

Taller Simulación de la Probabilidad

1. Suponga que se tienen 100 chips de memoria en una caja, de los cuales 90 se encuentran en “buen” estado y 10 se encuentran “dañados”. Debemos escoger 5 entre los 100 aleatoriamente para actualizar una computadora. ¿Cuál es la probabilidad que los cinco chips sean buenos?
 - Use la función *sample* para aproximar $P(X = 5)$ por simulación
 - Explore una alternativa diferente a la vista en clase para escoger los elementos “buenos” y los elementos “malos” en una muestra.
2. Suponga que se tienen 25 personas en un cuarto. ¿Cuál es la probabilidad que dos o mas de ellos tengan cumplan años el mismo día?
 - Verifique mediante una gráfica la hipótesis donde a mayor número de personas en un cuarto, mayor es la probabilidad de coincidencias.
 - Use la función *sample* para simular la probabilidad de emparejamiento para las 25 personas en 100000 cuartos.
3. La variable aleatoria X del ejercicio (1) corresponde a una distribución Hipergeométrica. En R las probabilidades de este modelo $P(X = x)$, se calculan con la función *dhyper*.
 - (a) ¿Cuál es la relación entre `sample(1:100,5)` y `choose(100,5)`?
 - (b) Calcule $P(X = 2)$ usando *dhyper* y resuelva también `choose`
 - (c) Corra el programa del ejercicio (1) siguiendo las instrucciones `mean(bueno)` y `var(bueno)`. ¿Cuál es el resultado de estas dos instrucciones? En términos de la variable aleatoria X , ¿qué aproximación es mejor?
 - (d) Ejecute `sum((0:5)*dhyper(0:5,90,10,5))`, ¿qué conexión tiene este resultado con el punto (3c)?
4. Basados en el ejercicio (2)
 - (a) Ejecute cada una de las siguientes instrucciones en el orden dado, explique cada una. Indique la longitud de cada vector.
`a = c(5, 6, 7, 6, 8, 7); length(a); unique(a)`
`length(unique(a)); length(a) - length(unique(a))`
`duplicated(a); length(duplicated(a)); sum(duplicated(a))`
 - (b) Basado en los resultados encontrados en (4a), proponga una manera de contar los cumpleaños coincidentes que no sea usando el comando `unique`. Implemente un programa alternativo del punto (2), reporte sus resultados.

5. Un juego en un casino consiste en lanzar 3 dados. Las ganancias son directamente proporcionales al número total de 6 que salgan en los dados. Un apostador ha jugado 100 veces realice una simulación de los resultados.
 - (a) del ejercicio anterior realice una simulación de 10, 100, y 1000 juegos, ¿Qué tan cerca se encuentra esta simulación de los datos esperados?
6. Simula $n=10000$ observaciones de una variable aleatoria: (a) Binomial (b) Poisson (c) Hipergeométrica (d) Binomial Negativa (e) Normal (f) Gamma (g) Weibull, para diferentes parámetros en la distribución y explique el comportamiento a partir de los gráficos (distribución y acumulados) realizados con las observaciones simuladas.
7. Un juego de poker consiste en seleccionar 5 cartas de un mazo de 52. (Hay cuatro Aces en el mazo)
 - (a) Use el combinatorio para calcular la probabilidad que el juego de poker no tenga Aces. Use R para encontrar la respuesta numérica correcta en cinco pasos.
 - (b) Use la metodología del ejercicio (1) para aproximar la probabilidad en la parte (7a) por simulación.
8. Utilice el método de integración aproximado Monte Carlo para estimar:

$$\int_0^{\pi/4} \log(1 + \tan^2(x)) dx$$

con $n = 1000000$

Metodología

- El trabajo debe ser realizado en grupos de 2 personas
- El informe debe ser desarrollado en L^AT_EX no debe exceder de 4 páginas (adjuntarlo en el enlace del campus virtual), incluyendo gráficos y bibliografía.
- La fecha de entrega y presentación de resultados es el día Febrero 20 de 2018, a las 10 am.
- El código de las simulaciones deben ser incluidas en el documento escrito, no representan páginas adicionales.