Probabilidad y Estadística - LCC

Trabajo Práctico Final 2022

Problema 1

Se sabe que ciertos errores aleatorios de medición $X_1,X_2,\ldots,X_n$, que contribuyen en forma independiente al error total, se distribuye uniformemente en el intervalo (0,1). Encuentre en forma empírica la función de densidad de probabilidad de la variable aleatoria E:

$$E= X_1+X_2+\ldots+X_{40}.$$

a. Muestre gráficamente los resultados obtenidos y concluya.

Problema 2

Se tiene un suceso para el cual la probabilidad de éxito es *p*. Se realizan intentos independientes de este suceso hasta obtener *k* éxitos consecutivos. Considere el proceso N_k que denota el número de ensayos necesarios para obtener k éxitos consecutivos.

a. Simule un cierto número de trayectorias del proceso N_k considerando un valor de $p > 0.5$ e indique en cada caso cuántas realizaciones de la experiencia fueron necesarias hasta obtener k éxitos consecutivos. (Elija un valor de k que satisfaga la siguiente desigualdad $1< k <3$).

b. A partir de los resultados obtenidos en a) estime, de ser posible, la $E(N_k)$.

Problema 3

Considere un proceso Bernoulli E_n : la señal emitida en el momento n es incorrecta y simule:

a. Una realización de dicho proceso (para n=10).

b. Una realización del proceso N_n: número de señales incorrectas emitidas al momento n.

c. Encuentre la Esperanza y la Variancia de ambos procesos.

Problema 4

Considere un jugador que tiene una apuesta inicial de k dólares y repetidamente apuesta 1 dólar en un juego en el cual la probabilidad de ganar es p y la probabilidad de perder es 1-p. La fortuna sucesiva del jugador es una simple caminata aleatoria iniciada en k (estado inicial del proceso $X_0= k$). Supongamos que el jugador decide detenerse cuando su fortuna alcanza S dólares (S > k), o cae a 0 dólar, lo que ocurra primero.

a. Siimule y visualice la evolución del capital del jugador para un valor de k y S pero para distintos valores de p ($p < 0.5$, $p=0.5$ y $p> 0.5$).

b. Estime, mediante la simulación de un número adecuado de trayectorias del capital del jugador, la probabilidad de ruina de dicho jugador para los distintos escenarios planteados en a).

Nota: Este es el clásico problema de la ruina del jugador, discutido por primera vez por los matemáticos Blaise Pascal y Pierre Fermat en 1656 y el cual sirve para describir muchos fenómenos similares en la actualidad.

Problema 5

El modelo Wright-Fisher describe la evolución de una población fija de k genes. Los genes pueden ser de uno de dos tipos, llamados alelos: A o a. Sea X_n el número de alelos A en la población en el momento n, donde el tiempo se mide por generaciones. Bajo este modelo, el número de alelos A en el momento n+1 se obtiene muestreando con reemplazo desde la población de genes en el momento n. Por lo tanto, habiendo i alelos de tipo A en el momento n, el número de alelos A en el momento n+1 tiene una distribución binomial con parámetros *k* y *p = i/k*. Esto resulta en una cadena de Markov con matriz de transición definida por:

$$P_{ij} = \binom{k}{j} \left(\frac{i}{k}\right)^j \left(1 - \frac{i}{k}\right)^{k-j}, \quad \text{for } 0 \leq i, j \leq k.$$

a. Simular este proceso para algún valor de k.
b. Observar qué valor toma P_{00} y P_{kk}.
c. Cuando la cadena progresa, la población en algún momento termina con todos alelos a (estado 0) o todos alelos A (estado k). Determinar cuál es la probabilidad de que la población evolucione al estado k.

Problema 6

Una persona está buscando cierta información en un universo de tan solo 5 páginas Web. Cada una de estas páginas puede tener uno o varios *links* a alguna otra. También puede haber páginas que no posean links.

La persona elegirá, a partir de la página actual que está mirando, la próxima a visitar seleccionando con igual probabilidad alguna de las linkeadas. Si la página actual no posee link alguno, entonces seleccionará con igual probabilidad cualquiera de las 5 páginas Web existentes. El esquema de red simplificado se presenta en la Figura 1:

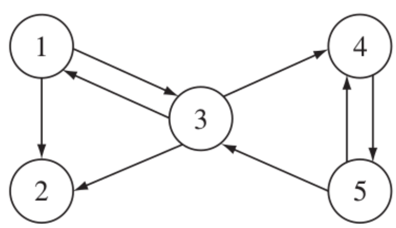

Figura 1. Red simplificada.

a. Modele el comportamiento de visitas a las páginas como una cadena de Markov. Especifique Matriz de transición en un paso y realice el grafo correspondiente.
b. Determine la probabilidad de visitar la página *j,* para *j=1,..,5.*

Nota: *Este modelo del navegador aleatorio forma las bases para el algoritmo PageRank que utiliza Google para determinar la importancia de una página en la Web. El ranking de una página está dado por la probabilidad estacionaria de la página en la cadena de Markov. El tamaño del espacio de estados en dicha cadena es de miles de millones de páginas.*

Probabilidad y Estadística - LCC
Trabajo Práctico Final 2022

Problema 7

En aplicaciones de seguridad informática, un *honeypot* (o sistema trampa) es una herramienta dispuesta en una red o sistema informático para ser el objetivo de un posible ataque informático, y así poder detectarlo y obtener información del mismo y del atacante. Los datos del *honeypot* son estudiados utilizando cadenas de markov. Se obtienen datos desde una base de datos central y se observan ataques contra cuatro puertos de computadoras - 80, 135, 139 y 445- durante un año. Los estados de la cadena de Markov son los cuatro puertos y se incluye un nodo indicando que ningún puerto está siendo atacado. Los datos de monitorean semanalmente y el puerto más atacado durante la semana es guardado. La matriz de transición para la cadena estimada para los ataques semanales es:

$$
\begin{array}{r}
 \\ 80 \\ 135 \\ \boldsymbol{P} = 139 \\ 445 \\ \text{No}
\end{array}
\begin{array}{c}
\begin{array}{ccccc} \;\;80\;\; & \;\;135\;\; & \;\;139\;\; & \;\;445\;\; & \text{No attack} \end{array} \\
\left(\begin{array}{ccccc}
0 & 0 & 0 & 0 & 1 \\
0 & 8/13 & 3/13 & 1/13 & 1/13 \\
1/16 & 3/16 & 3/8 & 1/4 & 1/8 \\
0 & 1/11 & 4/11 & 5/11 & 1/11 \\
0 & 1/8 & 1/2 & 1/8 & 1/4
\end{array}\right)
\end{array},
$$

con distribución inicial $\pi_0 = (0,0,0,0,1)$.

a. Después de 7 semanas, ¿cuáles son los puertos con más y menos probabilidad de ser atacados?

b. Encuentre la distribución límite (si es que existe) de los puertos atacados. Justifique.

Problema 8

Un cable submarino tiene defectos de acuerdo a un proceso de Poisson de parámetro $\lambda = 0.1$ por km.

- Simule una trayectoria de dicho proceso para una cierta cantidad de kms (máximo 10km).
- Grafique dicha trayectoria, tratando que la gráfica refleje las características fundamentales de dicho proceso.
- Simule una trayectoria del proceso para 100 km y registre la variable D: Distancia entre dos defectos sucesivos.
- Realice un análisis descriptivo gráfico y numérico conveniente para la variable D.

Este TP se finaliza con las respuestas a las preguntas que se realicen el día del examen final. Además puede requerir una defensa oral si los miembros de la cátedra lo consideran necesario.