Tarea 13 de Modelos Probabilistas Aplicados

Ley de los grandes números

5271

1 de diciembre de 2020

1. Introducción

En este documento se presenta las nociones básicas sobre la ley de los grandes números, ejemplos y aplicaciones.

2. Ley de los grandes números

LA ley de los grandes números plantea formalmente que con una sucesión de variables aleatorias independientes e idénticamente distribuidas con varianza finita se asegura que el promedio de las n primeras observaciones (variables aleatorias) se acerca a la media teórica cuando el número n de repeticiones tiende hacia infinito. Lo anterior se apoya en el teorema 8.2 del libro "*Introduction to Probability*"[1]:

Sean $X_1, X_2, ..., X_n$ un proceso de pruebas independientes e igualmente distribuidos con un valor esperado finito $\mu = E[X]$ y una y una varianza finita $\sigma^2 = \mathrm{Var}[X]$. Sea $S_n = X_1 + X_2 + \ldots + X_n$. Entonces para cualquier valor de $\epsilon > 0$ y n que tiende al infinito se tiene:

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0 \tag{1}$$

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) = 1 \tag{2}$$

Un caso particular de esta ley es el principio de estabilidad de las frecuencias, o teorema de Bernoulli, que ya hemos visto. Efectivamente, recordemos que una variable de Bernoulli es aquella que toma solo el valor 0 o 1 cuando no ocurre (u ocurre, respectivamente) un suceso A con probabilidades $1 - p$ y p. Sumar n variables de Bernoulli es contar el número de veces que se repite el suceso A en n pruebas. Una variable de Bernoulli tiene media p. Luego la media de n medias sera también p.

Para una mejor comprensión de lo anteriormente expuesto se tomará como ejemplo la resolución del ejercicio 3 de la página 312 del mismo libro, que dice:

Write a program to toss a coin 10,000 times. Let S_n be the number of heads in the first n tosses. Have your program print out, after every 1000 tosses, $S_n - \frac{n}{2}$. On the basis of this simulation, is it correct to say that you can expect heads about half of the time when you toss a coin a large number of times?

Para resolver este ejercicio se realiza el programa 2 en lenguaje R [3] que se muestra a continuación.

```r
n = 1000 # N mero de lanzamientos igual a 1000
Sn = c() # Numero de caras obtenidos
suma_caras = 0 # suma de los numeros de cara
n = c() # cantidad acomulada de lanzamientos
for(i in 1:10){
  simul = sample(0:1, 1000, rep = T) # Lanza una moneda 1000 veces
  suma_caras = suma_caras + sum(simul == 1) # Suma de las caras
  Sn = c(Sn, suma_caras / (1000 * i)) # N mero de caras
  n = c(n, 1000 * i) # N mero de lanzamientos.
}
data_frame = data.frame(n, Sn)
data_frame
```

Tarea13.R

El cuadro 1 de la página 2 se puede observar claramente, que para el programa 2 donde se ejecuta diez simulaciones de 1000 lanzamientos de una moneda donde la cara vale uno y la cruz 0. Es correcto decir que puede esperar cara la mitad de las veces cuando lanza una moneda una gran cantidad de veces. En la figura 1 de la página 3 se muestra los resultados para 100 repeticiones.

Cuadro 1: Resultados de la simulación de sacar cara en el lanzamiento de una moneda

	n	***Sn***
1	1,000	0.513
2	2,000	0.511
3	3,000	0.510
4	4,000	0.506
5	5,000	0.504
6	6,000	0.505
7	7,000	0.503
8	8,000	0.502
9	9,000	0.502
10	10,000	0.502

3. Aplicación en la estimación de subconjuntos en un método estocástico de ramificación y poda

En [2], se propone un método estocástico de ramificación y poda para resolver problemas de optimización global estocástica. Como en el caso determinista, el conjunto factible se divide en subconjuntos compactos. Para guiar el proceso de partición, el método utiliza estimaciones estocásticas superior e inferior del valor óptimo de la función objetivo en cada subconjunto. Y se demuestra la convergencia del método y se obtienen estimaciones de precisión aleatorias. Se discuten los métodos para construir límites estocásticos superior e inferior a través de la ley de los grandes números. Las consideraciones teóricas la ilustran con un ejemplo de un problema de ubicación de instalaciones.

En el método estocástico de ramificación y poda [2] se construye una secuencia de conjuntos $X^k(\omega) \subset X^{k-1}(\omega)$, y se tiene que estimar el valor de cota inferior $L(\cdot)$ en el límite que establece $X^* = \lim_k X(\omega)$,

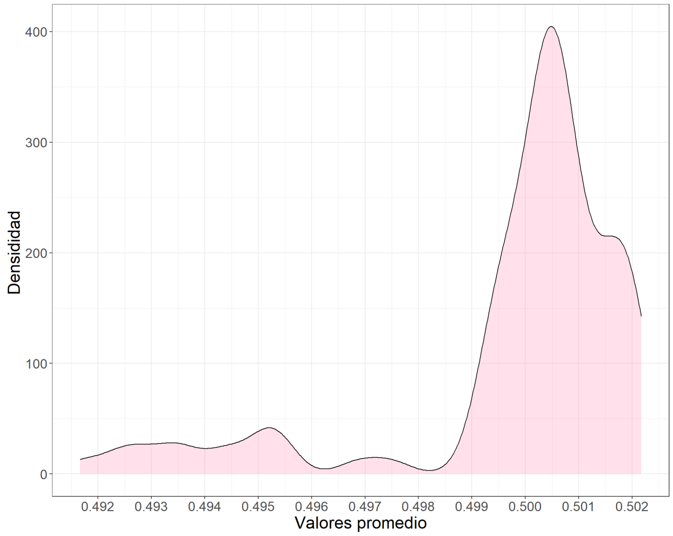

Figura 1: Diagrama de densidad de los resultados para 100 repeticiones del experimento.

usando observaciones independientes de variables aleatorias $\epsilon(X^k)$ tales que $\mathbb{E}[\epsilon(X^k)] = L(X^k)$. A tal efecto, en[2] se utiliza la siguiente estimación:

$$L_k(X^k) = \frac{1}{k}\sum_{s=1}^{[} k] \to L(X^*). \tag{3}$$

Referencias

[1] Grinstead, Charles M., Snell, J. Laurie. *Introduction to Probability.* American Mathematical Society, 2006.

[2] Vladimir I Norkin, Georg Ch Pflug, and Andrzej Ruszczyński. A branch and bound method for stochastic global optimization. *Mathematical programming*, 83(1-3):425–450, 1998.

[3] R Core Team. R: Un lenguaje y un entorno para la informática estadística, 2020.