TP 1.1 - Simulación de una Ruleta

Maria Paz Battistoni
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
battistonimpaz@gmail.com

Santino Cataldi
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
cataldisantinonanu@gmail.com

Marcos Godoy
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
mjgodoy2002@gmail.com

Matias Luhmann
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
luhmannm0@gmail.com

Matias Marquez
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
matiastmarquez@gmail.com

Tomás Wardoloff
Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
tomaswardoloff@gmail.com

22 de abril de 2025

Abstract

El presente documento tiene como objetivo describir el trabajo de investigación realizado como introducción a la cátedra Simulación de la carrera Ingeniería en Sistemas de Información. El mismo consiste en el desarrollo y análisis de un modelo simple de ruleta de casino, tomando como referencia la ruleta europea (con un único cero). A partir de la implementación de una simulación computacional en Python, se llevaron a cabo múltiples experimentos con el fin de observar el comportamiento estadístico del modelo, verificando su correcto funcionamiento mediante pruebas básicas. Finalmente, se presentan gráficos que permiten visualizar la evolución de distintas variables relevantes a lo largo de las simulaciones, facilitando la interpretación de los resultados.

1. Introducción

La ruleta es un juego de azar clásico, ampliamente conocido por su aparente simplicidad y su trasfondo matemático. Su dinámica, basada en la aleatoriedad, la convierte en un caso ideal para estudiar fenómenos estadísticos y probabilísticos en un entorno controlado.

Existen dos versiones principales del juego: la ruleta europea, que contiene 37 casillas numeradas del 0 al 36, y la ruleta americana, que añade una casilla adicional con el doble cero (00), alcanzando un total de 38 casillas. Para esta investigación se ha optado por utilizar la ruleta europea, caracterizada por tener un único cero, lo que permite un análisis más ajustado al modelo clásico de probabilidad.

En este estudio, se desarrollará una simulación por computadora del juego de la ruleta mediante un generador de variables aleatorias, con el objetivo de analizar cómo se comportan distintos indicadores estadísticos a lo largo de múltiples tiradas. Las simulaciones estarán estructuradas en corridas, cada una compuesta por un número determinado de tiradas.

A partir de los datos generados, se calcularán y analizarán medidas como la frecuencia relativa, el valor promedio, la varianza y el desvío estándar, observando su evolución en relación con el número de tiradas y comparándolos con los valores esperados desde el punto de vista teórico.

Finalmente, se presentarán los resultados obtenidos mediante representaciones gráficas que permitirán visualizar con claridad las tendencias y discrepancias surgidas en el proceso.

2. Conceptos teóricos empleados

Para poder desarrollar este trabajo, necesitamos definir algunos conceptos teóricos importantes pertenecientes a la rama de Probabilidad y Estadística. Entre ellos, tenemos los siguientes:

Probabilidad: es una medida de la certidumbre de que ocurra un evento.

Variable aleatoria: es una función que asigna un valor, usualmente numérico, al resultado de un experimento aleatorio.

Distribución uniforme discreta: es una distribución de probabilidad discreta simétrica que surge en espacios de probabilidad equiprobables, es decir, en situaciones donde de resultados diferentes, todos tienen la misma probabilidad de ocurrir. En este caso, cada número de la ruleta tiene la misma probabilidad de ser elegido.

Población: conjunto total de elementos u objetos que comparten una o más características de interés para una investigación estadística.

Muestra: subconjunto de la población, es decir, un grupo finito de elementos a ser estudiados.

Frecuencia Relativa: se refiere a la proporción o el porcentaje de veces que ocurre un evento particular.

$$f_{r_i} = \frac{f_i}{n} \tag{1}$$

donde:

- f_{r_i} es la frecuencia relativa del valor x_i.
- f_i es la frecuencia absoluta del valor x_i.
- n es el número total de observaciones.

Media (Promedio): es la suma de un conjunto de valores dividida entre el número total de sumandos.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

donde:

- $\bar{x}$ es la media o promedio de la muestra.
- n es el número total de observaciones en la muestra.
- x_i son los valores individuales en la muestra, con $i = 1, 2, \ldots, n$.
- $\sum_{i=1}^{n} x_i$ representa la suma de todos los valores en la muestra.

Varianza: es una medida de dispersión definida como la esperanza del cuadrado de la desviación de dicha variable respecto a su media.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{3}$$

donde:

- s^2 es la varianza muestral acumulada.
- x_i es cada valor individual en la muestra.
- $\bar{x}$ es la media de la muestra.
- n es el número total de observaciones.
- Para $i = 1$, la varianza no está definida, y se asigna 0 por convención.

Desvío Estándar: es una medida que se utiliza para cuantificar la variación o la dispersión de un conjunto de datos numéricos.

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{4}$$

donde:

- s es el desvío estándar muestral.
- s^2 es la varianza muestral.

3. Desarrollo

Para llevar a cabo esta investigación, desarrollamos un programa utilizando el lenguaje de programación **Python** (versión 3.12.3), junto con bibliotecas como `random`, `matplotlib` y `statistics`. Estas herramientas nos permitieron simular tiradas de ruleta, calcular los valores estadísticos relevantes y generar representaciones gráficas para su análisis.

El programa fue diseñado para ejecutarse desde la línea de comandos, permitiendo al usuario ingresar los parámetros deseados mediante la instrucción: `python main.py -c XXX -n YYY -e ZZ`. De este modo, se puede controlar la cantidad de tiradas por corrida (`-n`), el número total de corridas (`-c`) y el número apostado por el jugador (`-e`), lo cual brinda flexibilidad y precisión al momento de realizar las simulaciones.

Finalmente, los resultados son presentados de forma visual mediante gráficos que muestran la evolución de los estadísticos a lo largo de las tiradas y corridas. Esta representación facilita la interpretación de los datos obtenidos y permite extraer conclusiones fundamentadas en el comportamiento observado en la simulación.

3.1. Resultados Obtenidos

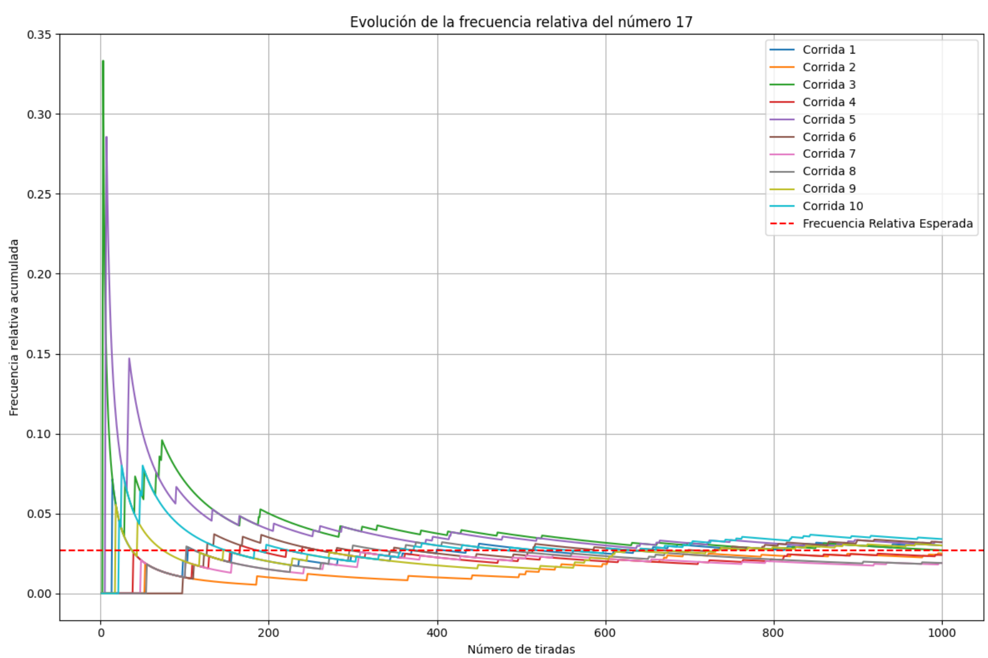

Figura 1: Frecuencia relativa del número 17 en 10 corridas de 1000 tiradas cada una. La línea roja punteada indica la frecuencia relativa teórica esperada, que es $1/37 \approx 0{,}027$

El gráfico de la Figura 1, evalúa si el número elegido aparece con la frecuencia que debería aparecer en una ruleta justa, es decir, en una ruleta donde todos los números tienen la misma probabilidad. A medida que se hacen más tiradas, se espera que la frecuencia relativa acumulada converja al valor esperado de 1/37. Si en una corrida se desvía mucho o no converge, puede indicar sesgo en la simulación o simplemente ruido estadístico (especialmente con pocas tiradas).

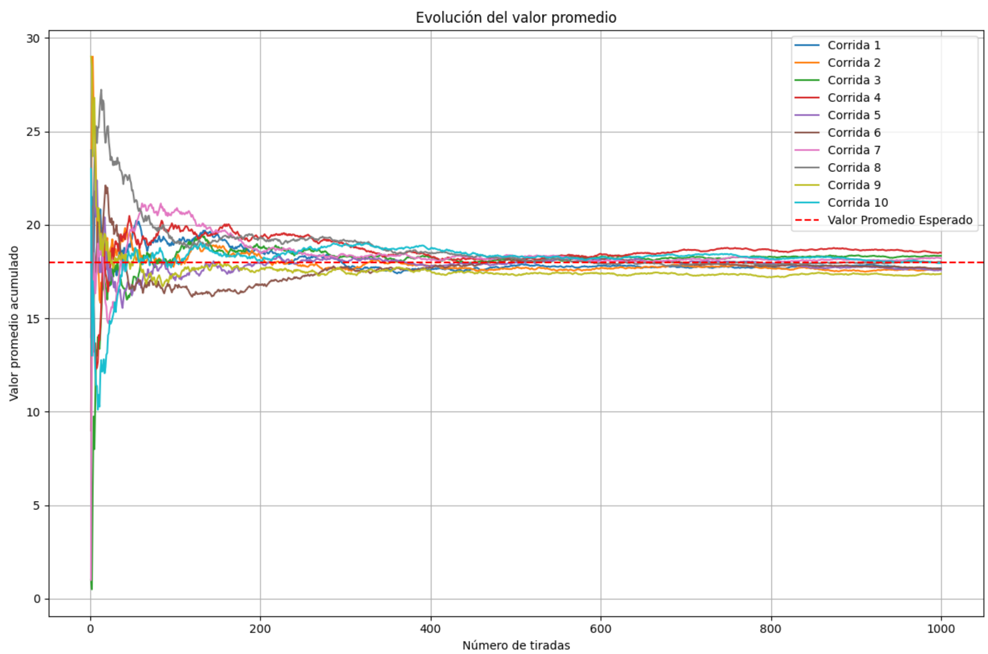

Figura 2: Promedio acumulado de los resultados de la ruleta en 10 corridas de 1000 tiradas cada una. La línea roja punteada indica el valor promedio esperado una ruleta europea uniforme, que es $((0+1+...+36)/37) = 666/37 \approx 18$

El gráfico de la Figura 2, permite verificar si el promedio de los valores obtenidos tiende hacia el promedio teórico. Si la simulación es correcta, las líneas de cada corrida deberían ir acercándose a 18 a medida que aumentan las tiradas. Al principio se puede observar que existe mucha variación, pero a medida que aumentan las tiradas, este tiende al valor esperado.

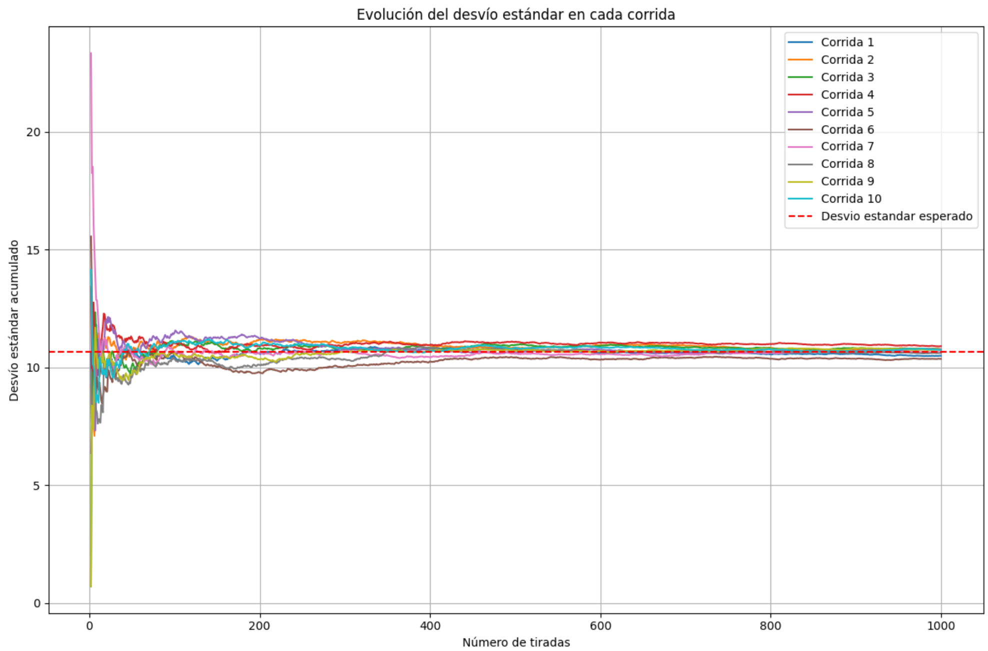

Figura 3: Variación del desvió estándar en 10 corridas de 1000 tiradas cada una. La línea roja punteada indica el desvío estándar teórico esperado para una variable aleatoria uniforme discreta entre 0 y 36

El gráfico de la Figura 3, permite ver cómo se estabiliza la dispersión de los resultados respecto al promedio. A medida que aumenta la cantidad de tiradas, el desvío estándar acumulado tiende a acercarse al valor teórico esperado, lo que indica que la simulación se comporta conforme a lo esperado por la teoría de probabilidad.

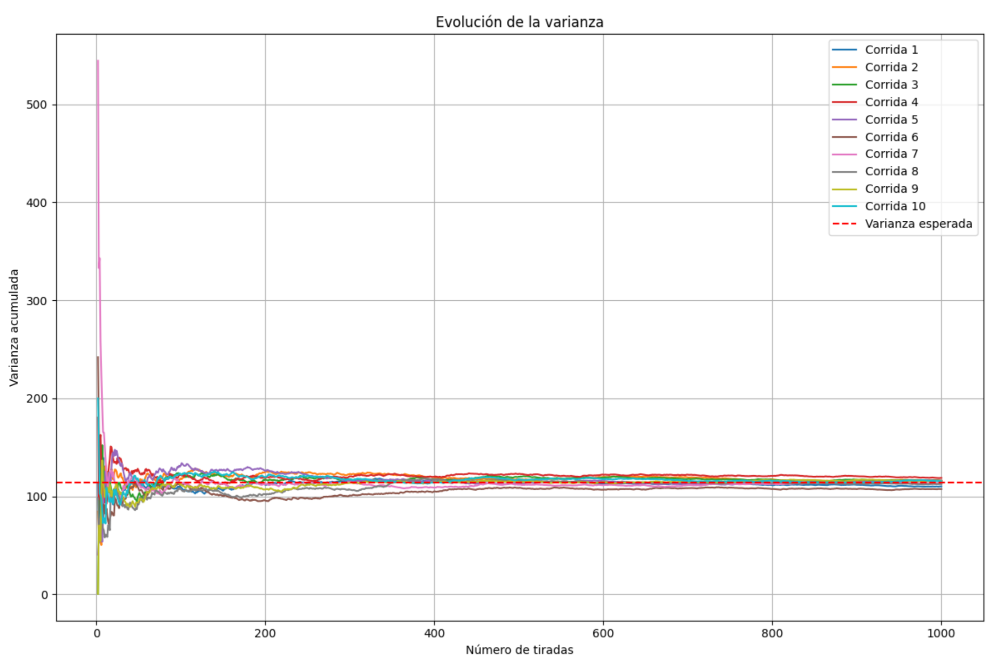

Figura 4: Varianza acumulada en 10 corridas de 1000 tiradas cada una. La línea roja punteada indica la varianza teórica esperada para una distribución uniforme discreta entre 0 y 36, que es $((37^2 - 1)/12) \approx 1132$.

El gráfico de la Figura 4, muestra cómo varía la dispersión de los datos respecto al promedio. Similar al desvío estándar, la varianza acumulada debería acercarse al valor teórico. Es útil para confirmar que no hay sesgos o errores sistemáticos en las simulaciones.

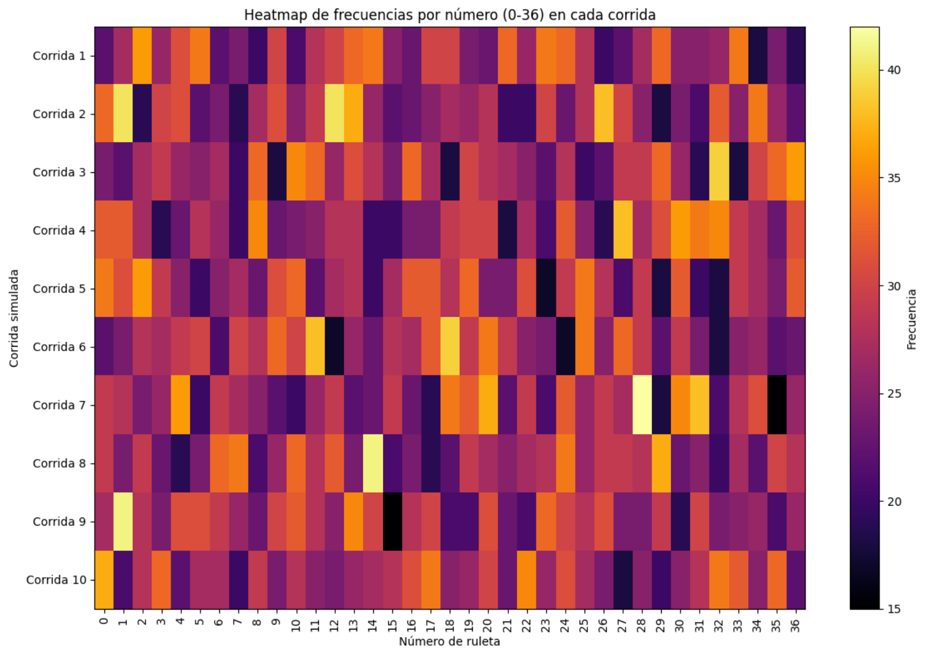

Figura 5: Mapa de calor 10 corridas de 1000 tiradas cada una

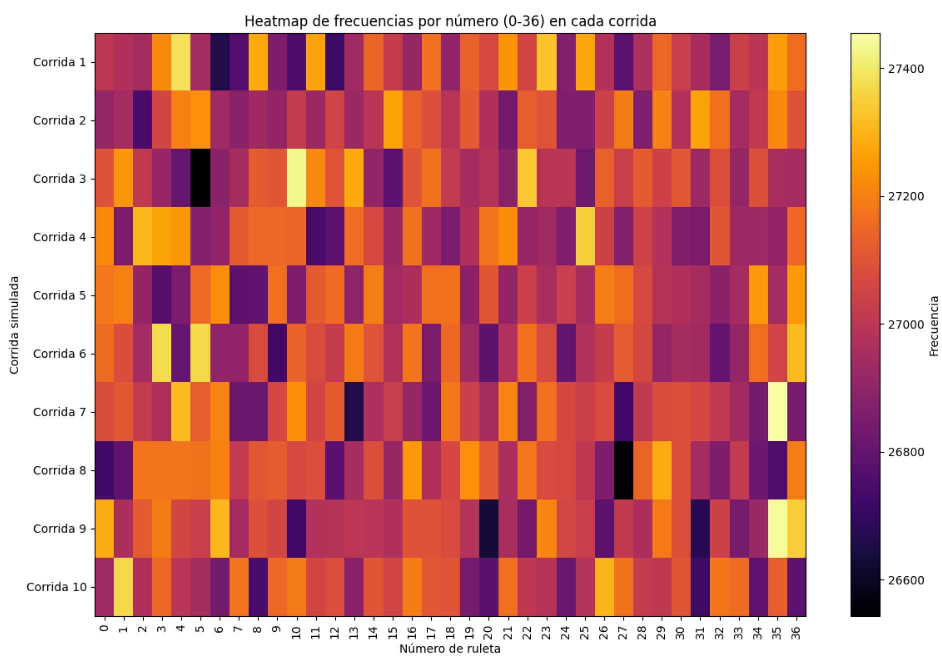

Figura 6: Mapa de calor 10 corridas de 1000000 tiradas cada una

Analizando los gráficos de la Figura 5 y de la Figura 6 podemos ver cómo afecta el número de tiradas elegido en la frecuencia de salida de los números. En el gráfico de la Figura 5, observamos que presenta mayores variaciones de colores, habiendo colores muy oscuros y muy claros al mismo tiempo. En cambio, en el gráfico de la Figura 6, notamos cierta uniformidad de los colores, lo que significa que al aumentar las tiradas, las frecuencias tienden al mismo número.

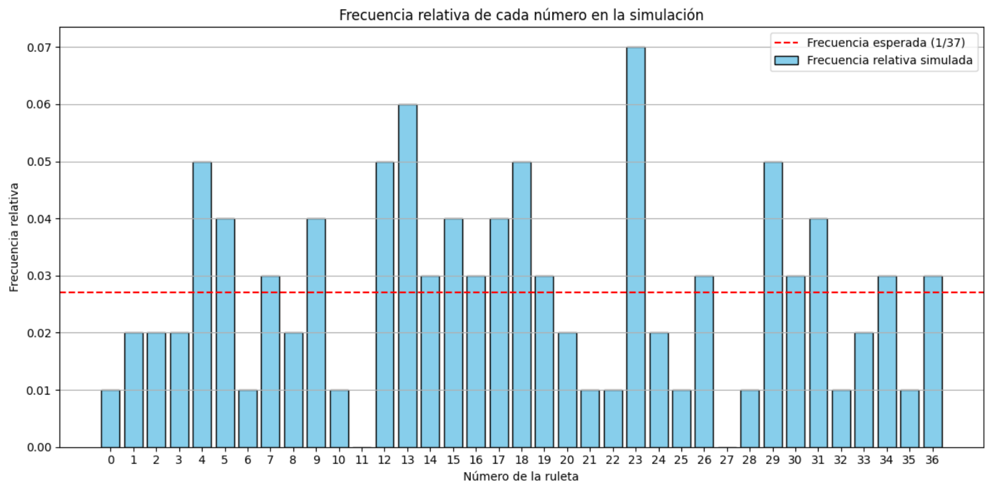

Figura 7: Frecuencia relativa de los números de una corrida de 100 tiradas

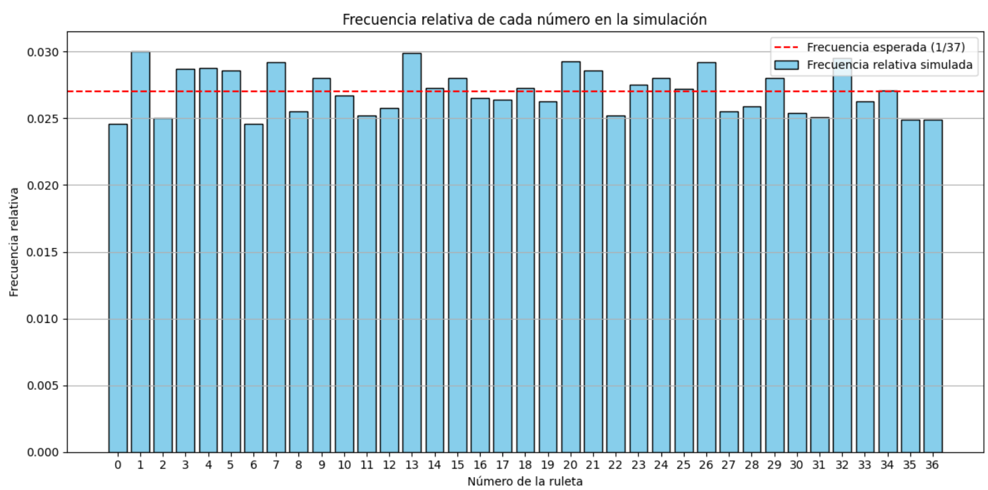

Figura 8: Frecuencia relativa de los números de una corrida de 10000 tiradas

Analizando los gráficos de la Figura 7 y de la Figura 8, podemos observar cómo la probabilidad acumulada de los números tiende a estabilizarse a medida que aumentan las tiradas. Con pocas tiradas, los valores de la probabilidad son más variados debido al factor del azar. Sin embargo, con un número mayor de tiradas, la probabilidad acumulada se ajusta hacia el valor esperado (alrededor de 1/37 por número en una ruleta justa).

4. Conclusiones

Los gráficos generados nos permitieron analizar el comportamiento estadístico de la simulación de una ruleta con 37 números (del 0 al 36). Al realizar múltiples corridas y observar cómo evolucionan diferentes medidas estadísticas a lo largo de las tiradas, pudimos evaluar la validez del modelo y su correspondencia con lo que esperábamos. A continuación, se resumen los principales aspectos observados:

4.1. Convergencia estadística a valores teóricos

A medida que se incrementa el número de tiradas, todas las métricas analizadas (frecuencia relativa, valor promedio, varianza y desvío estándar) tienden a estabilizarse y aproximarse a los valores teóricos esperados. Esto se refleja en:

- **Frecuencia relativa:** con más tiradas, la proporción de veces que sale un número determinado (por ejemplo, el 17) tiende a acercarse a 1/37.
- **Promedio acumulado:** se estabiliza en torno al valor esperado de la ruleta, que es 18.
- **Desvío estándar y varianza:** también convergen a sus respectivos valores teóricos.

4.2. Variabilidad entre corridas

Al graficar varias corridas en un mismo gráfico, se observa que, si bien hay cierta variabilidad al principio, todas tienden a comportarse de forma similar a medida que aumenta la cantidad de tiradas.

4.3. Importancia del tamaño de muestra

Los gráficos muestran claramente que con pocas tiradas los valores estadísticos pueden estar muy alejados de los valores esperados. Sin embargo, al aumentar la cantidad de tiradas, estas fluctuaciones disminuyen y los resultados se estabilizan. Esto pone en evidencia la importancia de trabajar con tamaños de muestra grandes cuando se analizan fenómenos aleatorios, para obtener conclusiones más fiables.

4.4. Resumen

En resumen, los gráficos confirman que el modelo de simulación de la ruleta se comporta según lo que predicen las leyes de la probabilidad. A través del análisis de frecuencia relativa, valor promedio, varianza y desvío estándar, se verifica que los resultados empíricos tienden hacia los valores teóricos esperados al aumentar la cantidad de tiradas. Esto permite concluir que la simulación es válida y que representa adecuadamente un proceso aleatorio equiprobable como el de una ruleta.

5. Bibliotecas Utilizadas

El desarrollo del programa se apoyó en una serie de bibliotecas del ecosistema de Python, ampliamente utilizadas tanto en ámbitos académicos como profesionales. Entre las bibliotecas utilizadas, se encuentran las siguientes:

Matplotlib es una biblioteca de visualización de datos que permite generar gráficos de alta calidad en distintos formatos. Es especialmente útil para representar series temporales, histogramas y distribuciones de probabilidad.

NumPy es una biblioteca especializada en el manejo de arreglos multidimensionales y en la ejecución eficiente de operaciones matemáticas sobre grandes volúmenes de datos.

Argparse facilita la construcción de interfaces de línea de comandos, permitiendo que el usuario configure parámetros de ejecución del programa de forma flexible y dinámica.

Random proporciona funciones para generar números aleatorios, esenciales en la creación de simulaciones que reproducen fenómenos basados en el azar.

Math contiene funciones matemáticas estándar, como exponenciales, logaritmos o raíces cuadradas, optimizadas para operar con precisión sobre números reales.

Statistics permite calcular medidas estadísticas clásicas como media, mediana, moda y desviación estándar, fundamentales para analizar el comportamiento de los datos simulados.

Referencias

[1] Wikipedia.

[2] Documentación Matplotlib

[3] Apuntes de la cátedra Probabilidad y Estadística UTN FRRO.