

## ACTIVIDAD N°3

22 de Marzo del 2015

**Máximo:** 3 estudiantes por grupo

**Fecha de Entrega:** 18 de Abril del 2015

Para esta actividad usted deberá tener en cuenta lo siguiente:

- Cualquier intento de plagio de internet o entre grupos, automáticamente anulará el trabajo.
- El trabajo debe ser sustentado en la fecha de entrega.
- **Especificaciones técnicas de entrega**
  - Deberá generar el archivo **.ipybn** (generado por IPython Notebook o Jupyter Notebook) y subirlo a **github**.
  - Deberá publicar la url de Github en **NBViewer** de Python.

### 1. (Valor 2) Teoría estadística y Simulación

- (a) Dada una variable aleatoria  $X \sim \text{Gamma}(k, \theta)$  con parámetro de forma  $k > 0$ , parámetro de escala  $\theta > 0$  y con FDP dada por:

$$f_X(x) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} I_{(0, \infty)}(x)$$

donde  $\Gamma(k)$  es la función gamma dada por

$$\Gamma(k) = \int_0^\infty u^{k-1} e^{-u} du$$

Demuestre matemáticamente que:

- $E[X] = k\theta$
- $\text{Var}[X] = k\theta^2$

Experimentalmente muestre que  $E[X]_{\text{Teorico}} \approx E[X]_{\text{Experimental}} \approx \bar{X}$ , cuando  $n \rightarrow \infty$  donde  $n$  es el tamaño de muestra.

- (b) Dada una variable aleatoria  $X \sim \text{Exp}(\theta)$  con parámetro de escala  $\theta > 0$  y con FDP dada por:

$$f_X(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}} I_{(0, \infty)}(x)$$

Demuestre matemáticamente que:

- $E[X] = \theta$
- $\text{Var}[X] = \theta^2$

Experimentalmente muestre que  $E[X]_{\text{Teorico}} \approx E[X]_{\text{Experimental}} \approx \bar{X}$ , cuando  $n \rightarrow \infty$  donde  $n$  es el tamaño de muestra.

- (c) Realice las gráficas correspondientes y analice los resultados.

## 2. (Valor 2.5) Algo más de valor esperado

- (a) Hallar el valor esperado de lanzar simultáneamente un par de dados de 6 caras.
- (b) Dada la siguiente función de densidad conjunta de dos variables aleatorias  $X, Y$  esta dada por:

$$f(x, y) = \begin{cases} \frac{xy}{96} & 0 < x < 4, 1 < y < 5 \\ 0 & \text{otro modo} \end{cases}$$

- Hallar  $E[X]$ ,  $E[Y]$ ,  $E[XY]$  y  $E[2X + 3Y]$ .
  - Encuentre  $f_{y|x}(Y|X)$  y  $P(Y > \frac{1}{2} | \frac{1}{2} < X < \frac{1}{2} + dx)$
- (c) Dada la siguiente función de densidad conjunta de una variable aleatoria  $X$  es

$$f(x) = \begin{cases} \frac{4x(9-x^2)}{81} & 0 \leq x \leq 8 \\ 0 & \text{otro modo} \end{cases}$$

- Hallar la moda, la mediana y la media.
- (d) Dadas  $X_1, X_2, \dots, X_n$  variables aleatorias mutuamente independientes (discretas o continuas), cada una con media  $\mu$  y varianza  $\sigma^2$  finitas. Entonces si  $S_n = X_1 + X_2 + \dots + X_n$  con  $n = 1, 2, 3, \dots$ , demuestre que

$$\lim_{n \rightarrow \infty} P(|\frac{S_n}{n} - \mu| \geq \epsilon) = 0$$

**Pista:** Use el teorema de Chebyshev.

- (e) Dadas  $\{Y_n : n \geq 1\}$  una secuencia de variables aleatorias y asumiendo que  $\lim_{n \rightarrow \infty} E[|Y_n|] = 0$ . Demuestre que la probabilidad de  $\{Y_n : n \geq 1\}$  converge a 0. **Pista:** Use la desigualdad de Markov.

## 3. (Valor 5.5) Aplicación de Cadenas de Markov

- (a) **(Valor 1)** Diseñe e implemente una estrategia computacional para coleccionar todas las páginas web del dominio @unal.edu.co (Puede usar librerías tales como [crawler4j](#), [Scrappy](#), [CrawlerPhp](#)), realice una descripción de la estrategia computacional implementada.
- (b) **(Valor 1)** Implemente una estructura computacional para almacenar la páginas web coleccionadas y permita representar en la máquina el concepto de cadena de markov del grafo web. (**Pista:** Puede usar matrices, listas enlazadas, tablas hash para llevar a cabo esta tarea).
- (c) **(Valor 3.5)** Calcule el ranking de las páginas web del dominio @unal.edu.co haciendo uso del algoritmo de [Page Rank](#) para los nodos del grafo web. Liste las páginas con el PageRank más alto, realice un informe y discuta a partir de los resultados obtenidos.

## 4. Un Poco más de Markov

- (a) **(Valor +3)** Compare los resultados obtenidos con el algoritmo de PageRank y compare los resultados con el algoritmo de [HITS](#).

# Bibliografía

1. Knill, Oliver. Probability Theory and Stochastic Processes with Applications.2009. Overseas Press.
2. Gallager, Robert. **Stochastic Processes: Theory for Applications**. 2014. Cambridge University Press.
3. Manning, C. D., Raghavan, P., & Schütze, H. **Introduction to information retrieval**. 2008. Cambridge: Cambridge University Press.
4. Hue, Tue Huynh. Stochastic simulation and Applications in Finance with Matlab Programs.2008.Wiley.
5. Blanco, Liliana. Probabilidad. Universidad Nacional de Colombia.2014.
6. Winston, Wayne. Investigación de Operaciones Aplicaciones y Algoritmos. 2004. Thomson Ed.
7. Benjamin. **Notebook Ipyhon Tutorial**. 2014.