

Universidad Distrital Francisco José de Caldas
Programa de Maestría en Ciencias de la Información y las Comunicaciones
Curso de Procesos Estocásticos
Primera Tarea
Fecha de Asignación: 10 de febrero de 2018
Fecha de entrega: 17 de febrero de 2018

Un experimento aleatorio consiste en medir el número de bytes que se reciben en el computador a través de su puerto de red durante un período de 10 segundos. ¿Cuál es el espacio muestral de este experimento, Ω ? Ante la pregunta "¿cuántos bytes se recibirán en el siguiente período de 10 segundos?", la respuesta correcta es "no sé". Sin embargo, intentaremos realizar al menos 180 repeticiones del experimento bajo condiciones idénticas para matizar un poco la respuesta a esa pregunta. Para eso llevaremos a cabo el siguiente proceso de observación:

En Windows abrimos una ventana de comandos con la aplicación Símbolo del Sistema (cmd), desde donde ejecutamos la siguiente instrucción

```
c:\> netstat -e 10 > estadisticas.txt
```

En seguida volvemos a Windows a navegar intensamente por internet, bajando una gran cantidad de grandes archivos continuamente, sin pausas entre ellos, durante al menos media hora, ojalá una hora. Al terminar de navegar, volvemos a la ventana de comandos y oprimimos CTRL-C. Entonces dispondremos de un archivo de estadísticas, `estadisticas.txt`, con N párrafos como el siguiente

```
Estadísticas de interfaz
```

| | Recibidos | Enviados |
|----------------------------|------------|-----------|
| Bytes | 2642623476 | 762006234 |
| Paquetes de unidifusión | 11990196 | 6479916 |
| Paquetes no de unidifusión | 22146 | 164304 |
| Descartados | 0 | 0 |
| Errores | 0 | 0 |
| Protocolos desconocidos | 0 | |

Si estuvimos media hora bajando archivos sin pausas entre ellos, lograremos hacer $N=180$ repeticiones del experimento; si estuvimos una hora, tendremos $N=360$ repeticiones (entre otros, podemos aprovechar y bajar los documentos del curso ☺). Lo importante es estar bajando archivos continuamente durante todo el tiempo que duren las repeticiones del experimento para intentar acercarnos a esa premisa imposible de "hacer repeticiones bajo condiciones idénticas". Cualquier pausa cambiará las condiciones de una repetición a otra y el ejercicio no será válido. También es importante que no haya procesos adicionales corriendo en el computador que interfieran con la captura de datos de `netstat`.

El campo Bytes recibidos dice cuántos bytes han sido recibidos desde alguna fecha particular hasta el momento del reporte. Si B_n es el número de bytes recibidos hasta el n -ésimo período y B_{n-1} es el número de bytes recibidos hasta el período n , la diferencia $X_n = B_n - B_{n-1}$ es el número de bytes recibidos durante el n -ésimo período, que es la cantidad que queremos observar con nuestro experimento "medir el número de bytes que se reciben en el computador a través de su puerto de red durante un período de 10 segundos".

Una gráfica de X_n en función de n para $n=1,2,3,\dots,N$ nos deberá revelar el concepto de aleatoriedad: ¿Podemos saber cuántos bytes habrían llegado en el período $N+1$?

Ahora calculemos el promedio muestral de los N datos que obtuvimos,

$$\underline{X} = \frac{1}{N} \sum_{n=1}^N X_n$$

Con este dato podemos dividir el espacio muestral Ω en dos subconjuntos: $A = \{x \in \Omega \mid x \geq \underline{X}\}$ y $A^c = \{x \in \Omega \mid x < \underline{X}\}$. En cada una de las N repeticiones que hicimos del experimento pudo ocurrir A o no, de manera que podemos contar cuántas veces ocurrió A hasta la repetición n , NA_n y, así, calcular la frecuencia relativa de A en las primeras n repeticiones del experimento:

$$g(n) = f_n(A) = \frac{NA_n}{n}, \quad n = 1, 2, \dots, N$$

Una gráfica de $g(n)$ para $n=1,2,3,\dots,N$ nos deberá revelar el concepto de regularidad estadística: ¿Podemos estimar un nivel de creencia apropiado de que en el período $N+1$ hubiera ocurrido el evento A ?

La tarea consiste en generar los datos $\{X_n, n=1,2,\dots,N\}$ para un valor de N mayor o igual a 180 (ojalá mucho mayor **cuidando preservar las condiciones del experimento**), presentar las gráficas de X_n vs n y de $g(n)$ vs n , discutir en un párrafo la relación que existe entre esas gráficas y los conceptos de aleatoriedad y regularidad estadística, y presentar un informe de una página que contenga las dos gráficas, las conclusiones, el nombre del autor, su código y su correo electrónico.

Recomendaciones: (1) La precisión conceptual en la redacción del párrafo de conclusiones será muy importante. (2) Si el informe toma más de una página, sólo se evaluará la primera. (3) No olvidar el correo electrónico.

Muchas gracias y diviértanse.

Marco.