

Primer Informe (4 de mayo de 2020)

1. *Cálculo de probabilidades mediante simulación.* La simulación del esquema de Bernoulli es una herramienta que permite el cálculo aproximado de probabilidades. La idea es simular un esquema de Bernoulli de parámetros (n, p) con n grande, y utilizar a la frecuencia como estimación de la probabilidad p . Vamos a utilizar entonces el esquema de Bernoulli para explorar numéricamente las cotas de grandes desvíos y de Bernoulli. Lo que queremos es trazar una tercera curva en color rojo con la probabilidad estimada y agregarla al siguiente gráfico, y agregar además la curva de probabilidades teórica en otro color diferente.

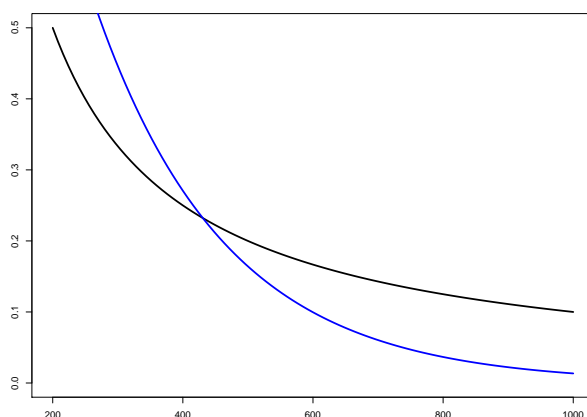


Figura 1: En negro la cota de Bernoulli, en azul la de grandes desvíos

Para eso:

1. Ir al teórico y obtener las fórmulas que están graficadas. Corresponden al problema de los nacimientos.
2. Graficar ambas cotas entre los valores de 200 y 1000.

3. El tercer paso es graficar una estimación de la probabilidad. Para eso hay que repetir una simulación con un número adecuado de experimentos, para distintos valores de n entre 200 y 1000. Generar un vector y plotearlo. Debe estar por debajo de las cotas graficadas primero. El valor de p y el intervalo para las frecuencias está en el ejemplo. Se sugiere usar 1000 simulaciones para cada valor de n .
4. Graficar el valor teórico de la probabilidad utilizando el comando `dbinom`, con los parámetros adecuados y sumando en el intervalo que corresponde.

2. Modelo binomial negativo. El modelo binomial negativo con parámetros (n, p) es un modelo de probabilidad en los naturales $\Omega = \{1, 2, \dots\}$ que cuenta en un esquema de Bernoulli el tiempo necesario para tener n éxitos.

1. Nos interesa mediante simulación calcular el tiempo medio de duración en este modelo. Para eso simulamos una cantidad $N = 10\ 000$ experimentos y calculamos el promedio en R. Elegimos dos casos: la cantidad de éxitos es $n = 50$ y $p = 0,5$ en el primero y $n = 10$ con $p = 0,05$ en el segundo.
2. Para completar el estudio, queremos determinar las probabilidades en el modelo para cada valor en Ω . Para eso vamos a graficar mediante el comando `barplot` dos cantidades para cada valor $\omega \in \Omega$. El primer valor es el teórico. Para eso hay que investigar qué paquete del R y qué comandos usar. El segundo es un valor simulado. Para graficar este valor hay que simular una gran cantidad de veces una variable binomial negativa, y graficar las frecuencias obtenidas de cada uno de los resultados posibles (hasta un valor razonable). Se espera obtener un gráfico de la forma

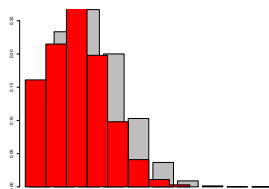


Figura 2: Cómo este pero mas prolijo y mas grande, que para cada valor se vean las dos barras

3. Estabilidad de las frecuencias. Mediante simulación comprobar el Teorema de Bernoulli. Para eso se requiere simular una serie de n variables aleatorias de Bernoulli, calcular las frecuencias sucesivas de éxito y graficarlas como función de n . Debe obtenerse un gráfico como el que sigue:

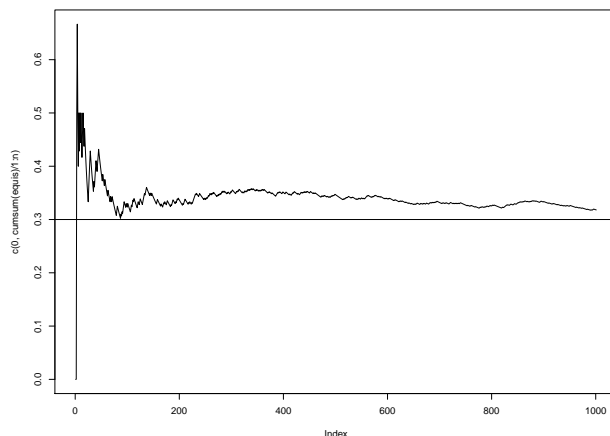


Figura 3: Estabilidad asintótica de las frecuencias

4. El recreo. En el patio del recreo los escolares juegan a la mancha. Al empezar, es uno el que mancha a los otros. Para eso los corre, pero puede tropezarse. Si se tropieza tiene que salir del juego. Los que son manchados empiezan a manchar a los otros, y también pueden tropezarse, y tienen que salir del juego. Tenemos entonces tres grupos de escolares: los libres (que mientras están libres no salen del juego), los que manchan, y los que ya no juegan porque tropezaron.

Para entender que pasa cuando termina el recreo, vamos a hacer los siguientes supuestos:

- Cuando empezamos hay uno que mancha, y todos los demás están libres.
- Contamos el tiempo en unidades, y actualizamos en cada una de esas unidades.
- Al pasar cada unidad de tiempo, cada manchado mancha a otro con una probabilidad β , y se tropieza con una probabilidad γ , que suponemos

pequeñas y constantes para todos los escolares durante todo el tiempo.

Queremos entonces graficar la evolución del juego. Para eso lo simulamos el juego una cantidad grande de veces y promediamos las curvas temporales para graficarlas. Suponemos en principio que $\gamma = 5 \times 10^{-3}$, que $\beta = 3 \times 10^{-4}$ y corremos el programa 2000 unidades de tiempo. Repetimos esto 1000 veces para promediar. Se puede también explorar el comportamiento para otros valores de los parámetros. (La idea es que cuando más pequeños son los parámetros, para mayor cantidad de unidades de tiempo, las gráficas que se obtienen son más suaves, más regulares.)

Datos sobre el informe

La fecha límite de entrega del informe es el viernes 22 de mayo. Se puede realizar de forma individual o en grupos de dos alumnos. Por favor avisen en el foro quienes formen grupos. Se entrega un documento R markdown y también el nootebook (el archivo .rmd). Son entonces dos archivos por cada informe que hay que entregar por mail, a los mails mordecki@cmat.edu.uy y juanpablo@cmat.edu.uy, con el tema: “Entrega del primer informe”.