
TP 1.1 - SIMULACIÓN DE UNA RULETA

Aldana Risso Patrón

Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
rissopatronaldana7@gmail.com

Ignacio Fierro

Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
nachofier@gmail.com

Lucía Gelmetti

Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
luligelmetti@gmail.com

Juan Cruz Bonanno

Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
bonanno2340@gmail.com

Franco Reggiardo Chuglar

Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
francoreggiardo15@gmail.com

Marcos Oldani

Universidad Tecnológica Nacional - FRRO
Zeballos 1341, S2000, Argentina
marcosoldani1360@gmail.com

23 de abril de 2025

ABSTRACT

En este informe se presenta un análisis de los resultados obtenidos a partir de la simulación de la ruleta europea de casino, cuyos números van del 0 al 36, utilizando un programa desarrollado en Python 3.x. A través de múltiples corridas y con un tamaño muestral significativo, se realizó un estudio respaldado por conceptos de Probabilidad y Estadística, complementado con representaciones gráficas de los datos simulados.

1. Introducción

Este trabajo práctico tiene como objetivo simular mediante Python el comportamiento estadístico de una ruleta, permitiendo analizar conceptos de probabilidad, la frecuencia relativa, la varianza y la distribución de los resultados. A través de múltiples corridas y tiradas, se busca observar cómo se comporta el número elegido frente a la probabilidad teórica, y visualizar estos datos mediante gráficos generados con Matplotlib. Además, se incorporan herramientas de análisis estadístico y visualización para interpretar los resultados de forma clara y fundamentada.

2. Metodología

Para el desarrollo de este programa, empleamos varias librerías del lenguaje Python, entre las cuales se destacan “*numpy*” que se utiliza para el cálculo numérico y el análisis de datos, “*matplotlib*” que permite crear y personalizar gráficos en dos dimensiones, “*sys*” para la gestión de argumentos en la línea de comandos, y “*getopt*” para procesar las opciones ingresadas por la línea de comandos.

La simulación consistirá en realizar n corridas de la ruleta con m tiradas, seleccionando el número x para evaluar su comportamiento. En cada una de las corridas, se calcularán la varianza, desvío estándar, promedio y frecuencia relativa del número x , proporcionando así un análisis detallado de su distribución y comportamiento en el contexto del juego de azar simulado.

3. Marco Teórico

3.1. Definiciones Básicas de Probabilidad y Estadística

Para la buena comprensión del informe, se presentan a continuación una serie de definiciones necesarias:

- **Estadística:** la Estadística se puede comprender como “el arte de aprender a partir de los datos. Está relacionada con la recopilación de datos, su descripción subsiguiente y su análisis, lo que nos lleva a extraer conclusiones”.
- **Población:** es el grupo o colectivo que se estudia. Este término cuenta aquí con un significado más amplio que en su uso habitual, ya que, se puede referir a personas, cosas o medidas abstractas como el tiempo.
- **Muestra:** cuando se conoce la característica que se quiere estudiar y se ha delimitado la población de estudio, se procede a observar los elementos para obtener la información necesaria.

3.2. Ruleta europea vs. ruleta americana

Diferencias estructurales. Existen dos versiones principales del juego de la ruleta: la **ruleta europea** y la **ruleta americana**. Aunque ambas comparten la misma mecánica de juego, presentan diferencias importantes en cuanto a su estructura numérica y, por ende, en sus probabilidades.

La ruleta europea contiene un total de **37 casillas**, numeradas del 0 al 36. Por otro lado, la ruleta americana incorpora una casilla adicional identificada como 00, totalizando **38 casillas**. Esta diferencia aparentemente menor tiene un impacto significativo en las probabilidades de ganar para el jugador.

Probabilidad de acierto. Dado que la apuesta se realiza a un solo número, la probabilidad teórica de acertar dicho número varía entre ambas versiones. En la ruleta europea, esta probabilidad es:

$$P_{\text{europea}} = \frac{1}{37} \approx 0,0270 \quad (2,70 \%) \quad (1)$$

Mientras que en la ruleta americana es:

$$P_{\text{americana}} = \frac{1}{38} \approx 0,0263 \quad (2,63 \%) \quad (2)$$

Ventaja del casino. La **ventaja de la casa** representa el porcentaje del dinero apostado que, en promedio, retiene el casino a largo plazo. Para un juego justo, esta ventaja debería ser nula; sin embargo, las ruletas están diseñadas para que el casino siempre tenga un margen estadístico positivo. En la ruleta europea, dicha ventaja es del 2,70 %, mientras que en la americana asciende al 5,26 % debido a la presencia del doble cero:

$$\text{Ventaja}_{\text{europea}} = \frac{1}{37} \times 100 \approx 2,70 \% \quad (3)$$

$$\text{Ventaja}_{\text{americana}} = \frac{2}{38} \times 100 \approx 5,26 \% \quad (4)$$

Elección de tipo de ruleta para la simulación. Dado que la ruleta europea ofrece una menor ventaja al casino y es la versión predominante en los casinos europeos, se ha optado por utilizarla como base para las simulaciones realizadas en este trabajo práctico. No obstante, se analizan ambas variantes con el fin de comparar su comportamiento estadístico.

3.3. Parámetros Estadísticos

Los parámetros estadísticos son medidas numéricas que resumen características de un conjunto de datos. Estos proporcionan información importante sobre la distribución, la tendencia central y dispersión de los datos. Entre ellos tenemos:

3.3.1. Frecuencia Absoluta

En estadística, la frecuencia absoluta F_i hace referencia a la ocurrencia de un suceso al realizar el experimento una cantidad determinada de veces k . La suma de estas frecuencias da como número el resultado total de datos N .

$$\sum_{i=1}^k F_i = N \quad (5)$$

3.3.2. Frecuencia Relativa

La frecuencia relativa f_i representa la proporción o el porcentaje de veces que ocurre un determinado valor en un conjunto de datos, en relación con el tamaño total del conjunto de datos. Indica la frecuencia de un valor en relación con el total de observaciones.

$$f_i = \frac{F_i}{N} \quad (6)$$

Donde F_i es la frecuencia absoluta del valor x_i y N es el número total de observaciones.

3.3.3. Media

La media \bar{x} es el valor promedio de un conjunto de datos numéricos, calculada como la suma del conjunto de valores dividida entre el número total de valores.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (7)$$

Siendo N el número total de observaciones y x_i el valor obtenido en cada observación.

3.3.4. Varianza

La varianza σ^2 de una variable aleatoria es una medida de dispersión definida como la esperanza del cuadrado de la desviación de dicha variable respecto a su media.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (8)$$

Donde:

- σ^2 es la varianza poblacional.
- N es el tamaño de la muestra.
- x_i son los valores de la muestra.
- \bar{x} es la media de la muestra.

3.3.5. Desviación estándar

La desviación estándar σ es una medida que ofrece información sobre la dispersión media de una variable. La desviación estándar es siempre mayor o igual que cero.

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (9)$$

3.4. Teorema Central del Límite

El teorema central del límite se puede definir como un principio que determina que las medias de una serie de muestras de una población tendrán una distribución normal si el tamaño de la muestra es suficientemente grande, independientemente de la forma de la distribución de la población original. Además, sus varianzas son aproximadamente iguales a la varianza de la población a medida que aumenta el tamaño de la muestra, según la ley de los grandes números.

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (10)$$

Donde:

- Z es la variable estándar normal.
- \bar{X} es la media muestral.
- μ es la media de la distribución original.
- σ es la desviación estándar de la distribución original.
- n es el tamaño de la muestra.

3.5. Ley de los Grandes Números

La Ley de los Grandes Números describe el comportamiento de la media muestral cuando el tamaño de la muestra aumenta. Para una muestra suficientemente grande, la media observada se aproxima a la media poblacional.

Sea X_1, X_2, \dots, X_n una secuencia de variables aleatorias independientes e idénticamente distribuidas, cada una con esperanza matemática μ y varianza σ^2 finita, la media muestral (7) converge en probabilidad a μ cuando $n \rightarrow \infty$.

$$\bar{X}_n \xrightarrow{P} \mu \quad (11)$$

4. Ejecución de la Simulación

En esta sección presentaremos los gráficos resultantes de las diferentes simulaciones ejecutadas en conjunto con un breve desarrollo del mismo. A modo comparativo, para enriquecer el análisis, ejecutaremos distintas simulaciones, variando la cantidad de corridas de la simulación y número de tiradas por corrida. En primera instancia, se realiza una corrida de n tiradas y se comparan los valores esperados de la frecuencia relativa, el promedio, el desvío y la varianza, con dichos valores de un número X con respecto a n . Luego, se realizan 4 corridas de n tiradas para comparar los comportamientos de dichos valores. Para elegir un valor adecuado de n se hizo uso de la siguiente fórmula

$$n = \left(\frac{Z_\alpha}{2e} \right)^2 \quad (12)$$

Siendo:

- Z_α : valor obtenido mediante niveles de confianza.
- $2e$: límite aceptable de error muestral.

En el caso de este proyecto, se eligieron los valores más utilizados usualmente: un nivel de confianza del 95 % y un 3 % de error de estimación máximo aceptado.

Para obtener el valor crítico asociado a un nivel de confianza determinado en una distribución normal estándar, se debe buscar el valor z en la tabla de la distribución normal estándar que deja un área específica en las colas. En el caso del nivel de confianza del 95 %, se desea que el 95 % del área bajo la curva esté dentro del intervalo de confianza, lo que significa que se debe dejar un 2.5 % en cada cola. El valor de z que deja un área de 0.025 (2.5 %) en cada cola de la distribución normal estándar es aproximadamente 1.96.

Podemos calcular entonces, el valor de n de la siguiente manera:

$$n = \left(\frac{1.96}{2 \times 0.03} \right)^2 \approx 1067$$

Se elige entonces $n = 1000$ para realizar este proyecto. A continuación se presentan las gráficas obtenidas luego de la ejecución.

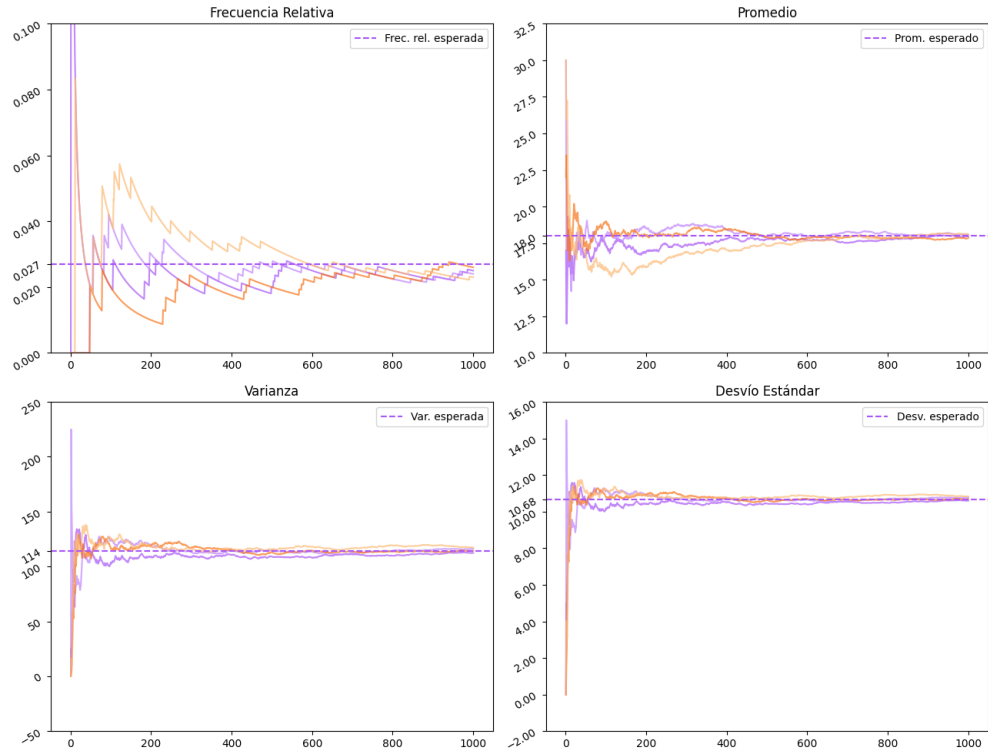


Figura 1: Parámetros estadísticos de varias corridas

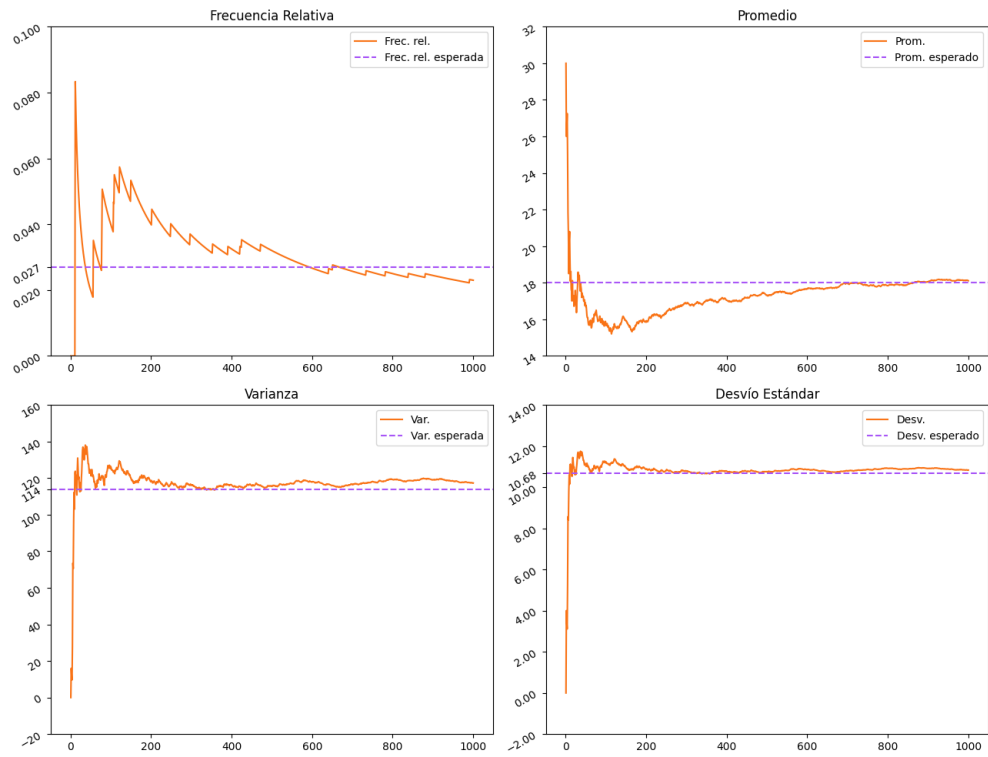


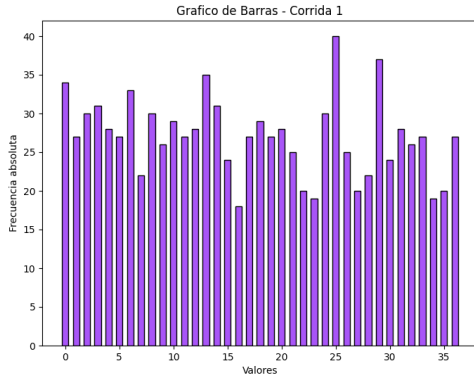
Figura 2: Parámetros estadísticos de una corrida

4.1. Gráficos de Barras

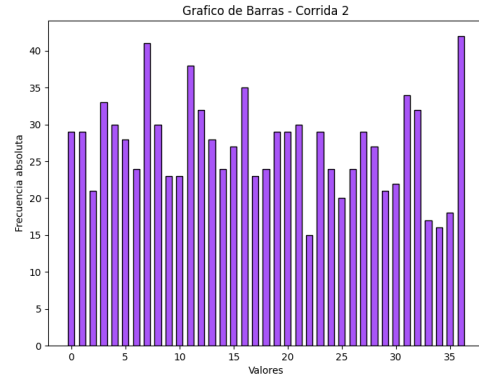
Descripción. Los siguientes gráficos de barras muestran los resultados de cuatro corridas independientes del experimento de simulación, cada una con 1000 tiradas. En cada gráfico se representa la frecuencia absoluta de aparición de los 37 números posibles en una ruleta europea (del 0 al 36).

Análisis Individual. En cada corrida individual, los números se distribuyen de forma relativamente uniforme, como es esperable en un sistema aleatorio justo. Se observan pequeñas fluctuaciones en la cantidad de apariciones de algunos valores, pero estas no siguen un patrón específico ni sugieren una desviación sistemática.

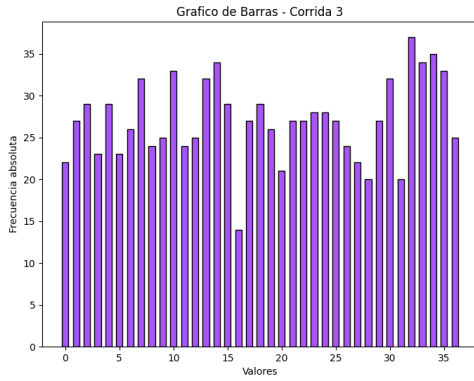
Estas diferencias responden a la variabilidad natural del azar y tienden a compensarse conforme aumenta el número de tiradas. Así, los gráficos de barras individuales muestran cómo la distribución de resultados se aproxima progresivamente a una frecuencia relativa cercana al 2.7 %, que es el valor teórico esperado para cada número.



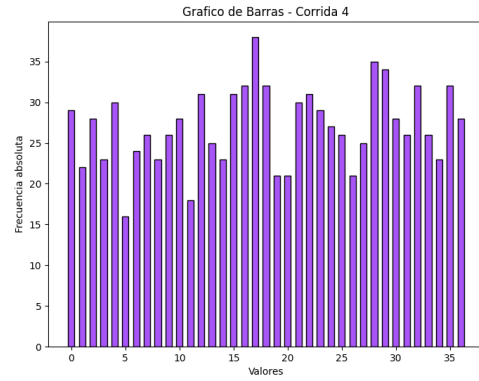
(a) Barras corrida 1



(b) Barras corrida 2



(c) Barras corrida 3



(d) Barras corrida 4

Figura 3: Barras individuales por corrida

Análisis Total. Este gráfico resume la frecuencia absoluta de cada número considerando el total de tiradas acumuladas (4000 en total, producto de las cuatro corridas de 1000 tiradas).

La distribución general de frecuencias es más uniforme que en los gráficos individuales. Las diferencias entre los valores se reducen y se alinean con la distribución teórica de una ruleta justa.

Este resultado ejemplifica claramente la Ley de los Grandes Números, que establece que al aumentar la cantidad de repeticiones de un experimento aleatorio, las frecuencias relativas tienden a estabilizarse en torno a sus valores teóricos. En este caso, cada número tiene una probabilidad de aparición cercana a $\frac{1}{37} \approx 2,7\%$, lo cual se refleja empíricamente en el gráfico.

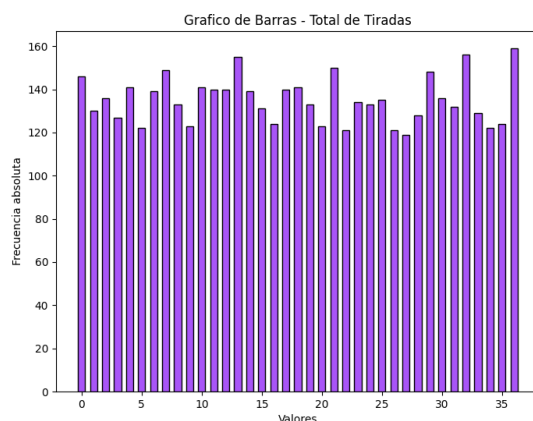


Figura 4: Barras total acumulado

4.2. Distribución por docenas

Descripción. El siguiente gráfico de torta agrupa los resultados de todas las tiradas según el sistema de docenas utilizado habitualmente en el juego de ruleta. Las categorías son:

- **Primera Docena:** números del 1 al 12.
- **Segunda Docena:** del 13 al 24.
- **Tercera Docena:** del 25 al 36.
- **Cero:** considerado aparte debido a su rol especial en la ruleta europea.

Esta organización responde a una forma común de apostar en los casinos, donde cada docena abarca 12 números y tiene una probabilidad teórica de acierto de $\frac{12}{37} \approx 32,43\%$. El cero no pertenece a ninguna docena y representa la ventaja de la casa.

Análisis Total. Al considerar la totalidad de las tiradas (4000), el gráfico muestra una distribución más equilibrada y ajustada a las proporciones teóricas. Las tres docenas convergen hacia valores cercanos al 32.43 %, y la aparición del número cero se estabiliza alrededor del 2.7 %. Este resultado confirma nuevamente que el modelo simulado reproduce adecuadamente el comportamiento probabilístico de una ruleta europea real.

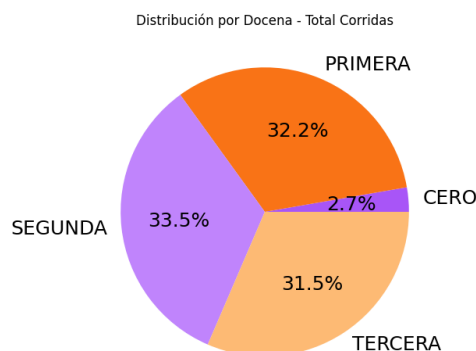
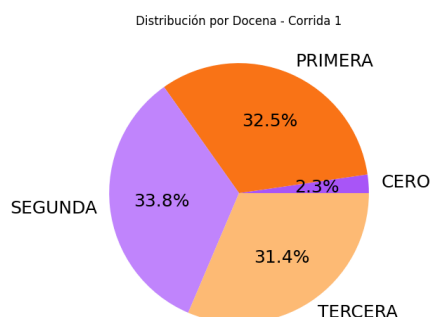
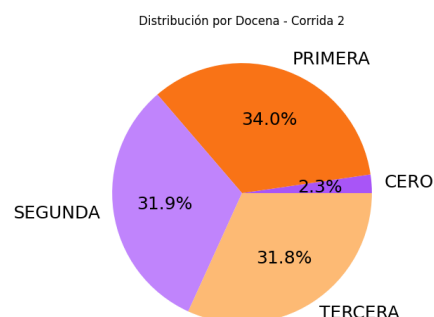


Figura 5: Gráfico de docenas, total corridas

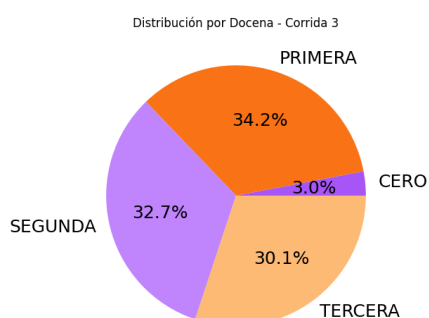
Análisis Individual. En los gráficos de cada corrida individual, se observa que las tres docenas presentan porcentajes que oscilan en torno al valor teórico esperado. Sin embargo, estas proporciones varían levemente entre corridas, lo cual es coherente con la variabilidad natural del azar. Aun así, no se detectan sesgos significativos ni patrones que indiquen una ruptura de la aleatoriedad: las fluctuaciones son aleatorias y no se concentran en una docena específica.



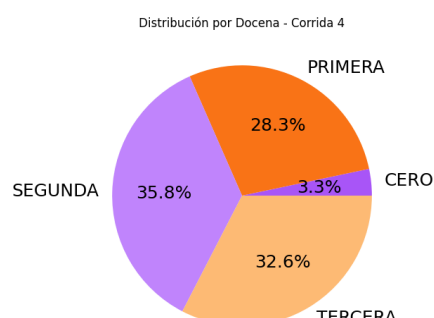
(a) Gráfico de docenas, corrida 1



(b) Gráfico de docenas, corrida 2



(c) Gráfico de docenas, corrida 3



(d) Gráfico de docenas, corrida 4

Figura 6: Gráficos de docenas individuales por corrida

4.3. Distribución por color

Descripción. Estos gráficos de torta ilustran la distribución porcentual de los resultados según el color de las casillas en la ruleta europea:

- **Rojo:** 18 números.
- **Negro:** 18 números.
- **Verde:** únicamente el número 0.

Esta división también forma parte de una modalidad típica de apuestas en ruleta. Los colores rojo y negro presentan una probabilidad teórica de $\frac{18}{37} \approx 48,65\%$, mientras que el verde (el cero) representa una probabilidad de $\frac{1}{37} \approx 2,7\%$.

Este tipo de análisis puede extenderse también a otras clasificaciones clásicas del juego, como par/impar o alto/bajo, que dividen el conjunto de resultados en mitades simétricas con la misma lógica estadística. En todos estos casos, el número cero opera como un elemento de desequilibrio que garantiza la ventaja del casino, ya que no pertenece a ninguno de los grupos apostables.

Análisis Individual. En cada corrida individual, se observa un reparto cercano al 50 % entre rojo y negro, con ligeras variaciones atribuibles a la aleatoriedad del experimento. La proporción del color verde, asociada al número cero, es notablemente menor, como se espera teóricamente.

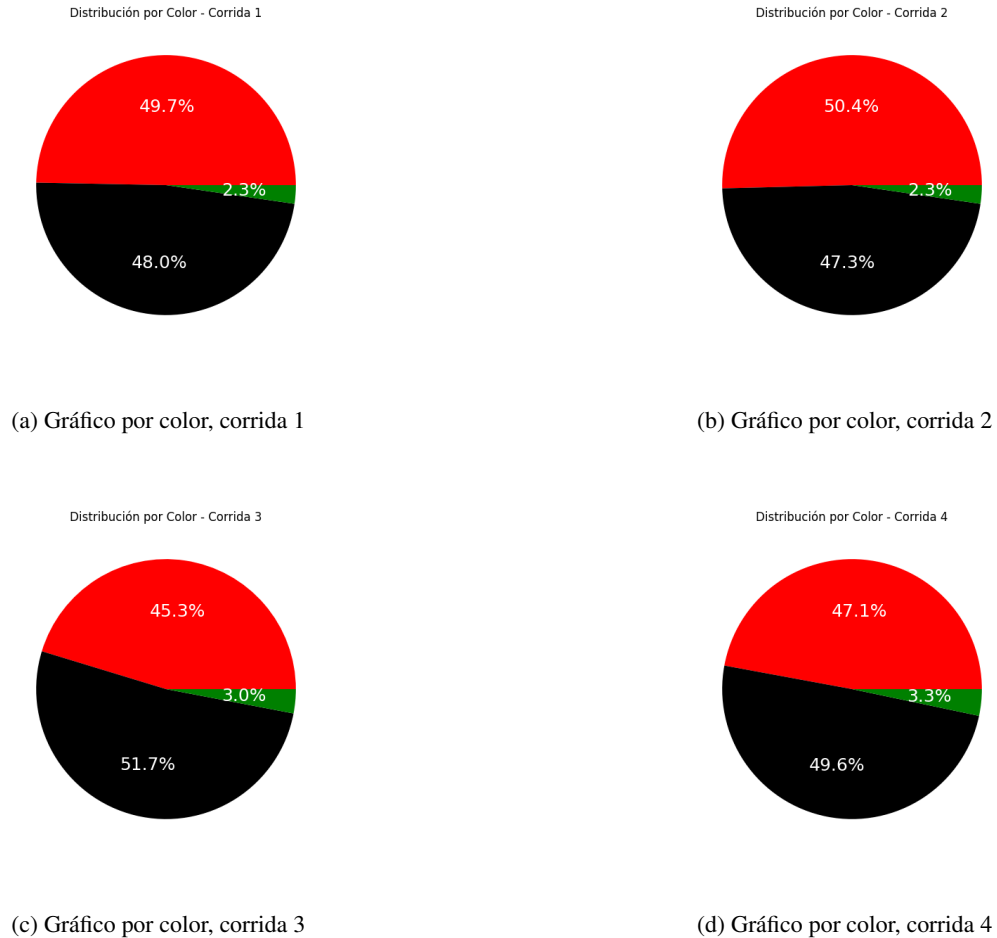


Figura 7: Gráfico por color individuales por corrida

Análisis Total. Al considerar la totalidad de las tiradas acumuladas, el gráfico general muestra una convergencia aún más clara hacia los valores esperados: rojo y negro tienden a estabilizarse cerca del 48.65 %, mientras que el verde se aproxima al 2.7 %. Esto refuerza la validez del modelo de simulación y su capacidad para reflejar el comportamiento real de una ruleta europea.

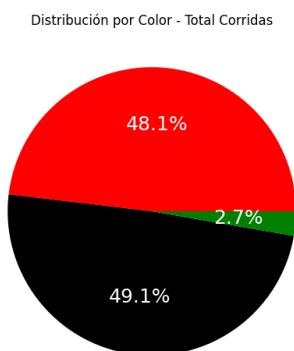


Figura 8: Gráfico torta, total corridas

4.4. Ilustración del Teorema Central del Limite

Para demostrar el Teorema Central del Límite, se generaron 500 corridas de 1000 tiradas, se calculó la media de cada corrida y se representaron en un histograma, con una curva de distribución normal (con una media y desvío estándar calculados a partir de estos datos).

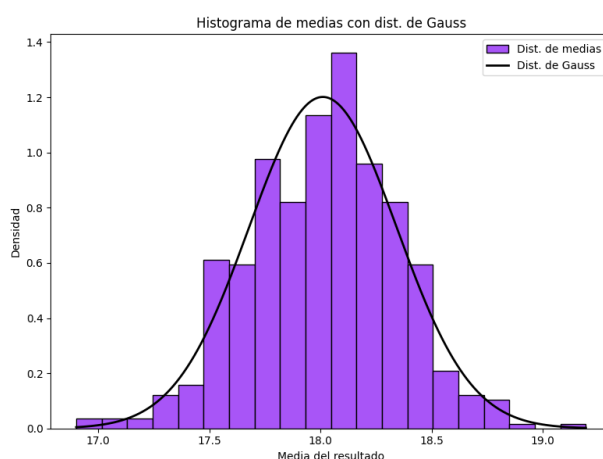


Figura 9: Histograma de medias con distribución de Gauss

Análisis. Esto muestra empíricamente que, aunque la variable original (un entero uniforme de 0 a 36) no sea normal, la distribución de sus medias muestrales tiende a serlo a medida que aumenta el número de observaciones.

5. Conclusión

A lo largo de este trabajo se logró simular de forma exitosa el comportamiento estadístico de una ruleta europea mediante técnicas de programación en Python y el uso de herramientas de análisis gráfico. Las simulaciones ejecutadas permitieron contrastar los resultados obtenidos con los valores teóricos esperados, brindando evidencia empírica de fenómenos fundamentales de la estadística, como la Ley de los Grandes Números y el Teorema Central del Límite.

Se observó que, si bien las corridas individuales presentan fluctuaciones en la frecuencia de aparición de los distintos números, estas diferencias tienden a estabilizarse a medida que aumenta la cantidad de tiradas. Este comportamiento es consistente con lo que se espera de un sistema aleatorio justo.

Asimismo, el análisis de las distribuciones por docenas, colores y otros criterios permitió validar que la simulación respeta las proporciones teóricas de la ruleta europea, donde cada número tiene una probabilidad aproximada del 2.7 %, los colores rojo y negro se reparten equitativamente, y el cero introduce la ventaja del casino.

Referencias

- [1] Aldana Risso Patrón. *TP 1.1 - Simulación de una ruleta (código fuente)*.
Disponible en: <https://github.com/AldanaRP/TPSimulacion>

- [2] Bacchini, H. *Introducción a la Probabilidad y a la Estadística*.
Universidad de Buenos Aires, Facultad de Ciencias Económicas, 2018.
Disponible en: http://bibliotecadigital.econ.uba.ar/download/libros/Bacchini_Introduccion-a-la-probabilidad-y-a-la-estadistica-2018.pdf