

Herramientas Computacionales - Taller 11

SEMANA 13 - HERRAMIENTAS ESTADÍSTICAS
2018-I

La solución debe subirse a SicuaPlus en un único archivo de **IPython Notebook** con el nombre `NombreApellido_taller11.ipynb`, el cual debe contener toda la solución del taller.

1. (0.5 puntos) **Comentarios**

Comente su código.

El Teorema del Límite Central

En un juego de mesa hay un momento crucial de la partida en que un jugador debe lanzar 10 veces un dado de 6 caras y sumar todos sus resultados. El dado está cargado de tal manera que la probabilidad de obtener cada número es $P([1, 2, 3, 4, 5, 6]) = [0.2, 0.1, 0.2, 0.3, 0.1, 0.1]$. El jugador gana la partida si obtiene una suma mayor a 40.

En este taller queremos verificar cualitativamente cuánto es la probabilidad de obtener una suma mayor a 40. Para ello tenemos que simular muchas ($M \approx 10000$) sumas y realizar un histograma. La razón entre el número de sumas que caen en una barra y el número total es aproximadamente la probabilidad de obtener una suma en el intervalo de dicha barra.

2. (2.0 puntos) **Generar datos**

Para simular muchas sumas de lanzamiento de dados, primero tenemos que simular los lanzamientos del dado cargado. Para esto, cree una función llamada `get_sum(N)` que dado un número entero N retorne la suma de N números aleatorios siguiendo la distribución $P([1, 2, 3, 4, 5, 6]) = [0.2, 0.1, 0.2, 0.3, 0.1, 0.1]$.

Ayuda: Revise la documentación de `numpy.random.choice`.

3. (1.5 puntos) **Generar datos 2**

Ahora que podemos simular la suma de $N = 10$ lanzamientos de dados, debemos repetir el procedimiento para simular $M = 10000$ sumas y de esta manera poder realizar el histograma.

Para esto, cree una función llamada `repeat_sum(M,N)` que llame M veces la función `get_sum(N)`, cada vez guardando el resultado en una lista. Esta función debe retornar una lista de M elementos donde cada elemento es el resultado de una suma S_N , dada por la función `get_sum(N)`.

4. (1.0 puntos) **Análisis de datos y gráficas**

Una de las herramientas más poderosas de la estadística es el Teorema del Límite Central (TLC). Este teorema establece que si generamos un número aleatorio N veces y sumamos estos valores, la probabilidad de obtener cierto valor de la suma está aproximado por una gaussiana sin importar la distribución de probabilidad de donde vengan los números.

En este caso, el TLC nos dice que sin importar cómo el dado esté cargado, la probabilidad de obtener una suma S_N de N lanzamientos de dados, está dada por una gaussiana:

$$P(S_N) \approx \sqrt{\frac{1}{2\pi\sigma^2}} \exp \left[-\frac{(S_N - \bar{\chi})^2}{\sigma^2} \right], \quad (1)$$

donde σ es la desviación estándar y $\bar{\chi}$ es el promedio.

Usando la función del punto 3, obtenga $M = 10000$ valores S_N con $N = 10$. Imprima un mensaje que deje claro cuál el valor promedio y la desviación estándar de la lista de 10000 valores S_N .

Grafique el histograma normalizado de esta lista de valores. Adicionalmente grafique la curva gaussiana con los valores que acaba de obtener.

Así podemos verificar las posibilidades que tiene este jugador de ganar la partida.

Ayuda: Utilice la función `plt.hist(lista_vals, bins = 20, normed = True)` para graficar un histograma normalizado de 20 barras a partir de la lista llamada `lista_vals`

Ayuda: Grafique primero el histograma y luego verifique el rango de valores para poder graficar la gaussiana haciendo uso de un `linspace`.