

Informe: Importance Sampling

Integrantes:

- Facello Ignacio
- Nieto Manuel

Introducción

Los métodos tradicionales de integración por el método de Monte Carlo no siempre son los más eficientes.

El método de **Muestreo de Importancia** o **Importance Sampling** es una forma de hacer que las simulaciones de Monte Carlo converjan más rápido. La idea del mismo se centra en elegir una distribución diferente a la uniforme para muestrear puntos que puedan generar datos más relevantes.

Recordemos que Monte Carlo busca aproximar una expresión θ desconocida, sabiendo que este valor puede calcularse como $\theta = E[h(X)]$ para cierta variable aleatoria X que posee una distribución F .

Luego la teoría detrás de Importance Sampling se puede resumir en el siguiente resultado, teniendo en cuenta que f_X representa la función de densidad de una variable X y g_Y la función de densidad una variable Y . Denominamos a f_X como la función **objetivo** y a g_Y como la función de **importancia**.

$$\begin{aligned}\theta &= E[h(X)] \\ &= \int_{-\infty}^{\infty} h(t) f_X(t) dt \\ &= \int_{-\infty}^{\infty} h(t) f_X(t) \left(\frac{g_Y(t)}{g_Y(t)} \right) dt \\ &= \int_{-\infty}^{\infty} \frac{h(t) f_X(t)}{g_Y(t)} g_Y(t) dt \\ &= E \left[\frac{h(Y) f_X(Y)}{g_Y(Y)} \right]\end{aligned}$$

Cabe destacar que el método de Importance Sampling no siempre converge más rápido que Monte Carlo y depende mucho de la función de importancia que se elija.

A partir de la bibliografía recomendada^[1] para este trabajo y algunas investigaciones adicionales^{[2][3]}, vimos que para lograr una más rápida convergencia es importante elegir una distribución Y tal que g_Y cumpla que $Var \left[\frac{h(Y) f_X(Y)}{g_Y(Y)} \right] < Var[h(X)]$. O coloquialmente, $g_Y(t)$ debe ser alto donde $|h(t) f_X(t)|$

sea alto. Por lo general no es sencillo encontrar una función g_Y con esta propiedad.

El objetivo de este proyecto fue comparar el desempeño de Importance Sampling para distintas funciones de importancia entre sí y con respecto a Monte Carlo tradicional para estimar $P(Z > 3)$ con $Z \sim \mathcal{N}(0, 1)$ y $P(W > 10)$ con $W \sim \Gamma(9, \frac{1}{2})$.

Algoritmo

Para la implementación de este método nos basamos en la de Monte Carlo del apunte y en la justificación teórica de Importance Sampling que desarrollamos en la introducción.

Recordemos que para Monte Carlo, se genera una muestra de n puntos de una variable uniforme continua $U \sim \mathcal{U}(0, 1)$ y luego se evalúa la expresión $h(t)f_X(t)$ para cada valor de la muestra. El promedio de las evaluaciones tiende a θ .

Por otro lado, en nuestro algoritmo se genera una muestra de n puntos de una variable aleatoria continua Y y luego se evalúa la expresión $\frac{h(t)f_X(t)}{g_Y(t)}$ para cada valor de la muestra. Por la justificación teórica, el promedio de las evaluaciones sigue tendiendo a θ .

A continuación presentamos nuestro algoritmo de Importance Sampling simplificado en pseudocódigo. Sean:

- `Y()` un generador de variables Y
- `g_Y` la función de densidad para la distribución Y
- `f_X` la función de densidad para la distribución original X
- `h` la transformación correspondiente de Monte Carlo

```
def importance_sampling(n):  
    integral = 0  
    for _ in range(n):  
        y = Y() # Variable aleatoria con distribución Y  
        integral += f_X(y) * h(y) / g_Y(y)  
    return integral/n
```

Resultados

Para comparar la velocidad de convergencia entre los métodos, se realizaron 50 estimaciones con distintos valores de n equidistantes, que van desde 10 mil a 500 mil con un incremento de 10 mil por estimación.

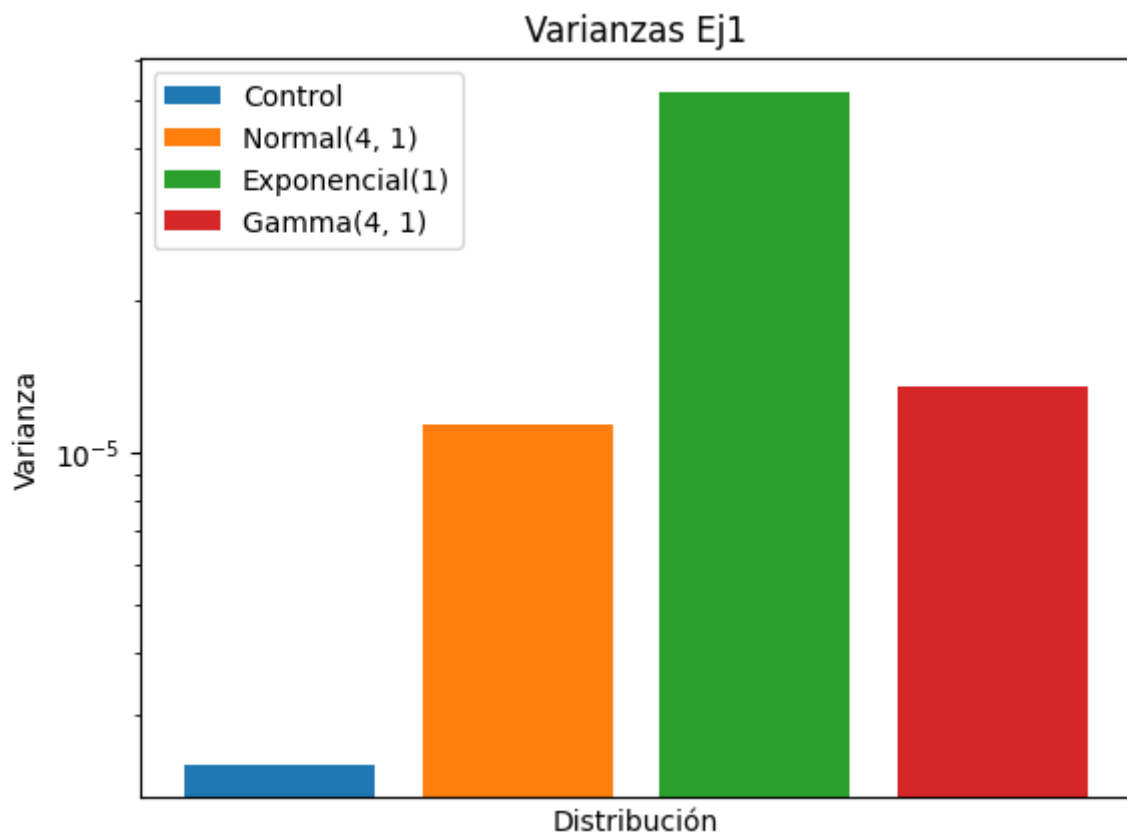
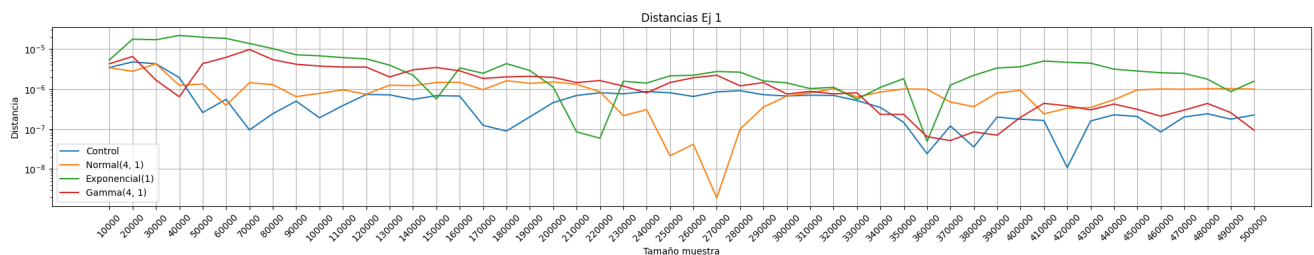
Definimos una estimación como: Ejecutar el algoritmo 10 veces para un mismo n y tomar el promedio sobre los valores obtenidos.

En los siguientes gráficos podemos observar la distancia entre dichas estimaciones y el valor real de θ para cada caso, calculado como el valor absoluto entre la diferencia de ambos. Cabe destacar que en los gráficos, *Control* se refiere a la estimación realizada utilizando el método de Monte Carlo estándar.

Ejercicio 1

Para este inciso buscamos aproximar $\theta = P(Z > 3)$ con $Z \sim \mathcal{N}(0, 1)$. Tomamos las siguientes distribuciones para analizar nuestra implementación del algoritmo de Importance Sampling:

- *Normal* $\sim \mathcal{N}(4, 1)$
- *Exponencial* $\sim \mathcal{E}(1)$
- *Gamma* $\sim \Gamma(4, 1)$



Observaciones:

Caso	Convergencia	Distancia	Varianza
Control	Media	Cercana al valor real	0.24006e-5
Normal	Lenta	Muy cercana en n=270.000	1.1345e-5
Gamma	Rápida	Alejada al principio	5.2045e-5
Exponencial	Muy lenta	Lejana al valor real	1.3476e-5

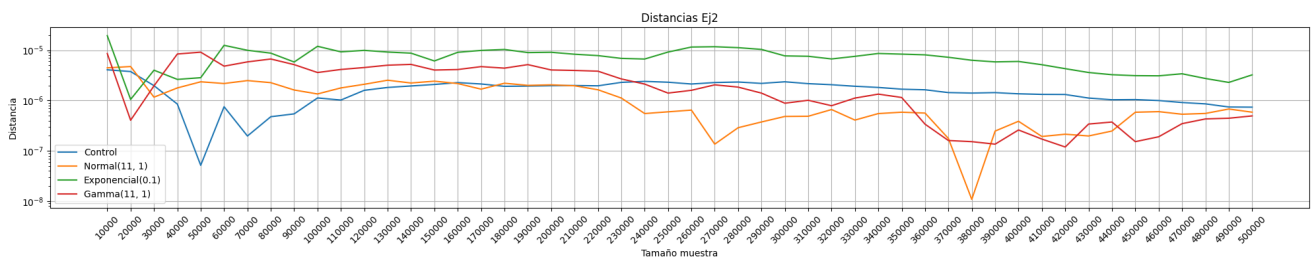
Conclusiones:

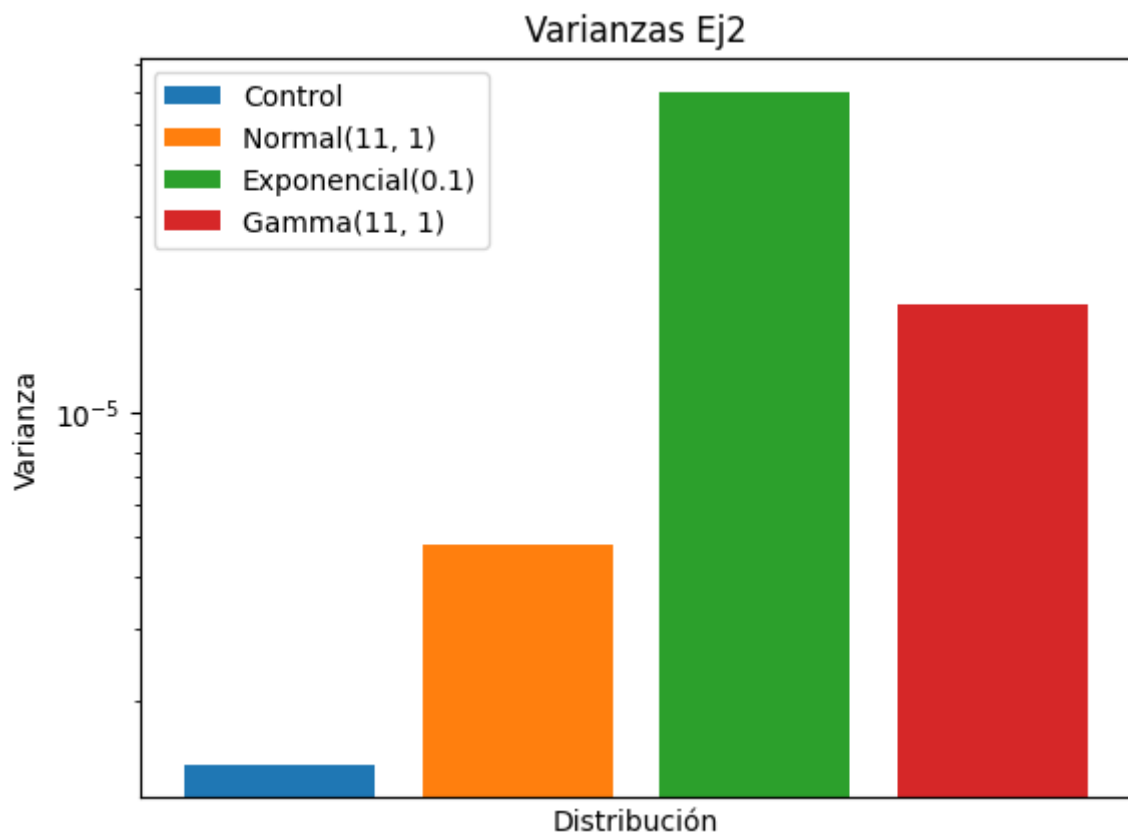
Vemos que para la mayoría de las distribuciones elegidas, Importance Sampling tiene una efectividad menor que la de Monte Carlo tradicional. Esto coincide con lo esperado al observar las varianzas muestrales de los distintos estimadores. Sin embargo, se observa un comportamiento extraño al comparar los estimadores Normal y Gamma, dado que viendo sus varianzas esperaríamos un mejor comportamiento de la primera con respecto a la segunda.

Ejercicio 2

Para este inciso tenemos $\theta = P(W > 10)$ con $W \sim \Gamma(9, \frac{1}{2})$. Las distribuciones que elegimos fueron:

- $Normal \sim \mathcal{N}(11, 1)$
- $Exponencial \sim \mathcal{E}(1/10)$
- $Gamma \sim \Gamma(11, 1)$





Observaciones:

Caso	Convergencia	Distancia	Varianza
Control	Lenta	Alejada al valor real	0.14126e-5
Normal	Rápida, pero más lenta que la Gamma	Es la más cercana en un punto	0.47747e-5
Gamma	Rápida	Alejada al principio	9.6210e-5
Exponencial	Lenta	Es la más lejana al valor real	1.8404e-5

Conclusiones:

En este inciso, observamos que con todas las distribuciones elegidas el método de IS tiene una efectividad mayor o igual a la de Monte Carlo tradicional, lo cual contradice con lo esperado al ver las varianzas muestrales de los estimadores.

Adicionalmente, podemos observar un comportamiento similar entre los estimadores Normal y Gamma (como el inciso 1), siendo en este caso aún más marcada la diferencia.

Conclusiones finales

Recordemos que anteriormente habíamos planteado la expectativa que dada una variable aleatoria $\frac{h(Y)f_x(Y)}{g_Y(Y)}$ con varianza pequeña esperaríamos que el

aproximador $\hat{\theta}_1 = E \left[\frac{h(Y)f_x(Y)}{g_y(Y)} \right]$ sea más eficiente que $\hat{\theta}_2 = E[h(x)]$.

Las observaciones que realizamos, especialmente para el inciso 2, no coinciden con lo que esperaríamos ver en relación con las varianzas muestrales de los distintos estimadores. Esto nos lleva a pensar que existe una relación más compleja entre el desempeño del método y las distribuciones que se elijen, pero desarrollar cuál es esta posible relación escapa a los objetivos de este trabajo.

En conclusión, el método de Importance Sampling no es inherentemente mejor que el método de Monte Carlo tradicional o viceversa. La realidad es que el método de IS es muy dependiente de las distribuciones que se elijan para el muestreo. Existen criterios para identificar distribuciones que puedan resultar en estimadores más eficientes, como el de la varianza, pero como pudimos observar su aplicación directa no asegura mejores resultados.

-
1. Capítulo 8 del libro Simulación (Segunda Edición ed.) de S. Ross (1999). ↩
 2. Importance Sampling (<https://www.youtube.com/watch?v=C3p2wl4RAi8>) ↩
 3. Importance sampling explained in 4 minutes (<https://www.youtube.com/watch?v=7A2jXWmnUFw>) ↩