

EST-24107: Simulación

Profesor: Alfredo Garbuno Iñigo — Primavera, 2022.

Objetivo. Estudiar integración numérica en un contexto probabilístico. Estudiar, en particular, el método Monte Carlo y entender sus bondades y limitaciones como un método de aproximación de integrales.

Lectura recomendada: Una lectura mas técnica sobre reglas de cuadratura se puede encontrar en la sección 3.1 de Reich and Cotter [1]. Y una buena referencia (técnica) sobre el método Monte Carlo lo encuentran en Sanz-Alonso et al. [2].

1. INTRODUCCIÓN

En muchas aplicaciones nos interesa poder resolver integrales de manera numérica. Estas pueden ser de cualquier forma. Por ejemplo, nos puede interesar resolver

$$\int_{\Theta} h(\theta) d\theta, \quad (1)$$

que bien puede ser reexpresada como una integral bajo una medida de probabilidad. Por ejemplo,

$$\int_{\Theta} f(\theta) \pi(\theta) d\theta, \quad (2)$$

de tal forma que podemos pensar en la ecuación de arriba como un valor esperado de una variable $\theta \sim \pi(\cdot)$ y calcular mediante un método numérico.

- La pregunta clave (I) es: ¿qué distribución podemos utilizar?
- La pregunta clave (II) es: ¿con qué método numérico resuelvo la integral?
- La pregunta clave (III) es: ¿y si no hay método numérico?

2. INTEGRACIÓN NUMÉRICA

Recordemos la definición de integrales Riemann:

$$\int f(x) dx \approx \sum_{n=1}^N f(u_n) \Delta u_n =: \hat{\pi}_N^R(f).$$

La aproximación utilizando una malla (cuadrícula) de N puntos sería:

$$\sum_{n=1}^N f(u_n) \Delta u_n.$$

El método es útil cuando las integrales se realizan cuando tenemos pocos parámetros. Es decir, cuando el dominio de integración es $\mathcal{X} \subseteq \mathbb{R}^p$ con p pequeña.

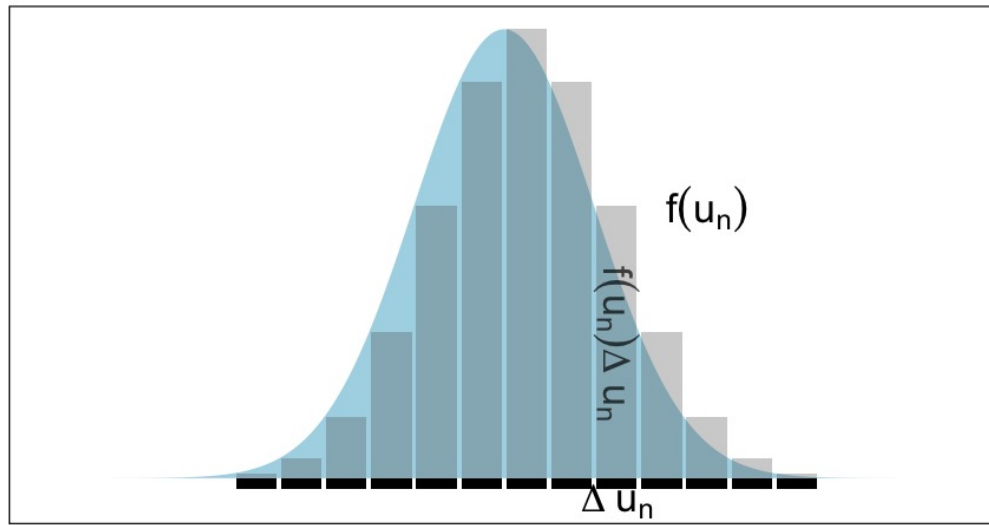


FIGURA 1. Integral por medio de discretización con $N = 11$.

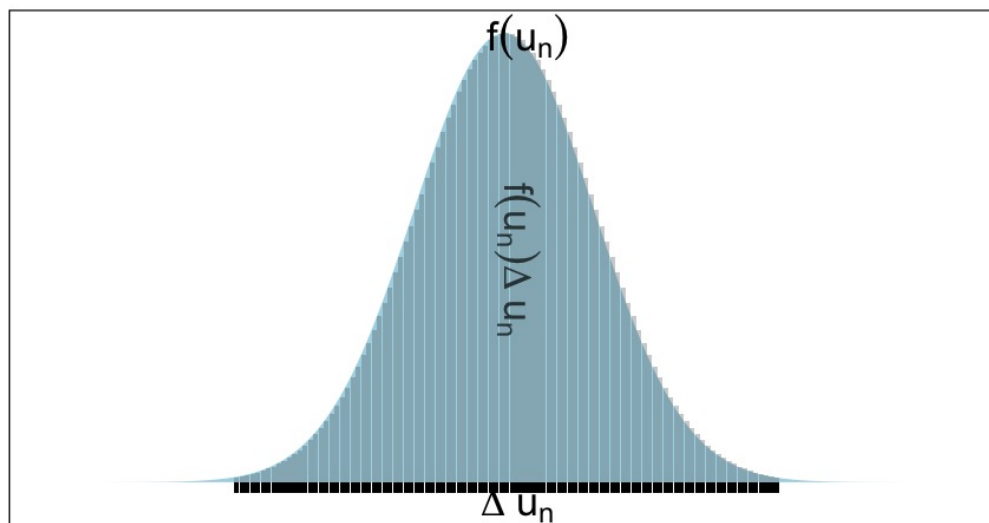


FIGURA 2. Integral por medio de una malla fina, $N = 101$.

3. ANÁLISIS DE ERROR

El concepto de integrabilidad de Darboux nos puede ayudar a acotar el error cometido por nuestra estrategia de integración. Por ejemplo, para una partición ρ_N (malla) del intervalo tenemos que

$$L_{f,\rho_N} \leq \hat{\pi}_N^R(f) \leq U_{f,\rho_N}. \quad (3)$$

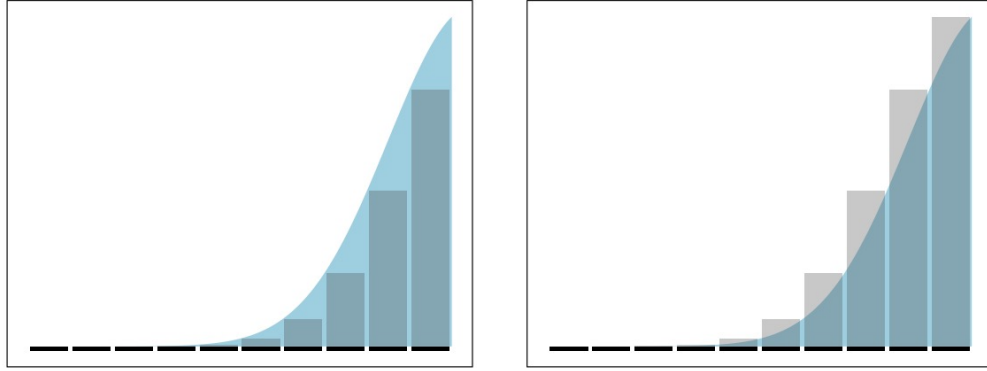


FIGURA 3. Integrales y cotas de Darboux.

Lo que recordarán de sus cursos de cálculo es que

$$\lim_{N \rightarrow \infty} |U_{f,\rho_N} - L_{f,\rho_N}| = 0, \quad (4)$$

y que además se satisface

$$\int f(x)dx = \lim_{N \rightarrow \infty} U_{f,\rho_N} = \lim_{N \rightarrow \infty} L_{f,\rho_N}. \quad (5)$$

3.1. Más de un parámetro

Consideramos ahora un espacio con $\theta \in \mathbb{R}^p$. Si conservamos N puntos por cada dimensión, ¿cuántos puntos en la malla necesitaríamos? Lo que tenemos son recursos computacionales limitados y hay que buscar hacer el mejor uso de ellos. En el ejemplo, hay zonas donde no habrá contribución en la integral.

3.2. Reglas de cuadratura

Por el momento hemos escogido aproximar las integrales por medio de una aproximación con una malla uniforme. Sin embargo, se pueden utilizar aproximaciones

$$\int f(x)dx \approx \sum_{n=1}^N f(\xi_n) \omega_n.$$

Estas aproximaciones usualmente se realizan para integrales en intervalos cerrados $[a, b]$. La regla de cuadratura determina los pesos ω_n y los centros ξ_n pues se escogen de acuerdo a ciertos criterios de convergencia.

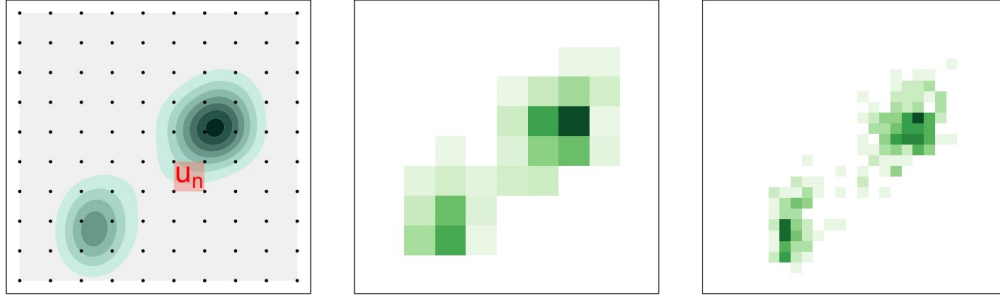


FIGURA 4. Integral multivariada por método de malla.

Por ejemplo, se consideran polinomios que aproximen con cierto grado de precisión el integrando. Los pesos y los centros se escogen de acuerdo a la familia de polinomios. Pues para cada familia se tienen identificadas las mallas que optimizan la aproximación. Ver sección 3.1 de Reich and Cotter [1].

4. INTEGRACIÓN MONTE CARLO

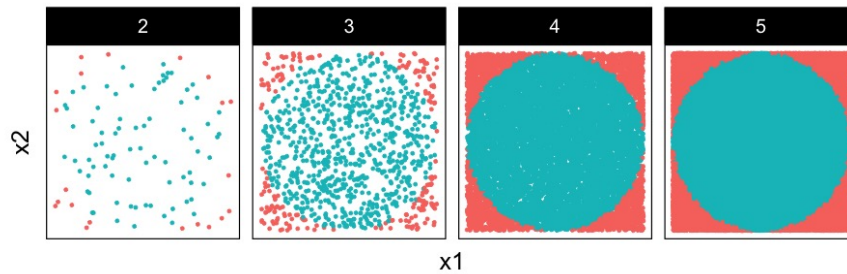
$$\pi(f) = \mathbb{E}_\pi[f] = \int f(x)\pi(x)dx,$$

$$\pi_N^{\text{MC}}(f) = \frac{1}{N} \sum_{n=1}^N f(x^{(n)}), \quad \text{donde } x^{(n)} \stackrel{\text{iid}}{\sim} \pi, \quad \text{con } n = 1, \dots, N,$$

$$\pi(f) \approx \pi_N^{\text{MC}}(f).$$

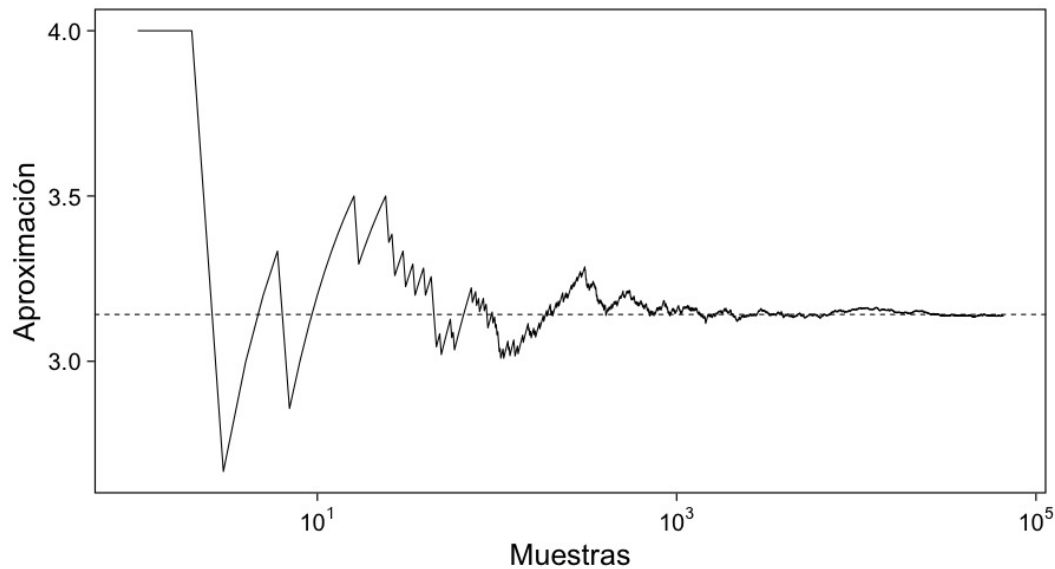
4.1. Ejemplo: Dardos

Consideremos el experimento de lanzar dardos uniformemente en un cuadrado de tamaño 2, el cual contiene un círculo de radio 1.

FIGURA 5. Integración Monte Carlo para aproximar π .

Si escogemos N suficientemente grande entonces nuestro promedio converge a la integral. En Fig. 6 se muestra para cada n en el eje horizontal cómo cambia nuestra estimación $\hat{\pi}_n^{\text{MC}}(f)$.

También podemos replicar el experimento unas M veces y observar cómo cambiaría nuestra estimación con distintas semillas. Por ejemplo, podemos replicar el experimento 10

FIGURA 6. Estimación $\pi_N^{MC}(f)$ con $N \rightarrow \infty$.

veces. En R y python lo usual es utilizar arreglos multidimensionales para poder guardar muestras bajo distintas replicas.

```

1 set.seed(108)
2 nsamples ← 10**4; nexp ← 50
3 U ← runif(nexp * 2 * nsamples)
4 U ← array(U, dim = c(nexp, 2, nsamples))
5 apply(U[1:5,,], 1, str)

```

```

1 num [1:2, 1:10000] 0.4551 0.7159 0.164 0.0627 0.5291 ...
2 num [1:2, 1:10000] 0.404 0.2313 0.9282 0.0426 0.0883 ...
3 num [1:2, 1:10000] 0.351 0.739 0.449 0.658 0.369 ...
4 num [1:2, 1:10000] 0.664 0.984 0.627 0.762 0.185 ...
5 num [1:2, 1:10000] 0.4635 0.6107 0.0115 0.7251 0.0117 ...
6 NULL

```

```

1 resultados ← apply(U, 1, function(x){
2   dardos ← apply(x**2, 2, sum)
3   exitos ← ifelse(dardos ≤ 1, 1, 0)
4   prop ← cummean(exitos)
5   4 * prop
6 })

```

Lo cual nos permite realizar distintos escenarios posibles.

Bajo ciertas consideraciones teóricas podemos esperar un buen comportamiento de nuestro estimador de la integral. E incluso podríamos (si el número de simulaciones lo permite) aproximar dicho comportamiento utilizando distribuciones asintóticas, (CLT).

4.2. Propiedades

4.2.1. Teorema [Error Monte Carlo] Sea $f : \mathbb{R}^p \rightarrow \mathbb{R}$ cualquier función bien comportada[†]. Entonces, el estimador Monte Carlo es **insesgado**. Es decir, se satisface

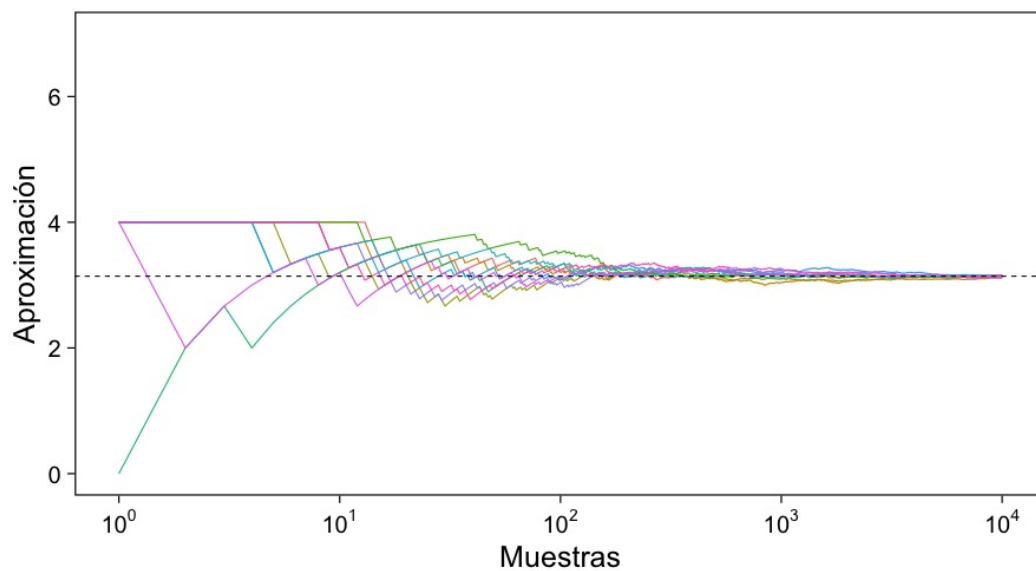


FIGURA 7. Réplica de las trayectorias de diversas realizaciones de la aproximación de la integral.

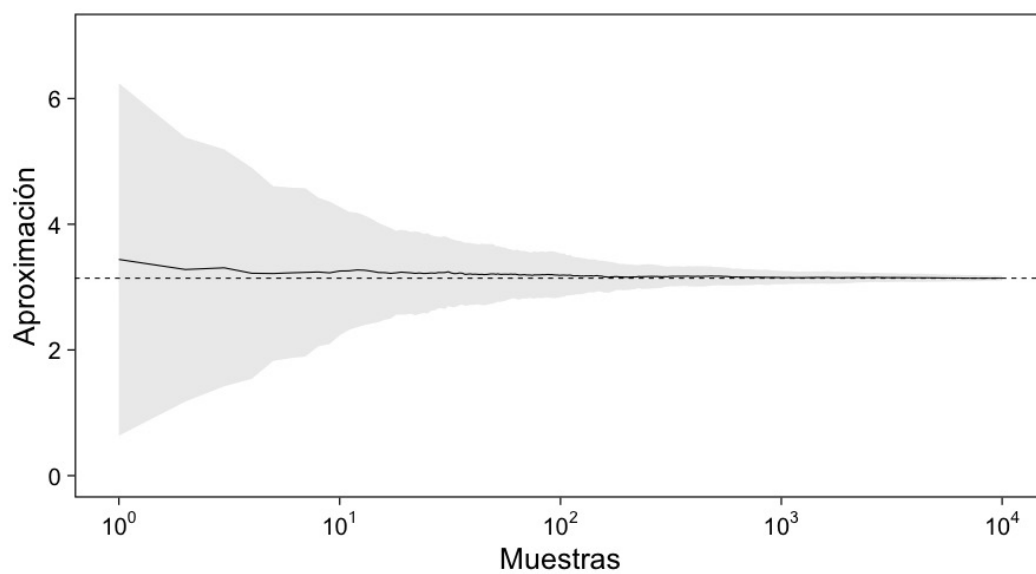


FIGURA 8. Comportamiento promedio e intervalos de confianza.

$$\mathbb{E} \left[\hat{\pi}_N^{\text{MC}}(f) - \pi(f) \right] = 0, \quad (6)$$

para cualquier N . Usualmente estudiamos el error en un escenario pesimista donde medimos el **error cuadrático medio** en el peor escenario

$$\sup_{f \in \mathcal{F}} \mathbb{E} \left[\left(\hat{\pi}_N^{\text{MC}}(f) - \pi(f) \right)^2 \right] \leq \frac{1}{N}.$$

En particular, la varianza del estimador (**error estándar**) satisface la igualdad

$$\text{ee}^2 \left(\hat{\pi}_N^{\text{MC}}(f) \right) = \frac{\mathbb{V}_\pi(f)}{N}.$$

4.2.2. Teorema [TLC para estimadores Monte Carlo] Sea f una función **bien comportada** ^{††}, entonces bajo una N suficientemente grande tenemos

$$\sqrt{N} \left(\hat{\pi}_N^{\text{MC}}(f) - \pi(f) \right) \sim \mathbf{N}(0, \mathbb{V}_\pi(f)). \quad (7)$$

4.2.3. Nota: El estimador Monte Carlo del que hablamos, $\hat{\pi}_N^{\text{MC}}(f)$, es una estimación con una muestra finita de simulaciones. En ese sentido podemos pensar que tenemos un *mapeo*

$$(x^{(1)}, \dots, x^{(N)}) \mapsto \hat{\pi}_N^{\text{MC}}(f), \quad (8)$$

con $x^{(i)} \stackrel{\text{iid}}{\sim} \pi$.

De lo cual es natural pensar: ¿y si hubiéramos observado otro conjunto de simulaciones?

En este sentido nos preguntamos por el **comportamiento promedio** bajo distintas muestras observadas

$$\mathbb{E}[\hat{\pi}_N^{\text{MC}}(f)] = \mathbb{E}_{x_1, \dots, x_N}[\hat{\pi}_N^{\text{MC}}(f)]. \quad (9)$$

De la misma manera nos podemos preguntar sobre la dispersión alrededor de dicho promedio (varianza)

$$\mathbb{V}[\hat{\pi}_N^{\text{MC}}(f)] = \mathbb{V}_{x_1, \dots, x_N}[\hat{\pi}_N^{\text{MC}}(f)]. \quad (10)$$

Al ser un ejercicio de **estimación** la desviación estándar del estimador recibe el nombre de error estándar. Lo cual denotamos por

$$\text{ee}[\hat{\pi}_N^{\text{MC}}(f)] = \left(\mathbb{V}[\hat{\pi}_N^{\text{MC}}(f)] \right)^{1/2}. \quad (11)$$

4.2.4. Nota: Para algunos estimadores la fórmula del error estándar se puede obtener de manera analítica (curso de **Inferencia Matemática**). Para otro tipo, tenemos que utilizar propiedades asintóticas (p.e. cota de Cramer-Rao).

Hay casos en los que no existe una fórmula asintótica o resultado analítico, podemos usar simulación [8] para cuantificar dicha dispersión (lo veremos en otra sección del curso).

4.2.5. Nota: Hay situaciones en las que la distribución normal asintótica no tiene sentido. Para este tipo de situaciones también veremos cómo podemos utilizar simulación para cuantificar dicha dispersión.

REFERENCIAS

- [1] S. Reich and C. Cotter. *Probabilistic Forecasting and Bayesian Data Assimilation*. Cambridge University Press, Cambridge, 2015. ISBN 978-1-107-06939-8 978-1-107-66391-6. [1](#), [4](#)
- [2] D. Sanz-Alonso, A. M. Stuart, and A. Taeb. Inverse Problems and Data Assimilation. *arXiv:1810.06191 [stat]*, jul 2019. [1](#)

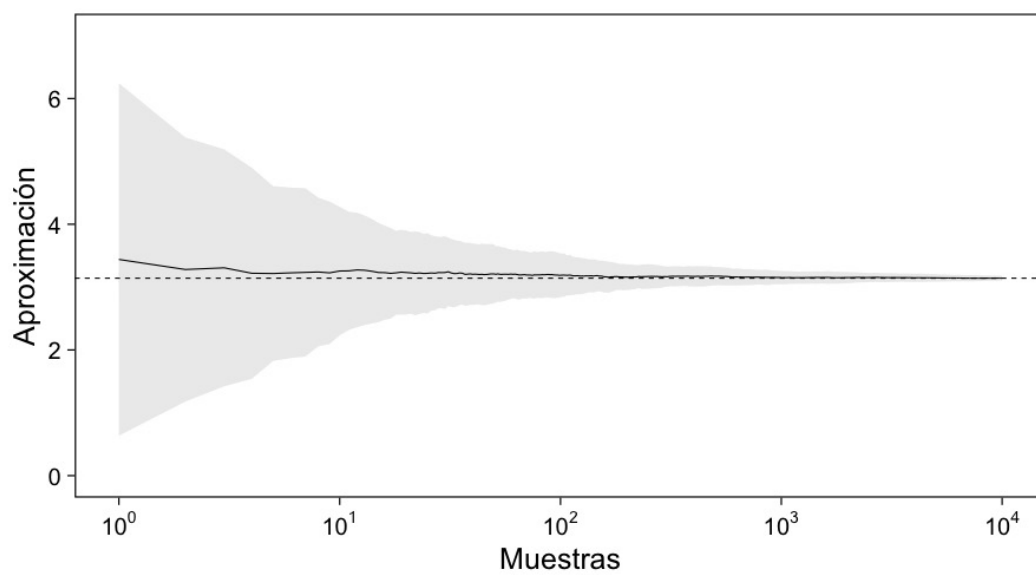


FIGURA 9. Comportamiento promedio e intervalos de confianza con aproximación asintótica.