Muestrador de Gibbs

Diego Uriarte

24/9/2018

¿Qué es el muestrador de Gibbs?

- Permite muestrar de una distribución sin tener que calcular la densidad.
- Queremos muestrar de f(x)

$$f(x) = \int \cdots \int f(x, y_1, \cdots, y_p) dy_1 \cdots dy_p$$

- ¿Qué podemos hacer?
- Resolver analítica o numéricamente, pero ¿si no se puede?

¿Cómo funciona el muestrador de Gibbs?

- Ilustraremos con un ejemplo bivariado. Queremos muestras de $f(x)$ (distribuida de manera conjunta con y). Conocemos las condicionales $f(x|y)$ y $f(y|x)$.
- Se especifica un valor inicial para $Y_0' = y_0'$.
- Se obtiene apartir de las condicionales una sucesión de muestras para x e y:

$$X_j' \sim f(x|Y_j' = y_j')$$

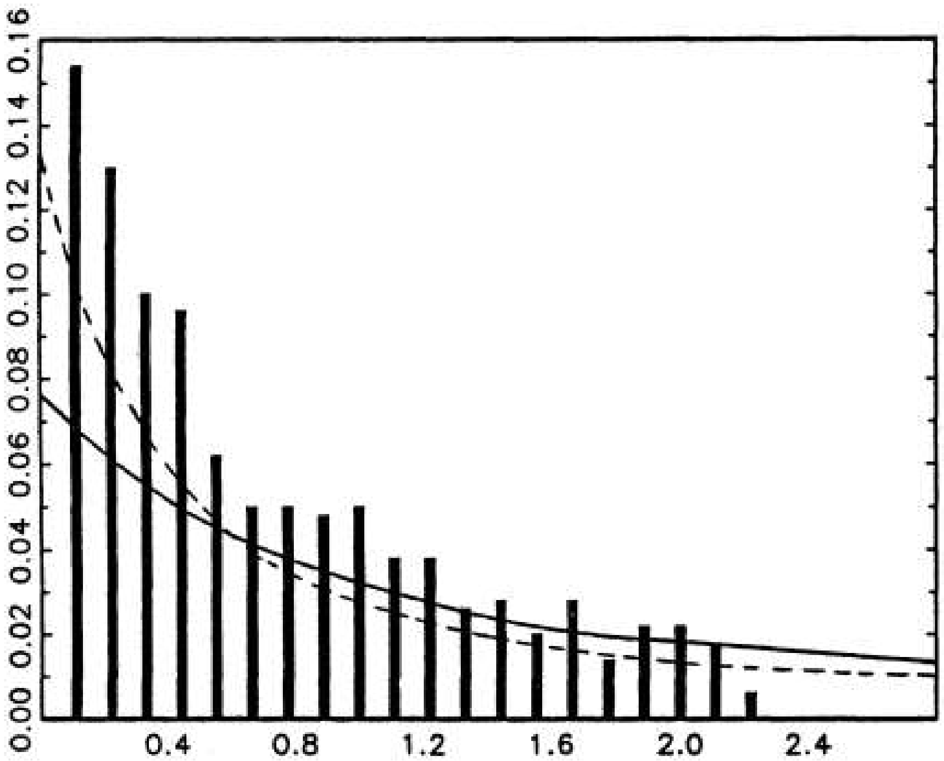
$$Y_{j+1}' \sim f(y|X_j' = x_j')$$

- Obtenemos así una secuencia: $Y_0', X_0', Y_1', X_1', \cdots Y_k', X_k',$
- Resulta que bajo algunas condiciones generales, cuando $k \to \infty$, X_k' es efectivamente una muestra de $f(x)$

Aplicación

Condicionales exponenciales truncas

$$f(x|y) = \frac{ye^{-yx}}{1-e^{-By}}, 0 < x < B < \infty$$
$$f(y|x) = \frac{xe^{-xy}}{1-e^{-By}}, 0 < y < B < \infty$$



Prueba simple de la convergencia del muestrador de Gibbs - Caso Discreto

Mostraremos como converge el muestreo tomando como ejemplo una distribución bivariada, donde X e Y se distribuyen Bernoulli de manera conjunta.

		X = 0	X = 1
Y	0	p_1	p_2
	1	p_3	p_4

$$p_i \geq 0,\ p_1 + p_2 + p_3 + p_4 = 1,$$

La distribución conjunta de X, Y está dada por:

$$\begin{bmatrix} f_{x,y}(0,0) & f_{x,y}(1,0) \\ f_{x,y}(0,1) & f_{x,y}(1,1) \end{bmatrix} = \begin{bmatrix} p_1 & p_2 \\ p_3 & p_4 \end{bmatrix}$$

La marginal de X es:

$$f_x = \begin{bmatrix} p_1 + p_3 & p_2 + p_4 \end{bmatrix}$$

Determinamos las probabilidades condicionales $X|Y = y$ e $Y|X = x$, formulando las siguientes matrices:

$$A_{y|x} = \begin{bmatrix} Prob(Y=0|X=0) & Prob(Y=1|X=0) \\ Prob(Y=0|X=1) & Prob(Y=1|X=1) \end{bmatrix} = \begin{bmatrix} \frac{p_1}{p_1+p_3} & \frac{p_3}{p_1+p_3} \\ \frac{p_2}{p_2+p_4} & \frac{p_4}{p_2+p_4} \end{bmatrix}$$

$$A_{x|y} = \begin{bmatrix} Prob(X=0|Y=0) & Prob(X=1|Y=0) \\ Prob(X=0|Y=1) & Prob(X=1|Y=1) \end{bmatrix} = \begin{bmatrix} \frac{p_1}{p_1+p_2} & \frac{p_2}{p_1+p_2} \\ \frac{p_3}{p_3+p_4} & \frac{p_4}{p_3+p_4} \end{bmatrix}$$

- Queremos muestrar de la marginal de X ($X'_0, X'_1, X'_2, \cdots$).
- Para ir de $X'_0 \rightarrow X'_1$ debemos muestrar Y'_1, es decir, $X'_0 \rightarrow Y'_1 \rightarrow X'_1$.

$$P(X'_1 = x_1 | X'_0 = x_0) = \sum_y P(X'_1 = x_1 | Y'_1 = y) \times P(Y'_1 = y | X'_0 = x_0)$$

Para este caso en particular, se tiene que la matrices de transición para x está dada por:

$$A_{x|x} = A_{y|x} A_{x|y}$$

Ahora, es directo calcular la distribución marginal de realización k, X'_k:

$$f_k = f_0 A_{x|x}^k = (f_0 A_{x|x}^{k-1}) A_{x|x} = f_{k-1} A_{x|x}$$

Bajo ciertas condiciones, se tiene un punto fijo tal que:

$$f = f A_{x|x}$$

En nuestro caso simple:

$$\begin{aligned}
f_x A_{x|x} &= f_x A_{y|x} A_{x|y} \\
&= \begin{bmatrix} p_1 + p_3 & p_2 + p_4 \end{bmatrix} \begin{bmatrix} \frac{p_1}{p_1+p_3} & \frac{p_3}{p_1+p_3} \\ \frac{p_2}{p_2+p_4} & \frac{p_4}{p_2+p_4} \end{bmatrix} \begin{bmatrix} \frac{p_1}{p_1+p_2} & \frac{p_2}{p_1+p_2} \\ \frac{p_3}{p_3+p_4} & \frac{p_4}{p_3+p_4} \end{bmatrix} \\
&= \begin{bmatrix} p_1 + p_2 & p_3 + p_4 \end{bmatrix} \begin{bmatrix} \frac{p_1}{p_1+p_2} & \frac{p_2}{p_1+p_2} \\ \frac{p_3}{p_3+p_4} & \frac{p_4}{p_3+p_4} \end{bmatrix} \\
&= \begin{bmatrix} p_1 + p_3 & p_2 + p_4 \end{bmatrix} \\
&= f_x
\end{aligned}$$

Ejemplo numérico

Tomemos $p_1 = 0.1, p_2 = 0.4, p_3 = 0.15$ y $p_4 = 0.35$. Por tanto:

$$f_x = [\, p_1 + p_3 \quad p_2 + p_4 \,] = [\, 0.25 \quad 0.75 \,]$$

Ahora, partamos de $f_0 = [0.99 \quad 0.01]$, es decir, alejados del valor de la marginal de x. Se tiene:

$$\begin{aligned}
f_1 &= f_0 \times A_{xx} = [0.99 \quad 0.01] \times \begin{bmatrix} 0.26 & 0.74 \\ 0.24667 & 0.7533 \end{bmatrix} \\
f_1 &= [0.2598667 \quad 0.7498684] \\
f_2 &= [0.2501316 \quad 0.7499982] \\
f_3 &= [0.2500018 \quad 0.7499982] \\
f_4 &= [0.25 \quad 0.75]
\end{aligned}$$

Intuición para el caso bivariado continuo

¿Cómo es que podemos llegar a partir de las condicionales a la marginal?

- Conocemos $f_{X|Y}(x|y)$ y $f_{Y|X}(y|x)$. Si queremos la marginal, por definición:

$$f_X(x) = \int f_{xy}(x,y)dy$$

- Por Bayes sabemos que: $f_{xy}(x,y) = f_{X|Y}(x|y)f_Y(y)$, por tanto:

$$f_X(x) = \int f_{X|Y}(x|y)f_Y(y)dy$$

De manera similar, tenemos $f_{xy}(x,y) = f_{Y|X}(y|x)f_X(y)$ con:

$$f_Y(y) = \int f_{Y|X}(y|x)f_X(x)dx$$

Reemplazando:

$$
\begin{aligned}
f_X(x) &= \int f_{X|Y}(x|y) \int f_{Y|X}(y|t) f_X(t)\,dt\,dy \\
&= \int \left[\int f_{X|Y}(x|y) f_{Y|X}(y|t)\,dy \right] f_X(t)\,dt \\
&= \int h(x,t) f_X(t)\,dt
\end{aligned}
$$

con $h(x,t) = \left[\int f_{X|Y}(x|y) f_{Y|X}(y|t)\,dy \right]$.

- Llegamos a lo que se conoce como una ecuación de punto fijo integral.
- La secuencia de Gibbs es una manera de converger a la solución de la ecuación.

¿Gibbs siempre converge a la solución?

Veamos el primer ejemplo, pero sin acotar las distribuciones condicionales.

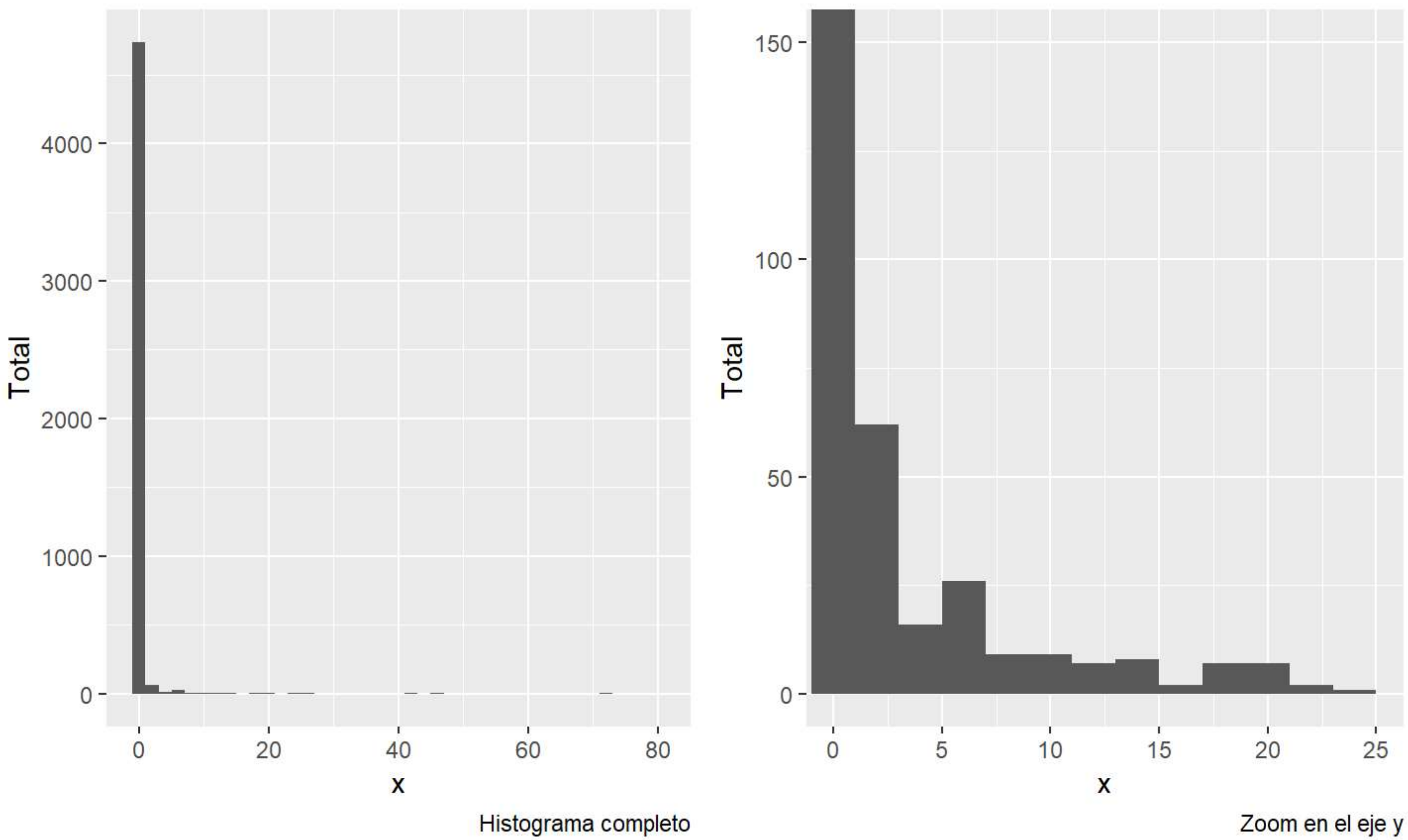
$$f(x|y) = ye^{-yx}, 0 < x < \infty$$
$$f(y|x) = xe^{-xy}, 0 < y < \infty$$

$$\begin{aligned} f_X(x) &= \int_0^\infty \left[\int_0^\infty ye^{-yx}te^{-ty}dy\right] f_X(t)dt \\ &= \int_0^\infty \left[\frac{t}{(x+t)^2}\right] f_X(t)dt \end{aligned}$$

Y notamos que $f_X(t) = 1/t$ es solución puesto que:

$$\begin{aligned}
\frac{1}{x} &= \int_0^\infty \left[\frac{t}{(x+t)^2} \right] \frac{1}{t} dt \\
&= \int_0^\infty \left[\frac{1}{(x+t)^2} \right] dt \\
&= \left[\frac{-1}{(x+t)} \right]_{t=0}^{t=\infty} \\
&= 0 - \frac{-1}{x} = \frac{1}{x}
\end{aligned}$$

Pero cuando intentamos aplicar el muestrador de Gibbs a estas condicionales, no obtenemos convergencia:



- Ahora el procedimiento de Gibbs no converge (no obtenemos una aproximación de $1/x$).
- Podemos evitar esto si truncamos las distribuciones condicionales de manera que se cumpla:

$$\int_0^B f_X(x)dx < \infty$$

- La marginal ya no integra a un número finito:

$$\int_0^\infty \frac{1}{x} dx \to \infty$$

Conclusión

Una condición para que algoritmo converga es que la marginal sea una función de densidad $\int f_X(x)dx < \infty$. Una manera de segurar esto es acotando las condicionales.

Muchas Gracias!