
TP 1.1 - SIMULACION DE LA TIRADA DE UNA RULETA

Ivo D'ursi
Victoria Merlino
Tomas Brasca
Jean Hilaire Messeroux
Luca Braccani
Cátedra Simulación
UTN - FRRO Zeballos 1341, S2000

13 de abril de 2022

Resumen

Simularemos las tiradas de una ruleta mediante un programa, generando diferentes conjuntos de datos pseudo-aleatorios y así poder analizar la variabilidad, el acercamiento al valor teórico esperado y la convergencia de estos conjuntos de manera gráfica y analítica.

Keywords— Ruleta · Tirada · Simulación · Estadística

1. Introducción

Vamos a intentar lograr una modelización que se encuentre lo más cercana posible de la ruleta real (el famoso juego de casino), simulando su comportamiento aleatorio o no determinista y adaptándonos a nuestras limitaciones de tiempo, números pseudo-aleatorios y muestras obtenidas no reales. Empezaremos por explicar que la ruleta consiste en 37 números secuenciales (0 al 36), en donde obtendremos uno de esos números de manera pseudo-aleatoria (con la librería random de python) simulando al juego real. En nuestro informe esto lo realizaremos en dos diferentes casos, el primer caso analizado corresponde a realizar **n = 100** veces la obtención de un número de la ruleta y en el segundo caso que presenta el informe lo realizamos con **n = 1000** y esto lo repetimos en **k = 30** a lo que llamamos **corridas**.º "tiradas". Hacemos esto porque la diferencia radica en que en las **k** veces que se corre el programa, nosotros tomamos diferentes valores y sus respectivos parámetros estadísticos para poder analizarlos luego y compararlos. Esto nos va a permitir llegar a mejores conclusiones ya que haciendo uso de la estadística, nos permite interpretar los datos obtenidos.

2. Marco Teórico

En el siguiente ensayo utilizaremos parámetros estadísticos para describir a las diferentes distribuciones de probabilidad. Lo que significa que todo valor obtenido va a ser a partir de una muestra.

2.1. Cálculo de promedio

El promedio no es una representación fija que aporte información útil al conjunto de datos, ya que, es muy sensible a valores extremos. Esto quiere decir que a valores muy grandes en el conjunto de datos, esta medida resulta inestable y tiende a variar mucho.

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_i^n X_i$$

- **Xi** : Cada valor del conjunto de observaciones
- **n**: Total de valores en el conjunto de observaciones

2.2. Calculo de la varianza

La unidad de medida de la varianza será siempre la unidad de medida correspondiente a los datos pero elevada al cuadrado. La varianza siempre es mayor o igual que cero. Al elevarse los residuos al cuadrado es matemáticamente imposible que la varianza salga negativa. Y de esa forma no puede ser menor que cero.

$$Var(x) = \frac{1}{n} \sum_i^n (x_i - E(x))^2$$

- **E(x)** esperanza matematica o promedio denotado en el parrafo anterior

2.3. Calculo del Desvio

La desviación estándar o desviación típica es una medida que ofrece información sobre la dispersión media de una variable. La desviación estándar es siempre mayor o igual que cero. Primero tomamos el número sobre el cual deseamos medir la desviación estándar y la cantidad de números que hay.

$$\sigma = \sqrt{\frac{1}{n} \sum_i^n (x_i - E(x))^2}$$

2.4. Calculo de la Frecuencia Relativa

La frecuencia de un evento es el número de veces en que dicho evento se repite durante un experimento o muestra estadística. Comúnmente, la distribución de la frecuencia suele visualizarse con el uso de histograma

$$F_i = \frac{n_i}{N}$$

- **N**: total de valores en el conjunto de observacion
- **n_i**: Un valor dentro del conjunto de observaciones, que va desde i.....N

2.5. Teorema central del limite

El teorema central del límite (TCL) es una teoría estadística que establece que, dada una muestra aleatoria suficientemente grande de la población (n > 30), la distribución de las medias muestrales seguirá una distribución normal. Es necesario que las distribuciones sean independientes.

2.6. Regla empirica

Esta regla en las estadísticas sugiere que cada dato que se puede observar caerá bajo tres desviaciones estándar diferentes de la media en una distribución normal. Siguiendo la regla:

- 68% — (μ ± σ)
- 95% — (μ ± 2σ)
- 99,7% — (μ ± 3σ)

donde **μ** es el promedio y **σ** es el desvio estandar ambos parametros pueden ser muestrales o poblacionales.

2.7. Desarrollo

En nuestro desarrollo tendremos dos programas, los cuales se diferenciaran entre un solo conjunto muestral y 30 conjuntos muestrales diferentes. A los cuales les calcularemos sus respectivas variables aleatorias y con el análisis de estas podremos acercarnos a una conclusión.

- **—** Representa el valor teorico esperado.
- **—** Representa valor calculado por nuestro programa.

3. Casos de estudio

Se haran dos casos de estudio, el primero constara de un experimento de 1000 repeticiones y el segundo se evaluara distintas cantidades de repeticiones como cantidades de experimentos.

3.1. Caso unica tirada

En este caso utilizamos un conjunto de 1000 tiradas de ruleta distintas en un solo conjunto muestral. Ademas de hacer varias corridas para el caso de la frecuencia muestral. Con esto, podemos visualizar como cada parametro estadistico.

3.1.1. Dispersion de las tiradas

En la figura 1 se observa como se encuentran dispersados los valores de un experimento en el cual se realizan N tiradas de ruletas(en este caso $N = 1000$), aunque no se puede observar que cada resultado no influya en el valor del siguiente. Debido a la utilizacion de un generador de numeros pseudoaleatorio, pero el estudio de esto va mas alla de alcance del trabajo. Entonces, esta grafica nos sirve para demostrar que efectivamente puede salir cualquier valor de la ruleta en una tirada de la misma.

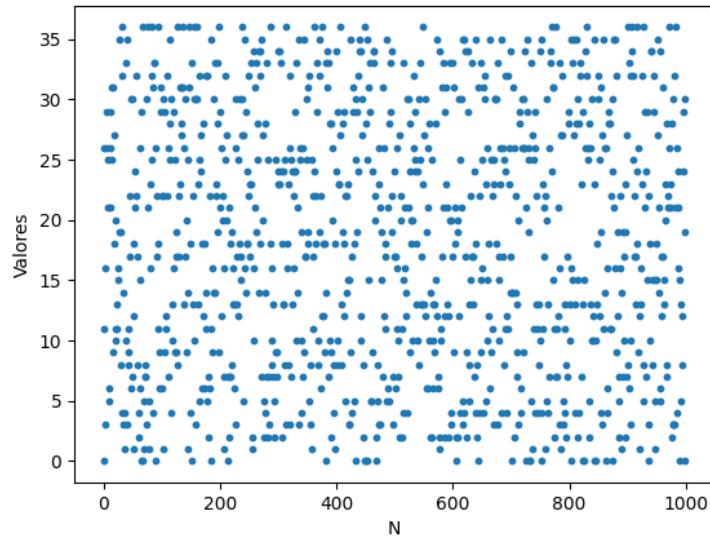


Figura 1: Dispersion de los resultados de cada tirada

3.1.2. Estabilidad en lo parametros

El estudio de los parametros estadisticos nos ayuda a comprender los valores en los cuales se agrupan mayoritariamente los datos.

En este estudio presentamos dos graficas que representan al desvio(a) y el promedio(b). Junto a ellos se encuentra en color rojo el valor "esperado" de cada parametro estadistico, los cuales fueron calculados en base a una muestra de 1000 tiradas.

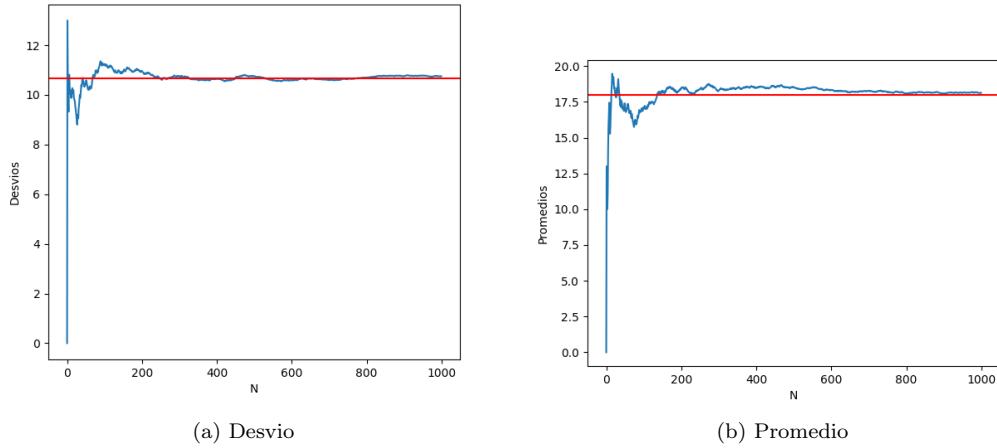
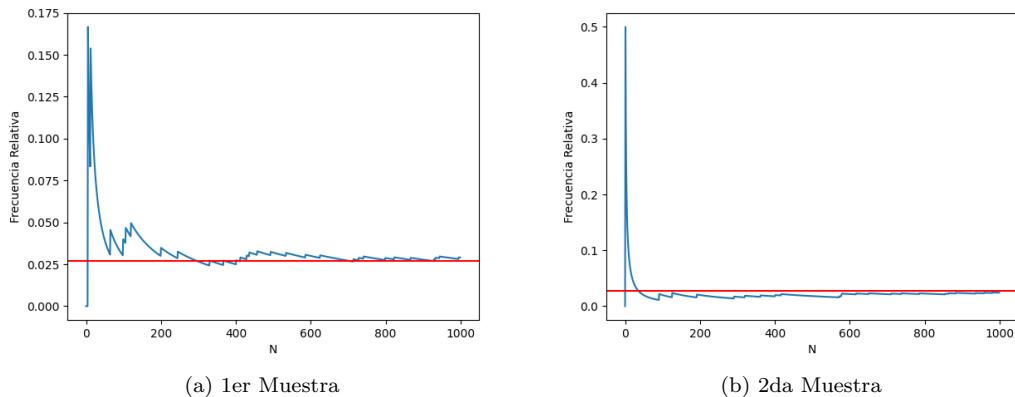


Figura 2: Valor de los parametros a medida que N aumenta

Volviendo con las graficas, podemos observar como a medida que aumenta el valor de N (tirada) los valores de las dos graficas en el eje y empieza a oscilar en nuestro valor “esperado”. Esto nos sirve para imponer una distribucion uniforme, gracias al teorema central del limite, con la cual tambien podremos calcular el valor de los parametros estadisticos, utilizando la formula adecuada para cada uno.

3.1.3. Análisis de distintas muestras

No es suficiente con analizar una solo experimento, por lo cual decidimos realizar mas muestras de las 1000 tiradas para ampliar nuestro estudio.



Las tres figuras de esta sección muestran la frecuencia relativa de un numero para 1000 tiradas. En el caso de las superiores (a) y (b) se observa un comportamiento parecido al de el desvio y el promedio (analizados anteriormente), el cual oscila en el valor “esperado”. Otra cosa a destacar, ademas de la estabilidad, es como, a diferencia de la figura (a), la figura (b) alcanza valores distintos, en este caso superiores a los de la Figura (a), antes de alcanzar el valor estable.

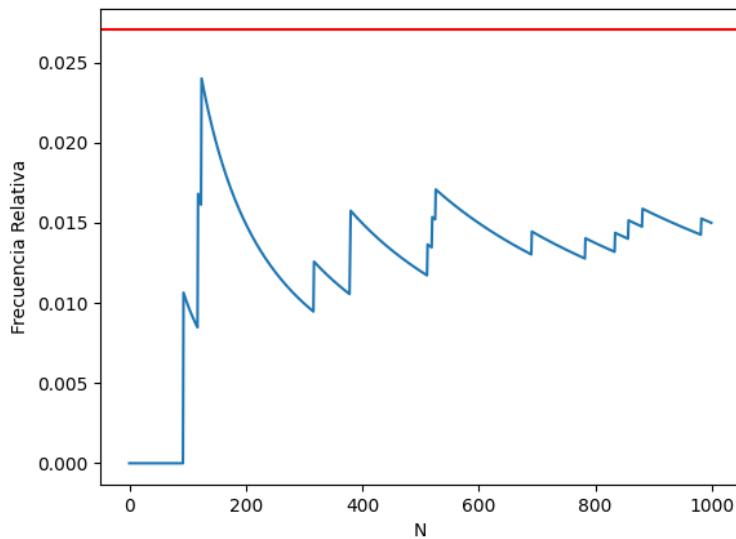


Figura 3: 3er muestra

Por otro lado, la figura 3 esta relacionado a una muestra de 1000 tiradas en la cual los valores, si bien ronda cerca del esperado, nunca lo supero y tampoco lo alcanzo. Aunque los valores no alcanzen el esperado, son tan relativamente pequeños comparados a los de la figura (a) y (b) que no podriamos decir que esta muestra a salido "malcon" respecto a las anteriores.

3.2. Caso multiples corridas

En este otro caso utilizamos un conjunto de N tiradas en cada conjuntos muestral(k), N y k variando dependiendo del analisis, y nos enfocaremos no en evaluar los valores de un parametro estadistico sino su entorno y llegar a distintas conclusiones. Por lo que, nos centraremos en los cambios observados a cambiar los parametros de entrada ($-$) de cada uno logrando como resultado final una variable aleatoria que representaria la evolucion de estos valores por cada tirada y que por el teorema central del limite tendra una distribucion normal. **Pudiendo aproximar junto con el desvio estandar y debido a la regla empirica el porcentaje de valores abarcados en la distribucion.(-)**

3.2.1. Aumento en la cantidad de muestras

Como se explico antes, se evaluara los impactos que tiene cambiar algun parametro, en este caso, en la entrada del programa y comparar los distintos resultados.

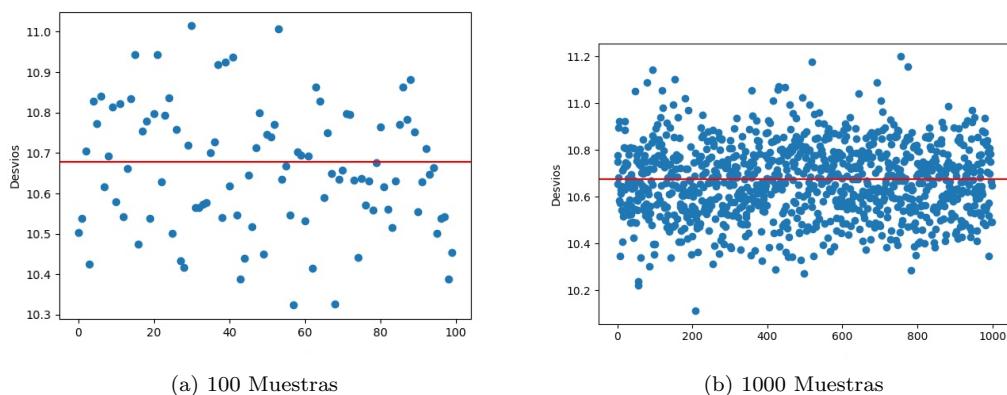


Figura 4: Valores del medida que varia k

Lo que podemos observar tanto en las figuras superiores (a) y (b), como en la inferior Figura 5. Es que a medida que aumentamos la cantidad de k , que simboliza a la cantidad de muestras, se puede observar una agrupacion por parte de los valores de las muestras, y como estos valores se acercan al “esperado” marcado con una linea de color rojo. Hay que recordar que en cada experimento la cantidad N , que hace referencia a la cantidad de tiradas, se mantiene constante a un valor de 1000

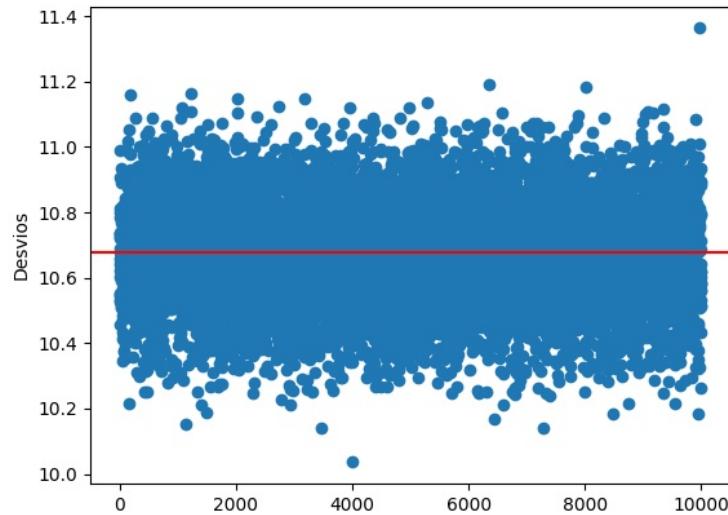


Figura 5: 10000 Muestras

Una cosa mas a observar es que como cada grafica es de un experimento diferente, se ven distintos valores que se alejaron de las medias, los cual podrian darse debido a que las tiradas fueron de muchos valores altos, o bajos.

3.2.2. Aumento en la cantidad de tiradas

No solo basta con cambiar la cantidad de la muestra, tambien decidimos probar incluyendo un experimento al cual se le haya dejado una cantidad de muestras igual a 1mil, pero cada una compuesta de 100mil tiradas diferentes. El resultado de este experimento se encuentra en la Figura 6 que, a simple vista, parece que tenemos un resultado parecido al de la figura de 1mil muestras en la seccion anterior. Pero, el cambio es observable en el dominio de valores que toma el desvio. Este dominio refleja los valores que el desvio tomo a lo largo de la iteracion. Y, como se puede observar, es mas acotado. Por lo que podemos concluir que al aumentar el numero de las tiradas en cada una de las muestra, los valores de sus parametros estadisticos estaran separados, pero a con valores muy pequeños.

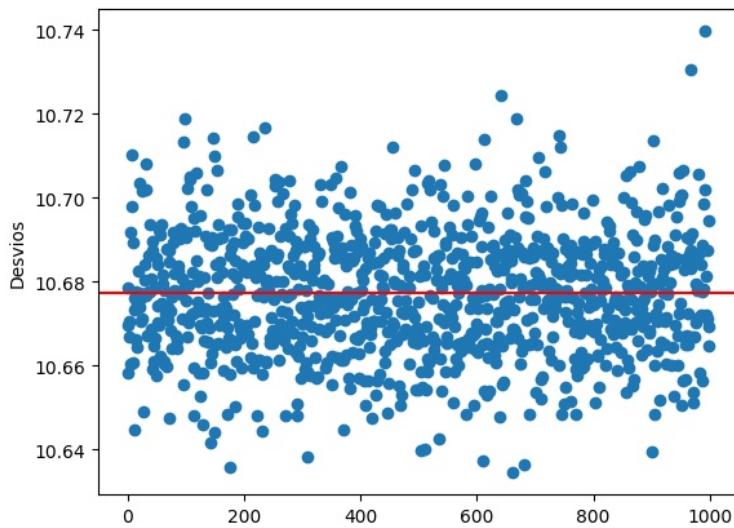


Figura 6: 1mil Muestras, 100mil Tiradas

4. Conclusión

Aunque parezca un simple juego de apuestas, la ruleta es ayuda para poner en practica conceptos como los de Probabilidad y Estadistica utilizados por nuestro grupo en el presente trabajo. Las formulas del Marco Teorico fueron de gran ayuda para el desarrollo del trabajo, ademas de ser el puente entre los numeros y la realidad como se ha demostrado en las distintas graficas realizadas por medio de una programacion. Por ultimo, destacar que los resultados obtenidos en las graficas se asemejan a los obtenidos mediante los calculos estadisticos reales. Ademas de dejar en evidencia la estabilidad de los resultados cuando se aumenta el numero de muestras

5. Discucion

5.1. Limitaciones

Para este caso asumimos que los sucesos son independientes, es decir, una observacion no influye en una futura ni es influenciada por una anterior, ademas las muestras obtenidas en nuestro caso son de caracter pseudoaleatorio, ya que utilizamos la libreria random que nos otorga python para "simular.^{el} numero que la ruleta saco.

5.2. Recomendaciones

Este experimento podria mejorarse obteniendo muchas muestras reales en diferentes ruletas, de esta manera obtendriamos valores reales, los cuales probablemente generarian unos resultados mas cercanos al valor teorico esperado.