Algoritmos de simulación

inm

Algunos recordatorios de las relaciones entre variables aleatorias.

Lema 1. *Las siguientes relaciones son verdaderas.*

1. *Si* $U \sim \mathrm{U}(0,1)$ *y* $\lambda > 0$ *entonces* $-\log(U)/\lambda \sim \mathrm{Exp}(\lambda)$.
2. *Si* $Z \sim \mathrm{N}(0,1)$ *entonces* $Z^2 \sim \chi_1^2 = \mathrm{Ga}(1/2, 1/2)$.
3. *Si* $Z_1, Z_2 \sim \mathrm{N}(0,1)$ *son independientes entonces* $Z_2/Z_2 \sim \mathrm{Cauchy}(0,1)$.

Demostración. Para establecer el inciso 1, consideremos $U \sim \mathrm{U}(0,1)$, denotemos $X = -\log(U)/\lambda$ y observemos que para $x < 0$, $\mathbb{P}(X \le x) = 0$, mientras que para $x \ge 0$,

$$\mathbb{P}(X \le x) = \mathbb{P}(\log(U) \ge -\lambda x) = \mathbb{P}(U \ge e^{-\lambda x}) = 1 - e^{-\lambda x}.$$

Para ver que el inciso 2 es verdadero supongamos que $Z \sim \mathrm{N}(0,1)$. Para $x < 0$, $\mathbb{P}(Z^2 \le x) = 0$; por otra parte, si $x \ge 0$ entonces

$$\mathbb{P}(Z^2 \le x) = \mathbb{P}(-\sqrt{x} \le Z \le \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}),$$

donde la última igualdad se da por la continuidad de Z y la definición de Φ. Así pues, cuando $x \ge 0$, derivando con respecto a x y usando la densidad de una normal estándar obtenemos que

$$f_{Z^2}(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{1}{2}x} = \frac{\left(\frac{1}{2}\right)^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} x^{\frac{1}{2}-1} e^{-\frac{1}{2}x},$$

donde hemos usado que $\Gamma(1/2) = \sqrt{\pi}$.

Para concluir, recordemos que si Y y W son variables aleatorias continuas e independientes, entonces

$$f_{W/Y}(x) = \int_{-\infty}^{\infty} |y|\, f_Y(y) f_W(wy)\,\mathrm{d}y.$$

En efecto, para $x \in \mathbb{R}$ tendremos que

$$\begin{aligned}
\mathbb{P}\left(\frac{W}{Y} \le x\right) &= \iint_{\{w/y \le x\}} f_Y(y) f_W(w)\,\mathrm{d}w\,\mathrm{d}y \\
&= \int_{-\infty}^{0}\int_{xy}^{\infty} f_Y(y) f_W(w)\,\mathrm{d}w\,\mathrm{d}y + \int_{0}^{\infty}\int_{-\infty}^{xy} f_Y(y) f_W(w)\,\mathrm{d}w\,\mathrm{d}y \\
&= \int_{-\infty}^{0}\int_{x}^{-\infty} y f_Y(y) f_W(zy)\,\mathrm{d}z\,\mathrm{d}y + \int_{0}^{\infty}\int_{-\infty}^{x} y f_Y(y) f_W(zy)\,\mathrm{d}z\,\mathrm{d}y \\
&= \int_{-\infty}^{x}\left(\int_{-\infty}^{0}(-y) f_Y(y) f_W(zy)\,\mathrm{d}y + \int_{0}^{\infty} y f_Y(y) f_W(zy)\,\mathrm{d}y\right)\mathrm{d}z \\
&= \int_{-\infty}^{x}\int_{-\infty}^{\infty} |y|\, f_Y(y) f_W zy\,\mathrm{d}y\,\mathrm{d}z,
\end{aligned}$$

donde la tercera igualdad se da por el cambio de variable $z = w/y$ y la cuarta igualdad es Fubini y un poco de álgebra. Así las cosas, si $C = Z_2/Z_1$, entonces para $x \in \mathbb{R}$,

$$f_C(x) = \int_{-\infty}^{\infty} |y| \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(xy)^2}{2}}\,\mathrm{d}y = \frac{1}{\pi}\int_0^{\infty} y e^{-\frac{(1+x^2)y^2}{2}}\,\mathrm{d}y = \frac{1}{\pi(1+x^2)},$$

que es la densidad de una variable aleatoria Cauchy(0, 1). ☹

En un curso de probabilidad suele verse el método de Box–Muller, el cual nos dice que si $U_1, U_2 \sim \mathrm{U}(0,1)$ son independientes entonces la transformación

$$\begin{aligned} Z_1 &= \sqrt{-2\log(U_1)}\cos(2\pi U_2), \\ Z_2 &= \sqrt{-2\log(U_1)}\operatorname{sen}(2\pi U_2), \end{aligned}$$

nos da un vector aleatorio (Z_1, Z_2) con $Z_1, Z_2 \sim \mathrm{N}(0,1)$ independientes. Analicemos estas expresiones detenidamente y observemos que

$$\begin{aligned} -2\log(U_1) &= Z_1^2 + Z_2^2, \\ \tan(2\pi U_2) &= Z_2/Z_1. \end{aligned}$$

Del recordatorio sabemos que $Z_i^2 \sim \mathrm{Ga}(1/2, 1/2)$, por lo que $Z_1^2 + Z_2^2 \sim \mathrm{Ga}(1, 1/2) = \exp(1/2)$, lo cual concuerda con que $-2\log(U_1) \sim \mathrm{Exp}(1/2)$. ¡Pero tenemos más! Observemos que $Z_2/Z_1 \sim$ Cauchy(0, 1), por lo que $\tan(2\pi U_2) \sim$ Cauchy(0, 1), es decir que hemos encontrado una forma de simular variables aleatorias Cauchy.

Ahora, podemos aprovechar el método de Box–Muller para obtener variables normales independientes no estándar sin morir en el intento. Pues si $\lambda_1, \lambda_2 > 0$ y $U_1, U_2 \sim \mathrm{U}(0,1)$ son independientes, entonces el vector aleatorio (X_1, X_2), definido por la transformación

$$\begin{aligned} X_1 &= \sqrt{-\frac{1}{\lambda_1}\log(U_1)}\cos(2\pi U_2), \\ X_2 &= \sqrt{-\frac{1}{\lambda_2}\log(U_1)}\operatorname{sen}(2\pi U_2), \end{aligned}$$

es tal que X_1 y X_2 son variables aleatorias independientes tales que $X_i \sim \mathrm{N}(0, (2\lambda_i)^{-1})$, puesto que se dará la igualdad $X_i = (2\lambda_i)^{-1/2} Z_i$.

Dejando de lado Box–Muller, algunas otras consecuencias del recordatorio, considerando una sucesión $\{U_i\}_{i\geq 1}$ de variables aleatorias independientes e idénticamente distribuidas U(0, 1), son las siguientes:

1. Si $\nu \geq 1$ es entero, $Y = -2\sum_{j=1}^{\nu}\log(U_j) \sim \chi^2_{2\nu} = \mathrm{Ga}((2\nu)/2, 1/2)$.
2. Si $\alpha \geq 1$ es entero y $\lambda > 0$, $Y = -\sum_{j=1}^{\alpha}\log(U_j)/\lambda \sim \mathrm{Ga}(\alpha, \lambda)$.
3. Si $\alpha, \beta \geq 1$ son enteras, entonces

$$Y = \frac{\sum_{j=1}^{\alpha}\log(U_j)}{\sum_{j=1}^{\alpha+\beta}\log(U_j)} \sim \mathrm{Beta}(\alpha, \beta).$$

Los primeros dos resultado son claros y se dejan al lector. Para ver que el tercer resultado es cierto, probaremos algo más general: para $\alpha, \beta > 0$, si $Y_1 \sim \mathrm{Ga}(\alpha, 1)$ y $Y_2 \sim \mathrm{Ga}(\beta, 1)$ son independientes, entonces

$Y_1/(Y_1+Y_2) \sim \mathrm{Beta}(\alpha,\beta)$ será independiente de $Y_1+Y_2 \sim \mathrm{Ga}(\alpha+\beta,1)$. Con este fin, definamos $W = Y_1/(Y_1+Y_2)$ y $S = Y_1 + Y_2$. Del teorema de cambio de variable,

$$\begin{aligned} f_{W,S}(w,s) &= s f_{Y_1}(ws) f_{Y_2}(s-ws) \\ &= s \times \frac{1}{\Gamma(\alpha)}(ws)^{\alpha-1}e^{-ws}\mathbb{1}_{\mathbb{R}_+}(ws) \times \frac{1}{\Gamma(\beta)}(s-ws)^{\beta-1}e^{s-ws}\mathbb{1}_{\mathbb{R}_+}(s-ws) \\ &= \left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}w^{\alpha-1}(1-w)^{\beta-1}\mathbb{1}_{(0,1)}(w)\right) \times \left(\frac{1}{\Gamma(\alpha+\beta)}s^{(\alpha+\beta)-1}e^{-s}\mathbb{1}_{\mathbb{R}_+}(s)\right). \end{aligned}$$

Que el inciso 3 es cierto es ahora inmediato. Para finalizar veremos un algoritmo en sumo útil para generar variables aleatorias continuas con solo conocer un múltiplo de la función de densidad: el *método de aceptación y rechazo*. Supongamos que f es una función de densidad y que g es otra función de densidad tal que $f \leq Mg$ para alguna $M > 0$. El algoritmo consiste en los siguientes pasos:

1. Generamos una variable aleatoria X de la función de densidad g y una variable aleatoria $U \sim \mathrm{U}(0,1)$.
2. Si $U \leq f(X)/(Mg(X))$, entonces definimos $Y = X$ y terminamos.
3. En otro regresamos a 1.

Proposición 2. *La variable aleatoria Y generada mediante el algoritmo de aceptación y rechazo tiene densidad f. Más aún, la cantidad de simulaciones que haremos para aceptar un valor X es una variable aleatoria geométrica con parámetro $1/M$.*

Demostración. Obtengamos la función de distribución de Y.

$$\mathbb{P}(Y \leq y) = \mathbb{P}\left(X \leq y \,\middle|\, U \leq \frac{f(X)}{Mg(X)}\right) = \frac{\mathbb{P}\left(X \leq y, U \leq \frac{f(X)}{Mg(X)}\right)}{\mathbb{P}\left(U \leq \frac{f(X)}{Mg(X)}\right)}.$$

Para el numerador tendremos, por independencia,

$$\begin{aligned} \mathbb{P}\left(X \leq y, U \leq \frac{f(X)}{Mg(X)}\right) &= \int_{-\infty}^{y} \mathbb{P}\left(U \leq \frac{f(x)}{Mg(x)}\right) g(x)\,\mathrm{d}x = \int_{-\infty}^{y} g(x) \int_0^{f(x)/Mg(x)} \mathrm{d}u\,\mathrm{d}x \\ &= \int_{-\infty}^{y} g(x)\frac{f(x)}{Mg(x)}\,\mathrm{d}x = \frac{1}{M}\int_{-\infty}^{y} f(x)\,\mathrm{d}x, \end{aligned}$$

mientras que el denominador estará dado por

$$\mathbb{P}\left(U \leq \frac{f(X)}{Mg(X)}\right) = \lim_{y\to\infty} \frac{1}{M}\int_{-\infty}^{y} f(x)\,\mathrm{d}x = \frac{1}{M}\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \frac{1}{M}. \tag{1}$$

Así las cosas, $\mathbb{P}(Y \leq y) = \int_{-\infty}^{y} f(x)\,\mathrm{d}x$ y por lo tanto $f_Y(y) = f(y)$. Más aún, de (1) vemos que la probabilidad de aceptar una simulación X como simulación de Y es $1/M$, de donde se sigue fácilmente el resultado. ☹

Una consecuencia importante de la demostración de la proposición 2 es que en realidad no es necesario conocer f en su totalidad, basta conocer un múltiplo de la función de densidad. Esto es útil en estadística bayesiana, pues para un valor de x fijo se tiene la relación de proporcionalidad

$$f_{\Theta\,|\,X}(\theta\,|\,x) \propto f_{X\,|\,\Theta}(x\,|\,\theta) f_\Theta(\theta) = f_{X,\Theta}(x,\theta),$$

por lo cual para realizar simulaciones de la *distribución posterior* de Θ, basta conocer la densidad conjunta y no hará falta normalizar. El costo a pagar será encontrar una función de densidad g y una $M > 0$ tales que $f_{X,\Theta}(x,\theta) \leq Mg(\theta)$ para todo θ.

Ejemplo 1 (Simulación de gaussianas estándar de una Cauchy)**.** Cuando $f(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ y $g(x) = (\pi(1+x^2))^{-1}$,

$$\frac{f(x)}{g(x)} = \sqrt{\frac{\pi}{2}}(1+x^2)e^{-\frac{x^2}{2}}$$

está acotado por $M = \sqrt{2\pi/e}$, valor que se alcanza en $x = \pm 1$, por lo que la probabilidad de aceptación es $1/M = \sqrt{e/2\pi} \approx 0.66$, lo cual implica que en promedio, una de cada tres simulaciones de la Cauchy es rechazada.

Ejemplo 2 (Generación de variables aleatorias gamma)**.** Supongamos que $\alpha \geq 1$ y, sin pérdida de generalidad, supongamos que $\lambda = 1$. Entonces $f(x) = x^{\alpha-1}e^{-x}/\Gamma(\alpha)$. Proponemos $g_b(x) = b^a x^{a-1}e^{-bx}/\Gamma(a)$, con $a = \lfloor \alpha \rfloor$ y $b < 1$. Entonces

$$\frac{f(x)}{g_b(x)} = \frac{\Gamma(a)}{\Gamma(\alpha)}b^{-a}x^{\alpha-a}e^{-(1-b)x} \leq \frac{\Gamma(a)}{\Gamma(\alpha)}b^{-a}\left(\frac{\alpha-a}{(1-b)e}\right)^{\alpha-a},$$

y vemos que la cota mínima se alcanza con $b = a/\alpha$.