TRABAJO PRÁCTICO 1.1 - SIMULACIÓN DE UNA RULETA

Danteo Elías
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
elias.danteo.tomas@hotmail.com

Cosentino Lucio N.
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
luciocosen@gmail.com

De Bernardo, Aarón
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
aarondebernardo@gmail.com

Fernandez Da Silva, Joaquín C.
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
joaquinfds13@gmail.com

Malizani Juan Pablo
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
juampi123.m@gmail.com

Pastorino Juan Jose
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
juanjosepastorino@gmail.com

April 21, 2025

ABSTRACT

El presente trabajo tiene como objetivo desarrollar un simulador de ruleta europea (con un solo cero) que permita analizar el comportamiento estadístico del juego. Mediante técnicas de simulación computacional, se estudian variables como la frecuencia relativa de aparición de un número, el valor promedio de las tiradas, la varianza y el desvío estándar. El simulador permite configurar múltiples corridas con diferente cantidad de tiradas y un número específico a analizar. Los resultados se comparan con los valores teóricos esperados, verificando las propiedades estadísticas de la ruleta como un generador de números aleatorios con distribución uniforme discreta.

***K*eywords** Simulación · Ruleta · Python · Estadística

1 Introducción

La ruleta es uno de los juegos de azar más emblemáticos de los casinos, cuya versión moderna fue desarrollada en el siglo XVIII. Según la historia, Blaise Pascal, mientras experimentaba con dispositivos de movimiento perpetuo, creó un prototipo temprano de lo que luego se convertiría en la ruleta moderna [1].

La ruleta europea, objeto de este estudio, consta de 37 números (0-36) con una distribución específica en el plato que busca equilibrar números pares e impares, altos y bajos, y colores rojos y negros. Desde el punto de vista probabilístico, cada número tiene la misma probabilidad de salir en una tirada ($^1/_{37}$), lo que la convierte en un práctico ejemplo de distribución uniforme discreta.

El objetivo de este trabajo es desarrollar un simulador computacional que permita:

- Generar tiradas aleatorias de ruleta siguiendo una distribución uniforme discreta
- Calcular estadísticas clave:
  - frecuencia relativa,
  - promedio,
  - varianza,

  - y desvío estándar.
- Comparar los resultados muestrales empíricos con los valores teóricos esperados
- Visualizar la convergencia de los estadísticos hacia los valores poblacionales

2 Marco Teórico

2.1 Método de Montecarlo

La esencia de este método radica en la simulación de eventos aleatorios o el muestreo de un espacio de posibilidades, seguido de la evaluación estadística de los resultados obtenidos. A medida que se aumenta el número de experimentos o muestras, las estimaciones obtenidas tienden a converger hacia los valores teóricos esperados.

El método de Montecarlo aplicado a la simulación de la ruleta puede proporcionar información valiosa como las probabilidades de ganancia, la distribución de resultados, la efectividad de las estrategias de apuestas y la validación de modelos teóricos.

2.2 Conceptos Básicos de Estadística

La estadística estudia la variabilidad sobre cuestiones de la realidad en los que interviene el azar. Gracias a la estadística se consiguen una serie de datos objetivables mediante los cuales se pueden extraer una serie de conclusiones.

La estadística son los métodos y procedimientos para recoger, clasificar, analizar y representar los datos, así como obtener conclusiones a través de ellos, con la intención de formular predicciones y ayudar en la toma de decisiones.

Existen dos tipos principales de estadística:

- **Estadística descriptiva:** es la parte de la estadística que se ocupa de ordenar, sintetizar y representar gráficamente los resultados recogidos durante la investigación. La estadística descriptiva no solo describe, sino también analiza y representa los datos utilizando elementos numéricos y gráficos.
- **Estadística inferencial:** es la estadística que tiene como objetivo obtener conclusiones sobre el total de la población a partir de los datos obtenidos en un subconjunto de la misma o grupo de elementos representativos (muestra).

Los principales conceptos en estadística utilizados son:

- **Población:** también conocido como universo o conjunto completo de individuos que cumplen una serie de características y al que harán referencia las conclusiones del estudio. A partir de la población de estudio se elegirá una muestra representativa.
- **Muestra:** es un grupo acotado o reducido de todos los individuos de forman la población. Se considera que una muestra es representativa cuando los individuos de la misma son seleccionados al azar.
- **Individuo:** son las personas o elementos que contienen la información del fenómeno que se pretende estudiar.
- **Muestreo:** es el procedimiento mediante el cual se obtiene una muestra. El muestreo puede ser probabilístico o aleatorio y no probabilístico o no aleatorio.
- **Aleatoriedad de una muestra:** es la característica mediante la cual todos los miembros de una muestra tienen las mismas posibilidades de formar parte de la misma.
- **Variables:** son las características de la población que se representan en los individuos que forman la muestra y que son susceptibles de ser medidas. Las variables pueden ser cuantitativa o cualitativas.
  - **Variables aleatorias:** se define al asignar un valor numérico a cada suceso elemental de una experiencia aleatoria. Es decir, una variable aleatoria numérica es un fenómeno de interés cuyos resultados se expresan con números.
- **Parámetro:** es un índice que resume una determinada característica de la población, representándose por las letras griegas μ y ρ. Un parámetro es la función definida sobre los valores numéricos de características medibles de una población
- **Estadístico:** es un índice que resume una determinada característica de la muestra, representándose por las letras del alfabeto latino “x” y “s”. Un estadístico es la función definida sobre los valores numéricos de una muestra.

- **Distribución de Probabilidad de una Variable Aleatoria:** es una función que asigna a cada suceso definido sobre la variable, la probabilidad de que dicho suceso ocurra. La distribución de probabilidad está definida sobre el conjunto de todos los sucesos y cada uno de los sucesos es el rango de valores de la variable aleatoria.

2.3 Distribución Uniforme Discreta

La **Distribución Uniforme Discreta** es una distribución de probabilidad en la que todos los valores posibles tienen la misma probabilidad de ocurrir. Es decir, cada resultado tiene la misma "chance" de salir dentro de un conjunto finito de valores.

Definición Formal: Sea X una variable aleatoria discreta que puede tomar n valores distintos $x_1, x_2, \ldots, x_n$. Se dice que X tiene una **distribución uniforme discreta** si:

$$P(X = x_i) = \frac{1}{n} \quad \text{para todo } i = 1, 2, \ldots, n$$

Características:

- **Esperanza (media):**
$$\mathbb{E}(X) = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{1}$$
- **Varianza** (cuando los valores son consecutivos de 1 a n):
$$\text{Var}(X) = \frac{n^2 - 1}{12} \tag{2}$$

Para una ruleta europea con números del 0 al 36, la función de probabilidad es:

$$P(X = k) = \frac{1}{37}, \quad \text{para } k = 0, 1, 2, \ldots, 36 \tag{3}$$

Los principales valores teóricos esperados son:

- **Valor esperado (media):**
$$E[X] = \frac{\sum_{k=0}^{36} k}{37} = \frac{0 + 1 + \cdots + 36}{37} = 18 \tag{4}$$
- **Varianza:**
$$Var(X) = \frac{(n^2 - 1)}{12} = \frac{(36^2 - 1)}{12} = 114 \tag{5}$$
donde $n = 36$ es el rango de valores (0 a 36).
- **Desviación estándar:**
$$\sigma = \sqrt{Var(X)} = \sqrt{114} \approx 10.677 \tag{6}$$
- **Frecuencia relativa esperada:** Para cualquier número específico:
$$fr = \frac{1}{37} \approx 0.027 \tag{7}$$

2.4 Ley de los Grandes Números

A medida que aumenta el número de tiradas ($n \to \infty$), las frecuencias relativas observadas deben converger a las probabilidades teóricas:

$$\lim_{n \to \infty} \frac{\text{Número de veces que sale } k}{n} = \frac{1}{37} \tag{8}$$

2.5 Inferencia Estadística de el numero de Tiradas Necesarias

Desarrollo teórico

Supongamos que se desea estimar la media poblacional μ de una variable aleatoria X con un intervalo de confianza del $100(1-\alpha)\%$ y un margen de error máximo tolerable E.

El estimador de la media es la media muestral $\bar{X}$, y bajo el teorema central del límite, si X tiene varianza finita, entonces:

$$\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \tag{9}$$

donde n es el tamaño total de la muestra y σ^2 la varianza poblacional.

El intervalo de confianza para la media poblacional está dado por:

$$\bar{X} \pm Z\alpha/2 \cdot \frac{\sigma}{\sqrt{n}} \tag{10}$$

donde $Z\alpha/2$ es el valor crítico de la distribución normal estándar asociado al nivel de confianza deseado.

Para garantizar que el error no supere E, imponemos:

$$Z\alpha/2 \cdot \frac{\sigma}{\sqrt{n}} \leq E \tag{11}$$

Despejando n:

$$\frac{\sigma}{\sqrt{n}} \leq \frac{E}{Z_{\alpha/2}} \Rightarrow \sqrt{n} \geq \frac{\sigma Z_{\alpha/2}}{E} \Rightarrow n \geq \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2 \tag{12}$$

2.6 Cálculo de la cantidad de tiradas necesarias

Dada una ruleta con una distribución uniforme $U(0,36)$, deseamos calcular la cantidad de tiradas necesarias para estimar la media con un intervalo de confianza $IC = 95\%$ y un margen de error absoluto $\varepsilon = 1 unidad$.

Paso 1: Cálculo de la varianza La ruleta sigue una distribución uniforme $U(0,36)$, por lo que la varianza de esta distribución es:

$$\sigma^2 = \frac{(b-a)^2}{12} \tag{13}$$

donde $a = 0$ y $b = 36$. Sustituyendo estos valores:

$$\sigma^2 = \frac{(36-0)^2}{12} = \frac{36^2}{12} = 108 \tag{14}$$

Entonces, la desviación estándar es:

$$\sigma = \sqrt{108} \approx 10.39 \tag{15}$$

Paso 2: Fórmula para el número de muestras necesarias La fórmula para calcular el número de muestras necesarias es:

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2 \tag{16}$$

Donde:

$Z_{\alpha/2} = 1.96$ (valor crítico para un intervalo de confianza del 95%),

$\sigma = 10.39$ (desviación estándar de la distribución),

$E = 1$ (margen de error).

Sustituyendo estos valores en la fórmula:

$$n = \left(\frac{1.96 \cdot 10.39}{1}\right)^2 = (1.96 \cdot 10.39)^2 \tag{17}$$

$$n = (20.37)^2 \approx 414.9 \tag{18}$$

Por lo tanto, se necesitarían aproximadamente 415 tiradas para estimar la media con un intervalo de confianza del 95

3 Implementación del Simulador

Esta investigación consiste en simular el funcionamiento del plato de una ruleta en Python y evaluar los valores obtenidos, implementando los conceptos teóricos mencionados en la sección anterior.

El experimento consiste en generar valores aleatorios que representan los números posibles en la ruleta, es decir, los enteros del 0 al 36. Cada valor generado corresponderá a una corrida y en cada corrida se efectuará un número fijo de tiradas.

Hacemos esto de esta manera para circunvenir una debilidad de los sistemas aleatorios computarizados, que es el sesgo debido a que las computadoras, por su propia naturaleza, no son aleatorias: son *deterministas*. Es decir, dado un mismo conjunto de instrucciones y condiciones, siempre producen el mismo resultado.

Lo que hacen es simular aleatoriedad usando un algoritmo llamado generador de números pseudoaleatorios (PRNG, por sus siglas en inglés) partiendo de un valor inicial llamado semilla (seed). Al utilizar varias corridas, obtendremos varias semillas lo que ayudara a evitar la pseudoaleatoriedad.

El objetivo es analizar la muestra resultante de estas corridas con un número seleccionado, en la cual aplicaremos los conceptos de nuestro marco teórico y generaremos las respectivas gráficas.

El simulador fue desarrollado en Python 3.x utilizando las siguientes bibliotecas principales:

- `random` para la generación de números aleatorios
- `matplotlib` para la visualización de resultados
- `numpy` para cálculos estadísticos
- `argparse` para leer los parámetros al ejecutar el programa

Para poder probar el simulador, se deja aquí un link a un repositorio de Github: `https://github.com/Luciocos/SimulacionTPs/tree/main`

3.1 Estructura del Código

El programa tiene la siguiente estructura:

1. Configuración de parámetros mediante argumentos de línea de comandos:
   - Número de corridas (`-c`) - por defecto: 10
   - Tiradas por corrida (`-t`) - por defecto: 415
   - Número a analizar (`-n`) - por defecto: 7
2. Generación de tiradas para una corrida aleatorias mediante `random.randint(0, 36)`
3. Cálculo de estadísticas para cada corrida:
   - Frecuencia relativa del número elegido
   - Valor promedio de las tiradas
   - Varianza y desvío estándar
4. Calculo de valores medios entre todas corridas
5. Visualización de resultados con comparación contra valores esperados

Nota: Los pasos 2 y 3 se repiten para todas las corridas.

3.2 Fórmulas Implementadas

Las principales fórmulas estadísticas implementadas son:

- **Frecuencia relativa** del número elegido k después de n tiradas:

$$fr(k) = \frac{\text{Conteo de } n \text{ en } t \text{ tiradas}}{t} \tag{19}$$

- **Promedio muestral**:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{20}$$

- **Varianza muestral**:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{21}$$

- **Desviación estándar muestral**:

$$s = \sqrt{s^2} \tag{22}$$

4 Resultados y Análisis

4.1 Configuración de las Simulaciones

Se realizo la simulación con los siguientes parámetros:

- Número de corridas: 10
- Tiradas por corrida: 415
- Número analizado: 7

4.2 Gráficos de Convergencia

A continuación se presentan los gráficos obtenidos para una de las corridas:

Los siguientes gráficos corresponden a la ejecución número siete del experimento. Esta se escogió de manera aleatoria para así evitar sesgo a la hora de la elección.

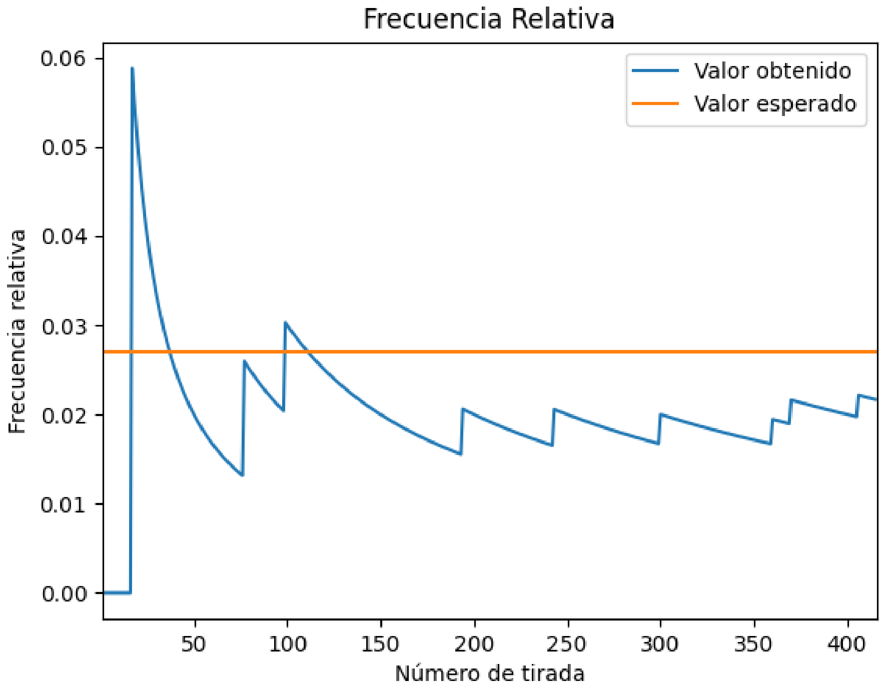

Figure 1: Convergencia de la frecuencia relativa del número 7 hacia el valor teórico (línea naranja) en la corrida siete

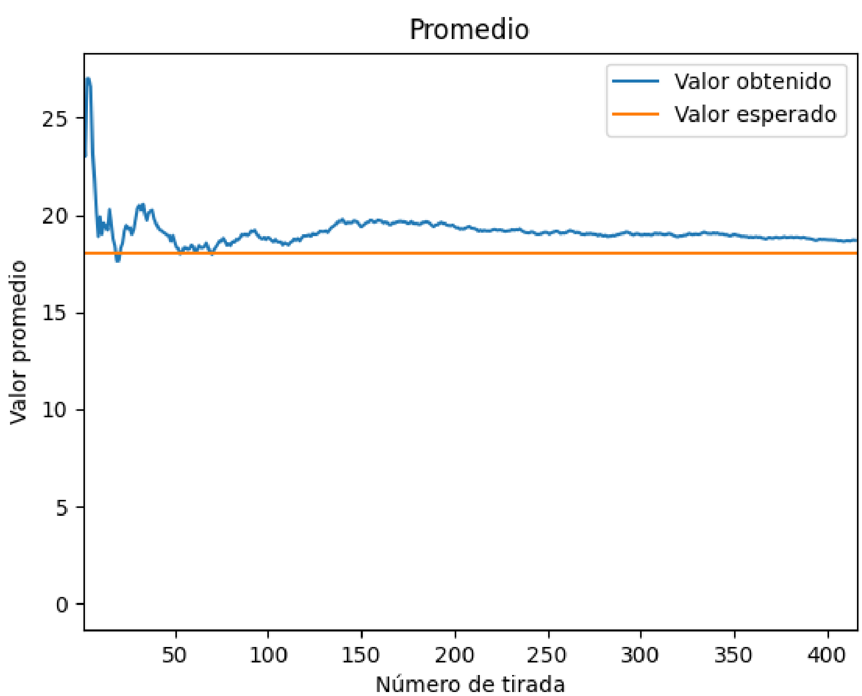

Figure 2: Convergencia del valor promedio de las tiradas hacia el valor teórico (18) en la corrida siete

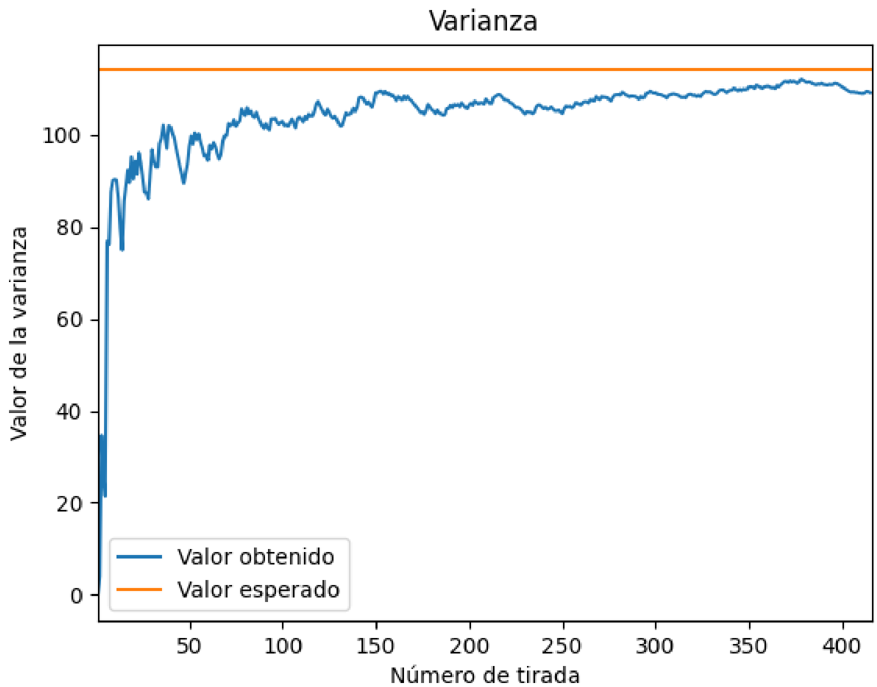

Figure 3: Convergencia de la varianza muestral hacia el valor teórico (114) en la corrida siete

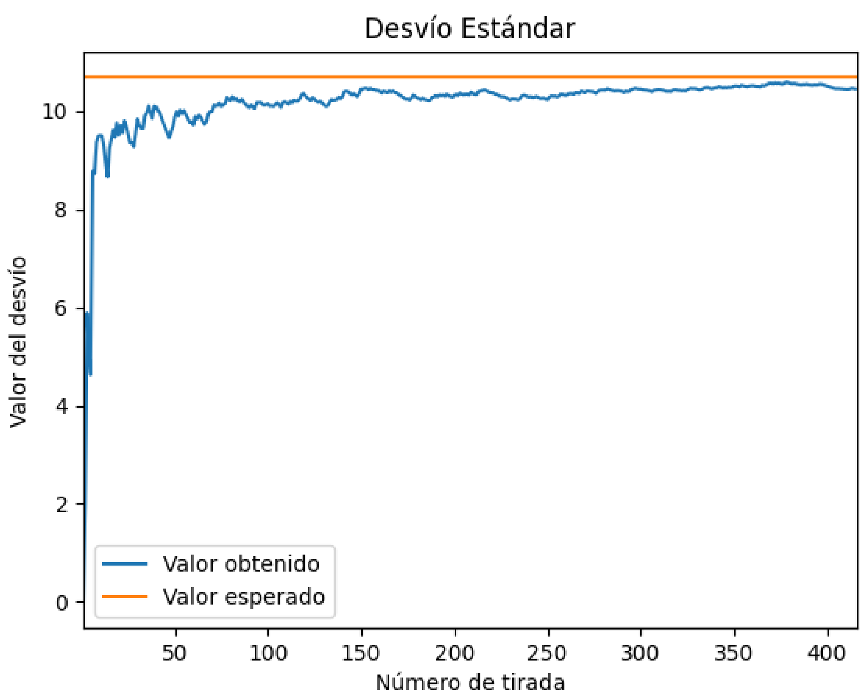

Figure 4: Convergencia del desvío estándar muestral hacia el valor teórico (10.677) en la corrida siete

Luego se presentan los gráficos obtenidos para el valor medio de todas las corridas:

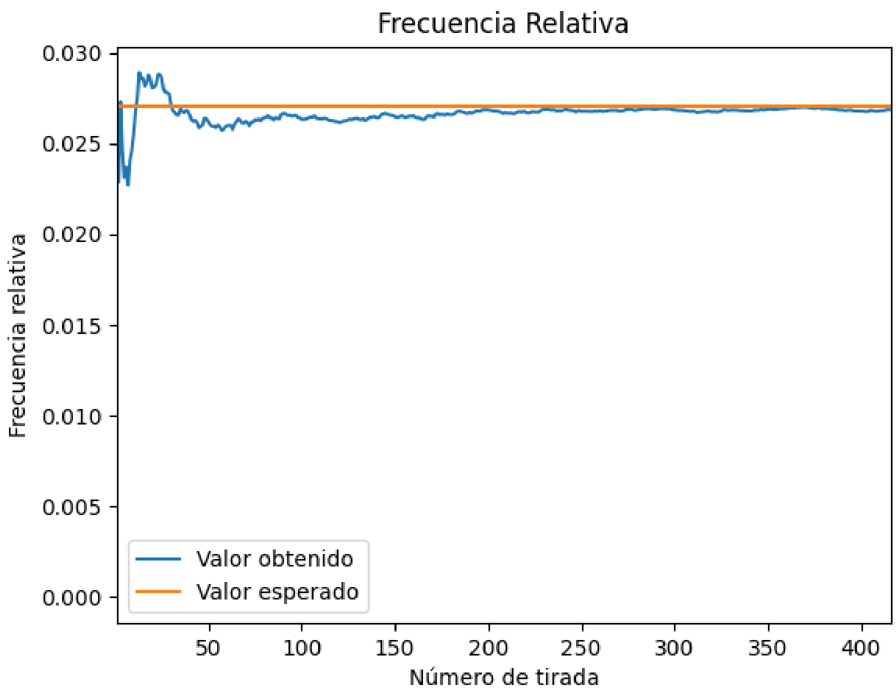

Figure 5: Convergencia de la frecuencia relativa del número 7 hacia el valor teórico (línea roja) para todas las corridas

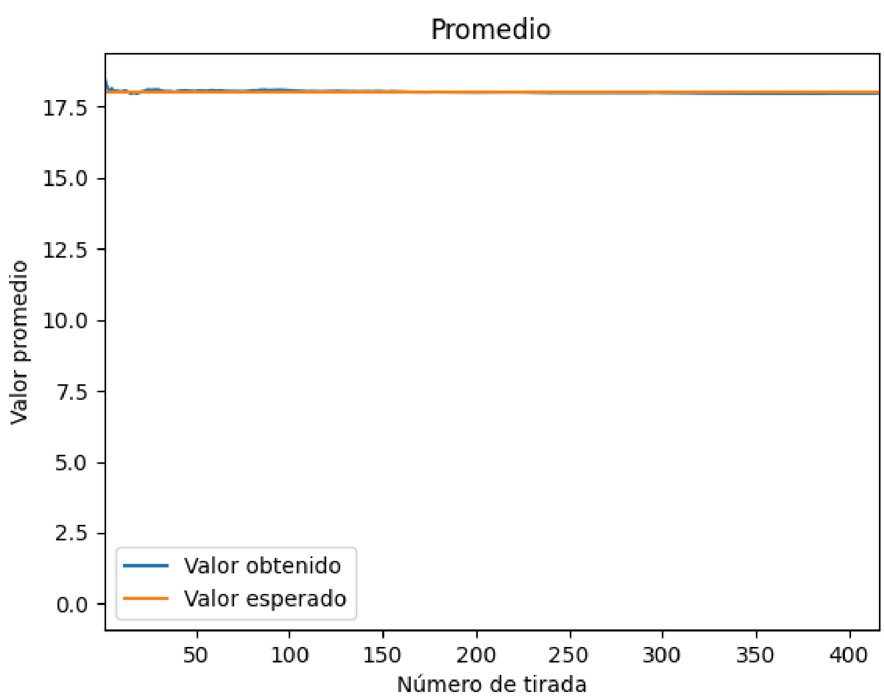

Figure 6: Convergencia del valor promedio de las tiradas hacia el valor teórico (18) para todas las corridas

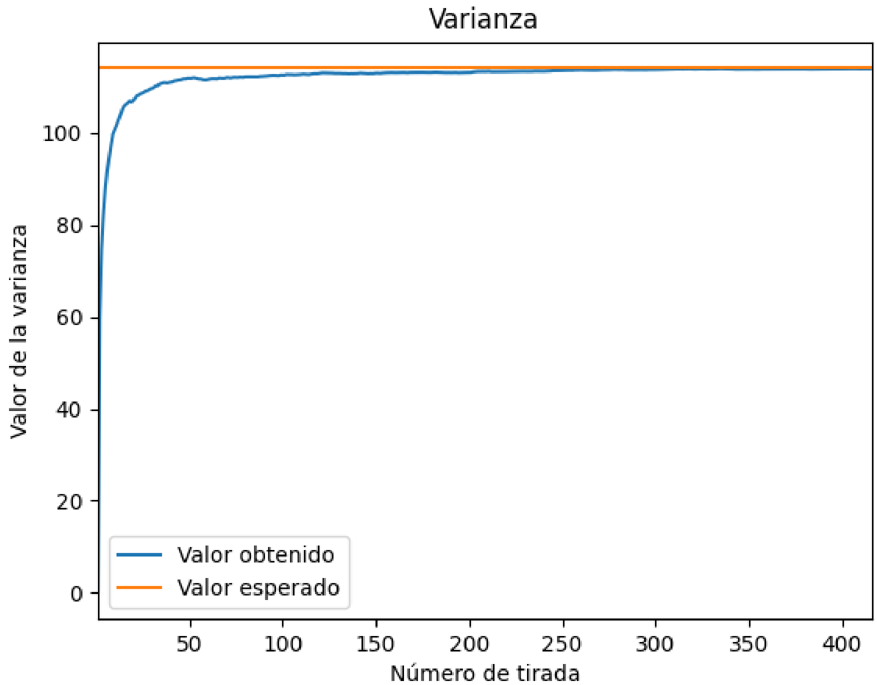

Figure 7: Convergencia de la varianza muestral hacia el valor teórico (114) para todas las corridas

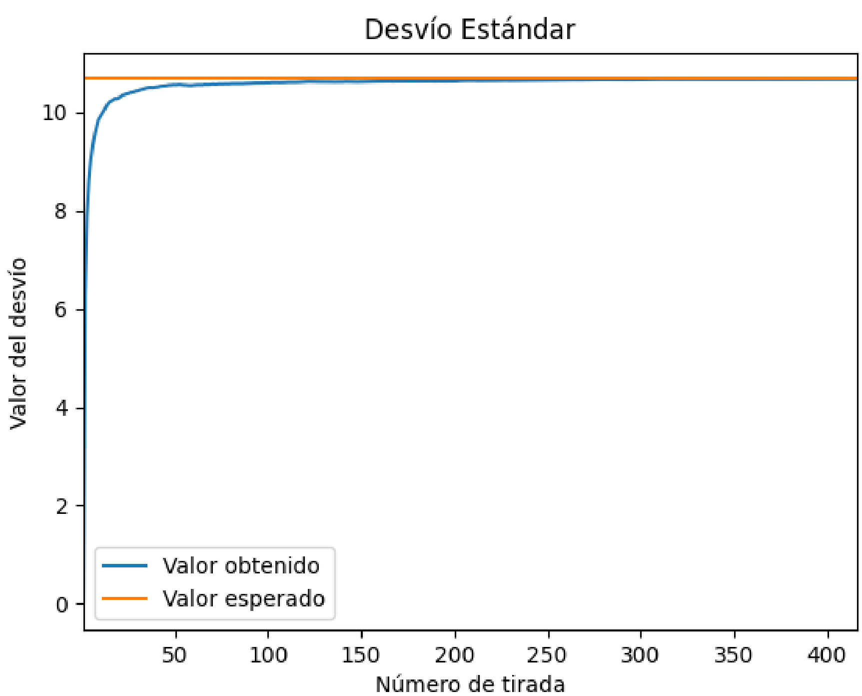

Figure 8: Convergencia del desvío estándar muestral hacia el valor teórico (10.677) para todas las corridas

4.3 Análisis de Resultados

Los gráficos muestran claramente cómo las estadísticas muestrales convergen hacia los valores teóricos a medida que aumenta el número de tiradas. Al alcanzar el valor calculado de 415 tiradas, se obtiene un nivel de confianza del 95% de que el error absoluto no supera una unidad. Esto es validado por nuestro simulador y confirma el comportamiento esperado bajo el modelo de una distribución uniforme discreta, como es el caso de una ruleta equilibrada.

- La frecuencia relativa del número 7 (Figura 5) oscila inicialmente pero se estabiliza alrededor de $1/37$ (0.027)
- El promedio muestral (Figura 6) converge rápidamente al valor teórico de 18
- La varianza (Figura 7) y desvío estándar (Figura 8) muestran mayor variabilidad inicial pero convergen a los valores esperados

Estas conclusiones son observando los gráficos para el valor medio de las tiradas, aun así. Si analizamos la corrida siete, podemos observar igual esta convergencia, aunque de manera mas oscilatoria.

5 Conclusiones

El simulador desarrollado permite estudiar empíricamente las propiedades estadísticas de la ruleta europea, evitando lo máximo posible el sesgo de la pseudoaleatoriedad de los lenguajes de programación. Los resultados obtenidos confirman:

- La ruleta sigue efectivamente una distribución de probabilidad uniforme discreta, en donde cada número tiene la misma probabilidad de ocurrencia.
- A medida que aumenta el número de tirada y se aproxima a la número 415, se ve que los estadísticos analizados convergen a los valores teóricos calculados.

Este trabajo sienta las bases para futuras investigaciones sobre simulaciones de juegos de azar y análisis de sistemas probabilísticos discretos.

Agradecimientos

Agradecemos al los profesores Juan I. Torres y a Jorge A. Flamini por la guía proporcionada y a la Universidad Tecnológica Nacional - FRRO por el apoyo institucional.

References

[1] Wikipedia. *Ruleta*. Disponible en: `https://es.wikipedia.org/wiki/Ruleta`

[2] Python Software Foundation. *Documentación de Python 3*. Disponible en: `https://docs.python.org/3/`

[3] Hunter, J. D. (2007). *Matplotlib: A 2D graphics environment*. Computing in Science & Engineering, 9(3), 90-95.

[4] Ross, S. M. (2014). *Introduction to Probability and Statistics for Engineers and Scientists*. Academic Press.

[5] Wikipedia. *Law of large numbers*. Disponible en: `https://en.wikipedia.org/wiki/Law_of_large_numbers`