Laboratorio 2
Estadística Computacional

Universidad Técnica Federico Santa María
Departamento de Informática

Héctor Allende Olivares
<hallende@inf.utfsm.cl>

Erick López Ovando
<elopez.ovando@gmail.com>

Camilo Valenzuela Carrasco
<camilo.valenzuela@alumnos.usm.cl>

11 de Noviembre del 2015

1. Objetivo

El alumno debe verificar la noción frecuentista de las probabilidades, para resolver algunos problemas no muy intuitivos, mediante un proceso de simulación. Además, debe verificar empíricamente el teorema de Glivenko-Cantelli[1]. En ambos casos, debe observar (y reportar) la exactitud y la precisión (a diferentes números de muestras), de forma análoga como fue presentado en la sesión presencial del LEC 2.

Para homogeneizar el trabajo de los grupos, todos usarán 100 iteraciones para cada tamaño de muestra.

2. Preguntas

1.- En un concurso, el participante escoge una puerta entre cinco, y su premio consiste en lo que se encuentra detrás. Una de ellas oculta un coche, y tras las otras cuatro hay una cabra. Sin embargo, antes de abrirla, el presentador, que sabe donde esta el premio, abre tres de las cinco puertas y muestra que detrás de ellas hay una cabra. Ahora el participante tiene una última oportunidad de cambiar la puerta escogida.

a) ¿Cuál es la probabilidad de ganar el auto si se cambia de puerta?

b) ¿Cuántos experimentos tuvo que realizar para lograr una convergencia razonable?. Justifique su razonamiento.
[Hint: obtenga la precisión de esta convergencia]

c) Realice un gráfico que muestre como la probabilidad empírica se acerca a la teórica.

2.- Dos estudiantes acurdan encontrarse en la biblioteca de la Universidad entre las 9 am y las 10 am un día Lunes. El primero que llega a la biblioteca espera al otro 10 minutos (dentro del intervalo de tiempo pactado). Si se supone que cada uno llega al azar en el intervalo de tiempo convenido y que los tiempos de llegada son independientes.

[1] https://en.wikipedia.org/wiki/Glivenko-Cantelli_theorem

a) ¿Cual es la probabilidad de que estos estudiantes se encuentren?

b) ¿Cuántos experimentos tuvo que realizar para lograr una convergencia razonable?. Justifique su razonamiento.

[Hint: obtenga la precisión de esta convergencia]

c) Realice un gráfico que muestre como la probabilidad empírica se acerca a la teórica.

3.- Una disputa entre apostadores en el año 1654 dio lugar a la creación de lo que hoy conocemos como teoría de probabilidades. Los involucrados en ese momento fueron dos famosos matemáticos: Blaise Pascal y Pierre de Fermat. Antoine Gombaud, caballero de Méré, noble interesado en los juegos de azar llamó la atención de Pascal con una aparente contradicción en un popular juego de dados. El juego consistía en arrojar un par de dados 24 veces; el problema era decidir si convenía o no apostar a que al menos una vez saldría un doble seis en las 24 tiradas. Una regla, aparentemente correcta, seguida por los apostadores, le hacía creer a de Méré que convenía apostar en favor de la ocurrencia de al menos un doble seis, sin embargo sus propios cálculos le indicaban lo opuesto.

a) ¿Qué respuesta le daría a de Méré?

b) Compliquemos un poco el problema: supongamos que se arrojan 3 dados 8 veces. ¿Conviene o no apostar a la ocurrencia de al menos un doble seis (al menos una vez, al menos en dos de los tres dados sale el número 6)?, ¿Cuál es aproximadamente la probabilidad de ganar?, ¿cuántos experimentos necesito para resolver el problema? (apóyese en un gráfico).

c) Responder nuevamente en el caso en que los tres dados se arrojan 9 veces en lugar de 8.

4.- Usando el comando `rbinom(n,10,.33)` genere muestras de diferentes tamaños `n` y estudie la convergencia de la función de distribución empírica hacia la función de distribución teórica.

a) Presente diferentes gráficos, a diferentes número de muestras, mostrando como se acerca la distribución empírica a la teórica.

b) Para alcanzar una convergencia con precisión menor a 0,01, ¿Qué tamaño debería tener la muestra?

5.- Usando el comando `rcauchy(n, location = 0, scale = 2.5)` genere muestras de diferentes tamaños `n` y estudie la convergencia de la función de distribución empírica hacia la función de distribución teórica.

a) Presente diferentes gráficos, a diferentes número de muestras, mostrando como se acerca la distribución empírica a la teórica.

b) Si cambiamos los parámetros de `location = 20` y `scale = 30`, ¿se necesita el mismo tamaño muestral para converger con la misma precisión anterior?. Si no es así, ¿cuánto debería ser el nuevo tamaño muestral para alcanzar una precisión comparable?. Justifique su respuesta de ser necesario.

6.- Usando el comando `rweibull(n, 10, 40)` genere muestras de diferentes tamaños `n` y estudie la convergencia de la función de distribución empírica hacia la función de distribución teórica.

a) Presente diferentes gráficos, a diferentes número de muestras, mostrando como se acerca la distribución empírica a la teórica.

b) ¿Cuánto es un tamaño muestral razonable para alcanzar la convergencia?. Justifique su respuesta.

7.- ¿Qué utilidad tienen el teorema de Glivenko-Cantelli en un sentido práctico?. Ejemplifíquelo con 3 aplicaciones reales.

3. Sobre la Entrega

1.- El informe debe ser realizado en LateX.

2.- Estructura del informe: El informe tiene que contener lo siguiente:

  2.1.- Introducción.

  2.2.- Desarrollo.

  2.3.- Conclusiones generales.

  2.4.- Referencias, si es que utilizaron. (Tienen que estar citadas en el informe).

3.- Los cálculos y gráficos deben realizarse utilizando R, este código no tiene que estar incluido dentro del informe.

4.- El informe y el código utilizado tiene que ser subido a moodle en formato nombreCompleto1-rol1-nombreCompleto2-rol2.tar.gz.

5.- Fecha de entrega: **Miércoles 18 de Noviembre**. Se descontarán 20 puntos por cada día de atraso.