Simulación Estocástica 2025-1

Taller 1: Probabilidad, Variables aleatorias, Generadores aleatorios, Simulación de Monte Carlo

Javier Riascos Ochoa, PhD

Indicaciones: Por grupo, subir al link en AVATA **un (1)** archivo Rmarkdown o Notebook de Python con el procedimiento analítico, códigos (de los puntos que lo requieran), resultados y análisis, y **un (1)** archivo pdf con la salida impresa.

Fecha de entrega: Domingo 9 de Marzo de 2025

1. **Events and probability (1 point)**(*Analytic exercise*)

   a) A box contains three marbles: one red, one green, and one blue. Consider an experiment that consists of taking one marble from the box then replacing it in the box and drawing a second marble from the box. What is the sample space? If, at all times, each marble in the box is equally likely to be selected, what is the probability of each point in the sample space?

   b) Repeat the previous exercise when the second marble is drawn without replacing the first marble.

2. **Congruential generators (1 point)**

   Find all of the cycles of the following congruential generator. For each cycle identify which seeds X_0 lead to that cycle:

   $$X_n = (9X_{n-1} + 3)\ mod\ 11.$$

3. **Uniformity and independence of the uniform generator (1 point)**

   Read section "2.1.1. Uniform simulation" from *Introducing Monte Carlo Methods with R* (Robert and Casella). Explain briefly each line of the code to produce Figure 2.1, the functions `runif, par, hist, acf`, and how the output allows to conclude the uniformity and independence of the uniform generator `runif`.

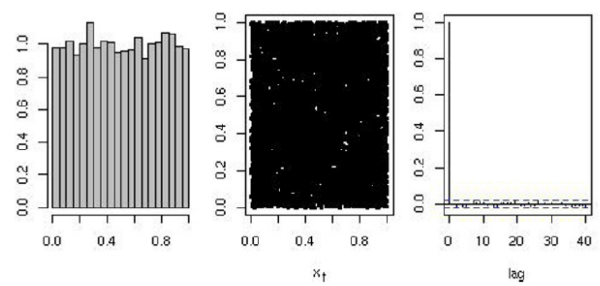


Fig. 2.1. Histogram *(left)*, pairwise plot *(center)*, and estimated autocorrelation function *(right)* of a sequence of 10^4 uniform random numbers generated by `runif`.

4. **Inverse method for a discrete r.v. (2 points)**

   Consider the discrete random variable (r.v.) X with probability mass function (*pmf*) given by:

   $$P(X = -1) = 0.2, \quad P(X = 0) = 0.5, \quad P(X = 1) = 0.3$$

   a) Calculate and plot the Cumulative Distribution Function (*CDF*) $F_X(x)$ of X.

   b) Calculate the expectation $E[X]$ and variance $Var(X)$ of X.

   c) Write a program to generate n values of this random variable using the inverse method i.e. by generating random numbers uniformly distributed in $(0, 1)$.

   d) Let $n = 100$, run the program, and determine: (i) the aritmetic mean $\overline{X}$ of the simulated values and (ii) the sample standard deviation S of these values. Compare these statistics with the quantities:

$E[X]$, and $\sqrt{Var(X)}$, respectively, in terms of the:

$$\textit{Absolute percentage error} = \frac{|\textit{Exact value} - \textit{Estimated value}|}{\textit{Exact value}} \cdot 100\,\%.$$

e) Repeat *d*) with $n = 1000$.

f) Repeat *d*) with $n = 10000$.

5. **Inverse method for a continuous r.v. (2 points)**

Consider the continuous random variable X with probability density function (*pdf*) given by:

$$f_X(x) = \begin{cases} 3(x-1)^2 & \text{for} \quad 1 < x \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

a) Find the following probabilities: (i) $P(X \leq 1)$, (ii) $P(1 < X \leq 1.5)$, (iii) $P(X \geq 1.5)$.

b) Calculate the expectation $E[X]$ and variance $Var(X)$ of X.

c) Find the *CDF* $F_X(x)$ of X. Plot this function.

d) Show how to simulate X with this *CDF* using the inversion method.

e) Write a program in `R` that draws 1000 samples of X. Plot a normalized histogram of the sample along with the PDF of X.

6. **Monte Carlo Integration (2 points)**

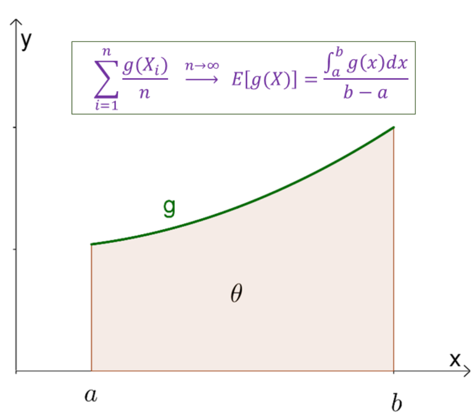
With respect to the following integrals: (i) Find their exact value analytically or with software (e.g., `WolframAlpha`), (ii) with Monte Carlo integration approximate the integrals and compare with the exact answer.

a) $\int_0^1 \exp\{e^x\}dx$

b) $\int_{-2}^2 e^{x+x^2}dx$

c) $\int_0^\infty x(1+x^2)^{-2}dx$

d) $\int_0^1 \int_0^2 e^{(x+y)^2}dydx$



7. **Estimating π (3 points)**

Suppose that X and Y are *iid* $U(-0.5, 0.5)$ random variables.

a) What is $P((X,Y) \in [a,b][c,d])$ for $-0.5 \leq a \leq b \leq 0.5$ and $-0.5 \leq c \leq d \leq 0.5$?

b) Based on your previous answer, what do you think you should get for $P((X,Y) \in A)$, where A is an arbitrary subset of $[-0.5, 0.5][-0.5, 0.5]$?

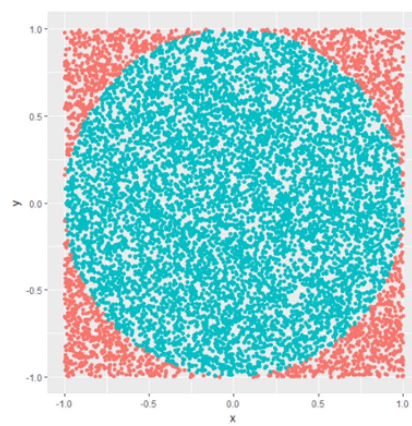
c) Let $A = \{(x,y) \in [-0.5, 0.5] \times [-0.5, 0.5] : x^2 + y^2 \leq 0.5^2\}$. What is the area of A?

d) Define the *r.v.* Z by:

$$Z = \begin{cases} 1 & \text{if} \quad X^2 + Y^2 \leq 0.5^2, \\ 0 & \text{otherwise.} \end{cases}$$

What is $E[Z]$?

e) Write a program to estimate π by simulating Z with n Monte Carlo independent simulations.



8. **Estimating expected values with Monte Carlo (3 points)**

For uniform $(0,1)$ random variables $U_1, U_2, \ldots$ define:

$$N = \text{mín}\left\{n : \sum_{i=1}^{n} U_i > 1\right\}$$

That is, N is equal to the number of random numbers that must be summed to exceed 1.

a) Estimate $E[N]$ by generating 100 values of N.

b) Estimate $E[N]$ by generating 1000 values of N.

c) Estimate $E[N]$ by generating 10000 values of N.

d) What do you think is the value of $E[N]$?