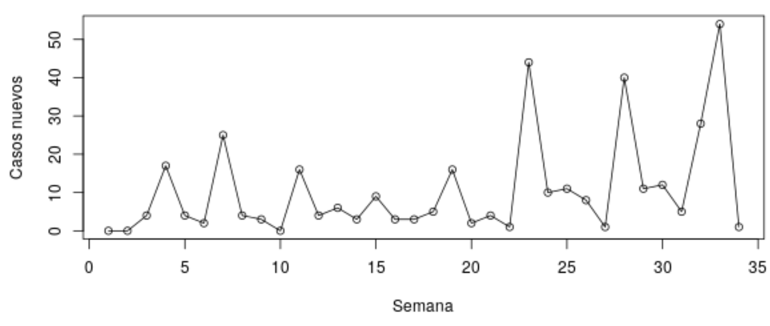

Figura 5: Casos reportados de Zika por semana.

Como primera aproximación se usa una distribución exponencial con media dada por los casos de Zika reportados por semana. Una comparación entre los

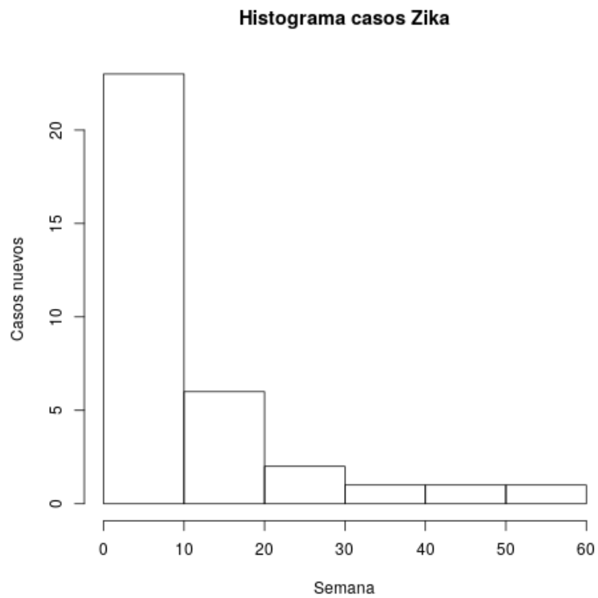

Figura 6: Histograma de los casos reportados por semana de Zika.

datos obtenidos del boletín y los obtenidos de la distribución se muestra en la figura 7. Los datos en azul son los obtenidos mediante la distribución.

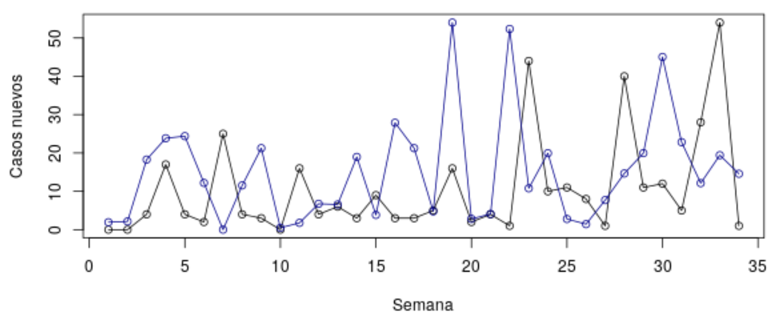

Figura 7: Comparación entre datos.

Resta utilizar el método de Monte Carlo para la predicción.