Simulación	ITAM
19 de febrero de 2018	Jorge de la Vega
2018-I	

Tarea 2. Fecha de entrega: Sábado 3 de marzo 2018

Lecturas

- Casella y Robert, capítulo 2
- Dagpunar, Capítulos 3 y 4. Secciones 7.1-7.3

Problemas

1. Proponer los algoritmos (método y pseudocódigo o código así como una corrida) para generar muestras de las siguientes densidades:

   (a) Cauchy
   $$f(x) = \frac{1}{\pi\beta\left[1+\left(\frac{x-\gamma}{\beta}\right)^2\right]}$$
   donde $\gamma, x \in \Re$, $\beta > 0$.

   (b) Gumbel (o de valor extremo)
   $$f(x) = \frac{1}{\beta}\exp\left(-e^{-(x-\gamma)/\beta} - \frac{x-\gamma}{\beta}\right)$$
   donde $\gamma, x \in \Re$, $\beta > 0$.

   (c) Logística
   $$f(x) = \frac{(1/\beta)e^{-(x-\gamma)/\beta}}{(1+e^{-(x-\gamma)/\beta})^2}$$
   donde $\gamma, x \in \Re$, $\beta > 0$.

   (d) Pareto
   $$f(x) = \frac{\alpha_2 c^{\alpha_2}}{x^{\alpha_2+1}}$$
   donde $c > 0, \alpha_2 > 0, x > c$.

   Para $\gamma = 0$ y $\beta = 1$ en cada inciso (a), (b) y (c), usen los algoritmos que obtuvieron para generar una muestra aleatoria de 5000 valores y obtengan $\bar{X}(n) = \sum_{i=1}^{n} X_i/n$ para $n = 50, 100, 150, \ldots, 5000$ para verificar empíricamente la ley fuerte de los grandes números. Hacer lo mismo para (d) con $c = 1$ y $\alpha_2 = 2$.

2. Para la distribución Pareto, crear funciones `rpareto`, `dpareto` `qpareto` y `ppareto` en R, para generar números aleatorios, obtener valores de la densidad, de la distribución y cuantíles de la familia Pareto de distribuciones. Mostrar su funcionalidad con ejemplos a través de una viñeta.

3. Una variable aleatoria discreta X tiene función de masa de probabilidad

x	0	1	2	3	4
$p(x)$	0.1	0.2	0.2	0.2	0.3

Utilicen el teorema de la transformación inversa para generar una muestra aleatoria de tamaño 1000 de la distribución de X. Construyan una tabla de frecuencias relativas y comparen las probabilidades empíricas con las teóricas.

Repitan considerando la función de `R sample`.

4. Graficar las siguientes densidades. Dar los algoritmos de transformación inversa, composición y aceptación-rechazo para cada una de las siguientes densidades. Discutir cuál algoritmo es preferible para cada densidad.

   (a)
$$f(x) = \frac{3x^2}{2} I(x)_{[-1,1]}$$

   (b) Para $0 < a < \frac{1}{2}$,
$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{a(1-a)} & 0 \leq x \leq a \\ \frac{1}{1-a} & a \leq x \leq 1-a \\ \frac{1-x}{a(1-a)} & 1-a \leq x \leq 1 \\ 0 & x \geq 1 \end{cases}$$

5. Considerando la transformación polar de Marsaglia para generar muestras de normales estándar, muestren que la probabilidad de aceptación de $S = V_1^2 + V_2^2$ en el paso 2 es $\pi/4$, y encuentren la distribución del número de rechazos de S antes de que ocurra una aceptación. ¿Cuál es el número esperado de ejecuciones del paso 1?

6. Obtengan una muestra de 10,000 números de la siguiente distribución discreta:
$$p(x) = \frac{2x}{k(k+1)}, x = 1, 2, \ldots, k$$
   para $k = 100$.

7. Desarrollen un algoritmo para generar una variable aleatoria binomial, usando la técnica de convolución (Hint: ¿cuál es la relación entre binomiales y Bernoullis?) Generar una muestra de 100,000 números. ¿Qué método es más eficiente, el de convoluciones o la función `rbinom` en R?

8. Probar que si X tiene función de distribución F, y si $h : \mathbb{R} \to B$ es una función estrictamente creciente donde $B \subseteq \mathbb{R}$, entonces $h(X)$ es una variable aleatoria con distribución $F(h^{-1}(x))$. Además, si F tiene densidad f y h^{-1} es absolutamente continua, entonces $h(X)$ tiene densidad $(h^{-1})'(x)f(h^{-1}(x))$ para casi toda x.

9. El kernel de Epanechnikov reescalado es una función de densidad simétrica:
$$f_e(x) = \frac{3}{4}(1 - x^2), \qquad |x| \leq 1.$$
   Consideren el siguiente algoritmo para generar muestras de esta distribución:

   (1) Generar uniformes iid $U_1, U_2, U_3 \sim \mathcal{U}(-1, 1)$.

   (2) Si $|U_3| \geq |U_2|$ y $|U_3| \geq |U_1|$, utilizar U_2. En otro caso, usar U_3.

   a. Escribir una función para generar variables aleatorias de f_e, y construir el histograma de una muestra de tamaño 1000.

b. Prueben que el algoritmo dado genera muestras de la densidad f_e.

10. Una compañía de seguros tiene 1000 asegurados, cada uno de los cuales presentará de manera independiente una reclamación en el siguiente mes con probabilidad $p = 0.09245$. Suponiendo que las cantidades de los reclamos hechos son variables aleatorias Gamma(7000, 1), hagan simulación para estimar la probabilidad de que la suma de los reclamos exceda $500,000.

11. Supongan que la densidad de una variable aleatoria X es $f(x) = \frac{1}{8} x I_{[0,4]}^{(x)}$. Supóngan que se toman muestras de tamaño 8 e interesa generar una muestra de 1000 observaciones del mínimo y del máximo de estas muestras. Obtener las 1000 muestras del minimo y el máximo y hacer los respectivos histogramas.

12. Obtener una muestra de tamaño 500 de la distribución conjunta de las siguientes tres variables aleatorias X_1, X_2, X_3 con distribución dada a continuación:

		$X_1 =$					
		0			1		
	$X_2 =$	1	2	3	1	2	3
	0	0.12	0.10	0.18	0.08	0.06	0.05
$X_3 =$	1	0.08	0.06	0.04	0.05	0.04	0.03
	2	0.06	0.04	0.02	0.04	0.03	0.02