

EST-24107: Simulación

Profesor: Alfredo Garbuno Iñigo — Otoño, 2023 — Introducción.

Objetivo: Estableceremos una idea general del alcance del curso. Esto nos llevará a encuadrar ciertos conceptos de simulación con el objetivo de cuantificar incertidumbre.

Lectura recomendada: El capítulo 1 de Sullivan [4] brinda una introducción al área de cuantificación de incertidumbre (UQ). Los capítulos 1–4 de Smith [3] tienen una bonita discusión (y un repaso) que sintoniza muy bien con lo expuesto en estas notas. El capítulo 1 de Hennig et al. [1] tiene una discusión interesante sobre el uso de simulación estocástica como proceso de cómputo numérico.

1. Motivación

Un curso típico de **simulación estocástica** hace énfasis en los *métodos de inferencia estadística* y *simulación de procesos estocásticos*. En contraste, este curso tomará énfasis en el uso de herramientas con el objetivo de **cuantificar incertidumbre**. Este marco de trabajo nos permitirá dejar a futuro aplicaciones de inferencia y procesos estocásticos.

Como modeladores matemáticos nos interesa **abstraer** y **representar** matemáticamente procesos o fenómenos.

Por ejemplo, nos interesa modelar el clima (procesos y características atmosféricas en el largo plazo) lo cual podemos simplificar por medio de ecuaciones diferenciales. Estas ecuaciones son un modelo matemático y la única manera que tenemos de resolverlas es por medio de métodos numéricos (ver Figura 1).

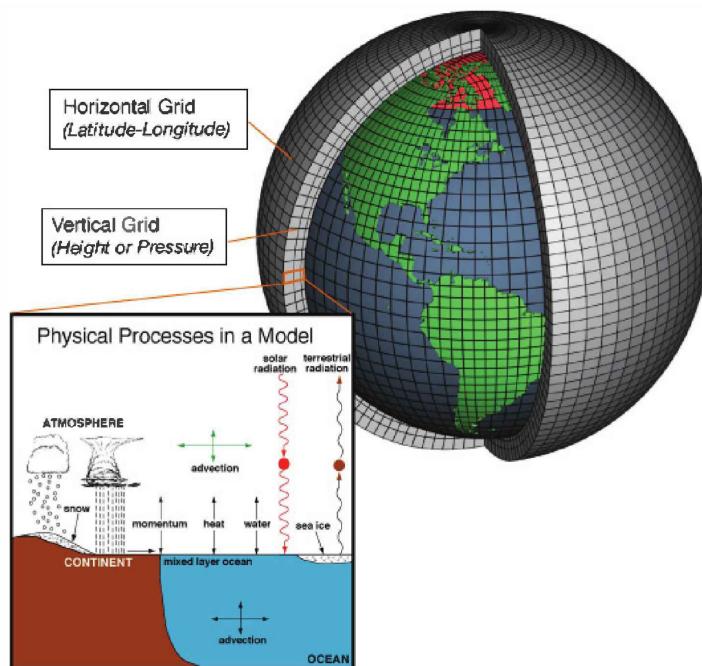


FIGURA 1. Imagen tomada de [3].

1 Motivación

Por supuesto la representación matemática tiene sus limitaciones. Además, los métodos numéricos que resuelven estos métodos son aproximaciones computacionales.

El modelado de procesos nos permite tomar decisiones informadas. Por ejemplo, el **modelado meteorológico** nos permite tomar informar sobre **eventos extremos** como huracanes (ver Figura 2).

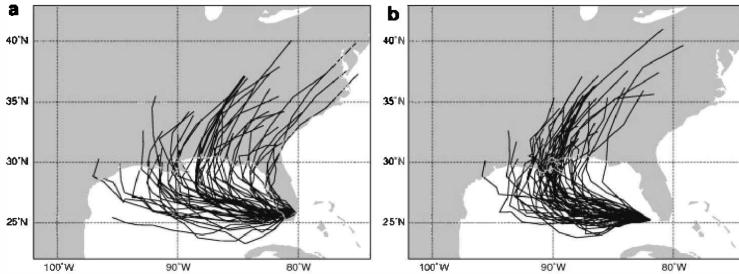


FIGURA 2. Imagen tomada de [3].

Por medio de herramientas computacionales podemos generar una colección de **situaciones ficticias** y con esto estudiar un conjunto de **escenarios hipotéticos** (ver Figura 2). Por ejemplo, nos interesa estudiar el **comportamiento promedio** y conocer la **dispersión** alrededor de este promedio.



FIGURA 3. Imagen tomada de [3].

Una gran cantidad de problemas de modelado se reduce en poder resolver **integrales de varias dimensiones**. Por ejemplo, en **generación de gráficos por computadora** se tiene que resolver la **ecuación de representación** bajo distintos escenarios (ver Figura 4).

Por último, en aplicaciones como finanzas, telecomunicaciones o seguros nos interesa estimar la probabilidad de ocurrencia de algún **evento muy raro** (probabilidad $\approx 10^{-3}$).

En resumen, los métodos que veremos en el curso serán su primera aproximación para resolver estos problemas con **énfasis en los algoritmos** y la **formulación basada en probabilidad**.

$$L_o(\mathbf{p}, \omega_0) = L_e(\mathbf{p}, \omega_0) + \int_{\mathcal{H}^2} f_r(\mathbf{p}, \omega_i \rightarrow \omega_0) L_i(\mathbf{p}, \omega_i) \cos \theta d\omega_i$$

FIGURA 4. Imagen tomada del material del curso *Gráficos por Computadora* impartido en Carnegie Mellon.

2. Formulación

Abordaremos el tipo de problemas descrito anteriormente con un poco de notación general. Denotaremos por z al resultado del fenómeno que nos interesa modelar. Nuestro conocimiento científico nos permite modelar dicho fenómeno a través de una abstracción la cual denotaremos por $f : X \rightarrow Z$ de tal forma que

$$z \approx f(x). \quad (1)$$

Esto nos permite hacer énfasis en que nuestro modelo matemático empata a la realidad hasta cierto nivel de precisión. Por otro lado, utilizamos el supuesto que nuestro modelo matemático recibe *entradas* específicas (x) que nos permiten *reproducir* los resultados con el modelo que hemos especificado.

Por ejemplo, consideremos el **movimiento de los planetas alrededor del sol** por medio de las ecuaciones de movimiento de Newton. En dicho sistema denotamos por $y \in \mathbb{R}^2$ el vector posición del planeta y por $p \in \mathbb{R}^2$ el vector de inercia. De tal forma que se satisface el sistema

$$\frac{dy}{dt} = \frac{p}{m}, \quad (2)$$

$$\frac{dp}{dt} = -\frac{k}{r^3}(y - y_\star), \quad (3)$$

donde m denota la masa del planeta; k , la constante de gravitación por masa planetaria y solar; y r la distancia entre el planeta y el sol. En este ejemplo, la función $f(x)$ es la solución del sistema de ecuaciones que nos da para cada x . En este caso consideremos x como el tiempo.

En nuestra caricatura anterior nos falta definir una colección de valores para que nuestro modelo represente, por ejemplo, la trayectoria de la Tierra alrededor del Sol. Para esto consideremos un **vector de parámetros** $\theta = (m, k, y_\star) \in \mathbb{R}^3$. Estos parámetros usualmente se ajustan de acuerdo a algún criterio o procedimiento de inferencia. En este sentido nuestro modelo lo podemos abstraer como

$$z \approx f(x; \theta), \quad (4)$$

donde $f : X \times \Theta \rightarrow Z$ y pensamos que existe una configuración de θ — que muchas veces le llamamos θ^* — que acerca proceso y que no podemos reducir por medio de un mejor modelo.

- Incertidumbre aleatoria:** también conocida como incertidumbre estadística, estocástica o irreducible. Se refiere a la incertidumbre que es natural para nuestro proceso y que no podemos reducir por medio de un mejor modelo.

3 Notación

2. **Incertidumbre epistémica:** se refiere a la incertidumbre derivada de nuestra simplificación del problema, nuestro estado de conocimiento o supuestos. En algunas ocasiones está asociada a los métodos numéricos con los que implementamos nuestros modelos. En otras, está asociada con los supuestos con lo que contamos para resolver un problema.

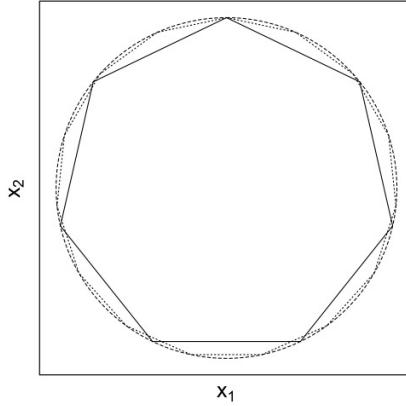


FIGURA 5. *Aproximación a un círculo mediante una trayectoria discretizada.*

Esta distinción nos ayuda a visualizar dos conceptos:

1. Identificar la necesidad de modelar incertidumbre en nuestros procesos.
2. Identificar el origen de dicha incertidumbre.

Lamentablemente en la práctica, al momento de generar simulaciones, nos olvidamos estas nociones y siempre es importante considerar las limitaciones de nuestros modelos para representar correctamente la realidad.

Ahora, la pregunta natural es ¿cómo modelamos la incertidumbre? En este curso (y en general en cualquier otras aplicaciones) utilizaremos el **lenguaje de probabilidad para expresar incertidumbre** ([2]). En este enfoque, es usual considerar incertidumbre aleatoria. Por otro lado, un curso como el de cálculo numérico nos permitirá cuantificar la incertidumbre epistémica. Sin embargo, también veremos en este curso que con herramientas probabilísticas podemos cuantificar ciertas nociones de incertidumbre de ambos tipos.

3. Notación

Denotamos por x una **variable aleatoria** y por $\mathbb{P}(\cdot)$ una **función de distribución**. Escribimos $x \sim \mathbb{P}$ para denotar que la variable aleatoria x tiene distribución $\mathbb{P}(\cdot)$. Denotamos por $\mathbb{E}[\cdot]$ el **valor esperado** del argumento con respecto a la distribución que estamos considerando. Durante el curso seremos explícitos en la variable aleatoria y usaremos

$$\mathbb{E}_x[\cdot] = \int_{\mathcal{X}} \cdot \pi(x) dx, \quad (5)$$

o bien, haremos énfasis en la distribución por medio de lo siguiente

$$\mathbb{E}_{\pi}[\cdot] = \int_{\mathcal{X}} \cdot \pi(x) dx, \quad (6)$$

de acuerdo al contexto.

Nota que en las ecuaciones anteriores estamos considerando el término $\pi(\cdot)$ como la **función de densidad de la función de probabilidad $\mathbb{P}(\cdot)$** .

Nos será útil la siguiente notación para evaluar valores esperados

$$\pi(f) := \mathbb{E}_\pi[f(x)] = \int_{\mathcal{X}} f(x) \pi(x) dx , \quad (7)$$

pues será el **objetivo general** para los métodos que estudiaremos en el curso.

Seremos cuidadosos cuando usemos $\pi(x)$ cuando hablamos de la función de densidad, o cuando usemos $\pi(f)$ cuando hablamos de la integral de $f(\cdot)$ ponderada por la densidad $\pi(\cdot)$. Nota que estamos utilizando la convención usual para denotar por las últimas letras del alfabeto— x, y, z por ejemplo—una variable aleatoria. Mientras que utilizamos las letras f, g ó h para denotar una función. El contexto indicará cual uso le daremos a $\pi(\cdot)$; ya sea como densidad o como un funcional.

Por ejemplo, utilizaremos la noción de **aproximar integrales** por medio de algún procedimiento de muestreo de tal forma que esperaremos construir un estimación $\hat{\pi}(f)$ con cierto grado de refinamiento. Por ejemplo, veremos el **método Monte Carlo** que utiliza una colección de N simulaciones para aproximar la integral anterior. Esto lo denotaremos por

$$\hat{\pi}_N^{\text{MC}}(f) \approx \pi(f) . \quad (8)$$

En general, nos interesa, y esperamos, que podamos:

1. Mejorar nuestra estimación con mas muestras (simulaciones)

$$\lim_{N \rightarrow \infty} \hat{\pi}_N^{\text{MC}}(f) = \pi(f) \quad (9)$$

2. Cuantificar la incertidumbre en nuestra aproximación por medio de alguna distribución de probabilidad. Por ejemplo,

$$\hat{\pi}_N^{\text{MC}}(f) \sim N \left(\pi(f), \frac{\mathbb{V}(f)}{N} \right) . \quad (10)$$

4. Repaso de probabilidad :

Consideraremos como requisitos el contenido de **Cálculo de Probabilidades II y Álgebra Lineal** (o equivalentes). En particular lo que requerimos como base es lo siguiente.

Definición 4.1 (**Espacio de probabilidad**). Un espacio de probabilidad está definido por la terna $(\Omega, \mathcal{X}, \mathbb{P})$:

1. El espacio muestral, Ω (elementos).
2. El espacio de eventos medibles, \mathcal{X} (subconjuntos).
3. La medida de probabilidad, $\mathbb{P} : \mathcal{X} \rightarrow [0, 1]$.

Definición 4.2 (**Variable aleatoria**). Una variable aleatoria es una función $X : \mathcal{X} \rightarrow \mathbb{R}$ con la propiedad de que las pre-imágenes bajo X son eventos medibles. Es decir,

$$\{w \in \mathcal{X} : X(w) \leq x\} \in \mathcal{X} \quad \forall x \in \mathbb{R}. \quad (11)$$

Definición 4.3 (Función de acumulación de probabilidad). Para toda variable aleatoria X tenemos una función de acumulación $\mathbb{P}_X : \mathbb{R} \rightarrow [0, 1]$ dada por

$$\mathbb{P}_X(x) = \mathbb{P}(\{w \in \mathcal{X} : X(w) \leq x\}). \quad (12)$$

Esto usualmente lo escribimos como $\mathbb{P}_X(x) = \mathbb{P}\{X \leq x\}$.

Definición 4.4 (Función de densidad). Una variable aleatoria es continua si su función de acumulación es **absolutamente continua** y puede ser expresada por medio de

$$\mathbb{P}_X(x) = \int_{-\infty}^x \pi(s) ds, \quad (13)$$

donde la anti-derivada $\pi : \mathbb{R} \rightarrow [0, \infty)$ se llama la **función de densidad** de la variable aleatoria X .

Las propiedades generales de las distribuciones de probabilidad se pueden especificar por medio de su centralidad (localización), su dispersión, su rango de valores, su simetría y el comportamiento de valores extremos.

En general esto lo podemos extraer de los momentos

$$\mathbb{E}(X^p) = \int_{\mathbb{R}} x^p \pi(x) dx, \quad (14)$$

o los momentos centrales. Por ejemplo: media y varianza.

Uno de los resultados que espero recuerden bien de sus cursos anteriores es el de la **Ley de los Grandes Números**. La cual podemos enunciar como:

Teorema 4.5 (Ley de los grandes números). Sea X_1, X_2, \dots una colección de variables aleatorias independientes e idénticamente distribuidas (**iid**) y sea \bar{X}_n el promedio de un subconjunto de n . Si denotamos por μ el valor promedio de X_i dentro de esa colección, entonces tenemos que

$$\bar{X}_n \rightarrow \mu \quad (\text{casi seguramente}). \quad (15)$$

Teorema 4.6 (Límite Central). Sea X_1, \dots, X_n una colección de n variables aleatorias **iid** con $\mathbb{E}[X_i] = \mu$ y $\mathbb{V}[X_i] = \sigma^2 < \infty$. Entonces tenemos que el comportamiento del promedio sigue la siguiente distribución

$$\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad (16)$$

para n suficientemente grande.

5. Conclusiones

En este curso estudiaremos las aproximaciones numéricas a problemas de integración y optimización utilizando herramientas estocásticas (basadas en variables aleatorias).

Cuantificar la incertidumbre de nuestras aproximaciones numéricas nos permite utilizar variables aleatorias para caracterizar dicha incertidumbre. Esto nos permite propagar incertidumbre a lo largo de todo un proceso de toma de decisiones.

Las ventajas de esta metodología (simulación estocástica) son:

1. utilizar probabilidad para caracterizar escenarios posibles;
2. utilizar nuestras computadoras para realizar dichas simulaciones;
3. determinar hasta qué momento estamos satisfechos con nuestra aproximación.

Referencias

- [1] P. Hennig, M. A. Osborne, and H. P. Kersting. *Probabilistic Numerics: Computation as Machine Learning*. Cambridge University Press, 2022. [1](#)
- [2] E. Jaynes and G. Bretthorst. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003. [4](#)
- [3] R. C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*, volume 12. SIAM, 2013. [1](#), [2](#)
- [4] T. Sullivan. *Introduction to Uncertainty Quantification*, volume 63 of *Texts in Applied Mathematics*. Springer International Publishing, 2015. [1](#)