

ESMA 5015: Accept Reject

Alejandro Ouslan

Spring 2025

Contents

1	notes for 2025-03-06	1
2	Consideren un algoritmo Accept-Reject para simular $X \sim N(0, 1)$	2
2.1	theory	2
2.2	algorithm	3
2.3	steps	3
2.4	algorithm	3
3	Generar una $X \sim Beta(\alpha, \beta)$	3
3.1	algorithm	4
4	Ejemplo generar $X \sim Beta(\alpha = 2.7, \beta = 6.3)$	4
4.1	algorithm	4
5	Bayesian Inference	4
5.1	Ejemplo	4
5.1.1	Frecuentista	5
5.1.2	Bayesiana	5
5.1.3	Calculation	5
5.2	Assignment	5
6	Markor Chain Monte Calro (MCMC)	5

1 notes for 2025-03-06

$$E(w) = \alpha\beta$$
$$w \sim \text{gamma}(\alpha, \beta) \quad \text{var}(w) = \alpha\beta^2$$

$$F_y(g) = \frac{1}{2}e^{-y} + \frac{1}{2}ye^{-y}$$
$$E[y] = \int_0^\infty yf_y(y)dy$$
$$= \frac{1}{2}E[y_1] + \frac{1}{2}E[y_2]$$
$$y_1 \sim \text{gamma}(1, 1)$$
$$y_2 \sim \text{gamma}(2, 1)$$
$$= \frac{1}{2}1 + \frac{1}{2}2$$
$$= \frac{3}{2}$$

$$\begin{aligned}
\text{var}(y) &= E[y^2] - E[y]^2 \\
E[y^2] &= \int_0^\infty y^2 f_y(y) dy \\
&= .5E[y_1^2] + .5E[y_2^2] \\
&= \frac{1}{2}2 + \frac{1}{2}6 \\
&= 4
\end{aligned}$$

$$\begin{aligned}
E[y_1^2] &= \text{var}(y_1) + E[y_1]^2 \\
&= 1 + 1^2 \\
&= 2
\end{aligned}$$

$$\begin{aligned}
E[y_2^2] &= \text{var}(y_2) + E[y_2]^2 \\
&= 2 + 2^2 \\
&= 6
\end{aligned}$$

$$\begin{aligned}
&= 4 - \frac{3^2}{2} \\
&= 4 - \frac{9}{2} \\
&= \frac{7}{2}
\end{aligned}$$

2 Consideren un algoritmo Accept-Reject para simular $X \sim N(0, 1)$

1. usando $Y \sim \text{cauchy}(0, 1)$

$$F_Y(y) = \frac{1}{\pi} \frac{1}{1+y^2} \quad y \in \mathbb{R}$$

2. usando $Y \sim \text{double-exponential}(0, 1)$

$$f_Y(y) = \frac{1}{2} e^{-|y|} \quad y \in \mathbb{R}$$

2.1 theory

1. calculate $M = \sup_x \frac{f_X(x)}{f_Y(y)}$

$$\begin{aligned}
&= \sup_x \frac{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}{\frac{1}{\pi} \frac{1}{1+x^2}} \\
&= \sup_x \frac{\pi}{\sqrt{2}} e^{-x^2/2} (1+x^2)
\end{aligned}$$

sea $h(x) = e^{-x^2/2}(1+x^2)$

$$\begin{aligned}
\frac{dh(x)}{dx} &= -xe^{-x^2/2}(1+x^2) + e^{-x^2/2}2x \\
&= \vdots \\
&= -xe^{-x^2/2}(x^2-1)
\end{aligned}$$

$$\frac{dh(x)}{dx} = 0 \Rightarrow x = 0, \pm 1 \text{ puede demostrar que el máximo ocurre en } x = \pm 1 \Rightarrow M = \frac{\pi\sqrt{\pi}}{\sqrt{2}} e^{-1/2} = 1.52$$

2.2 algorithm

1. simular $Y \sim \text{cauchy}(0, 1)$ y $U \sim U(0, 1)$ independientes.
2. si

$$\begin{aligned} U &< \frac{f_X(Y)}{M f_Y(Y)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-1/2} \frac{\sqrt{\pi}}{\sqrt{\pi}} (1 + Y^2) \\ &= \frac{1}{2} (1 + Y^2) e^{\frac{-Y^2+1}{2}} \end{aligned}$$

aceptar $X = Y$

3. si no, regresar a paso 1.

2.3 steps

- 1.

$$\begin{aligned} M &= \sup_x \frac{f_X(x)}{f_Y(y)} = \sup_x \frac{\int \frac{1}{\sqrt{2\pi}} e^{-x^2/2}}{\frac{1}{2} e^{-|x|}} \\ &= \sup_x \frac{\sqrt{2}}{\sqrt{\pi}} e^{-x^2/2+|x|} \\ h(x) &= \ln \left(e^{-x^2/2+|x|} \right) \\ \frac{dh(x)}{dx} &= \frac{d}{dx} \left(-\frac{x^2}{2} + |x| \right) = \begin{cases} -x - 1 & x < 0 \\ -x + 1 & x > 0 \end{cases} \\ \frac{dh(x)}{dx} &= 0 \Rightarrow x = \pm 1 \end{aligned}$$

debe demostrar que el supremo ocurre en $x = \pm 1$

2.4 algorithm

1. $M = \frac{\sqrt{2}}{\sqrt{\pi}} e^{-1/2+1} = 1.31$
2. simular $Y \sim \text{double-exponential}(0, 1)$ y $U \sim U(0, 1)$ independientes.
3. si

$$\begin{aligned} U &< \frac{f_X(Y)}{M f_Y(Y)} \\ &= e^{\frac{-Y^2}{2} + |Y| - 1/2} \end{aligned}$$

entonces aceptar $X = Y$

4. si no, regresar a paso 2.

3 Generar una $X \sim \text{Beta}(\alpha, \beta)$

Donde $0 < x < 1$ y $\alpha, \beta > 0$

Utilizar $y \sim U(0, 1)$.

$$M = \sup_x \frac{f_X(x)}{f_Y(y)} = \frac{\frac{\gamma(\alpha+\beta)}{\gamma(\alpha)\gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}}{1}$$

$$h(x) = \ln (X^{\alpha-1} (1-x)^{\beta-1})$$

$$\frac{dh(x)}{dx} = \frac{\alpha-1}{x} - \frac{\beta-1}{1-x}$$

$$\frac{dh(x)}{dx} = 0 \Rightarrow x = \frac{\alpha-1}{\alpha+\beta-2} \text{ y } \beta! = 1, \alpha! = 1$$

puede demostrar que en esta X OCURRE UN maximo

3.1 algorithm

1. Generar $y \sim U(0, 1)$ y $U \sim U(0, 1)$ independientes.
2. Si

$$u < \frac{f_X(y)}{M f_Y(y)} = \frac{\frac{\gamma(\alpha+\beta)}{\gamma(\alpha)\gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1}}{\frac{\gamma(\alpha+\beta)}{\gamma(\alpha)\gamma(\beta)} \frac{\alpha-1}{\alpha+\beta-2} y^{\alpha-1} (1-\frac{\alpha-1}{\alpha+\beta-2})^{\beta-1}}$$

define $X = y$

3. si no, regresar a paso 1.

4 Ejemplo generar $X \sim \text{Beta}(\alpha = 2.7, \beta = 6.3)$

$M = 2.6667444$ que ocurre en $x = \frac{2.7-1}{2.7+6.3-2} = 0.2428$

4.1 algorithm

1. Generar $y \sim U(0, 1)$ y $U \sim U(0, 1)$ independientes.
2. Si

$$u < \frac{f_X(y)}{M f_Y(y)} = \frac{\frac{\gamma(\alpha+\beta)}{\gamma(\alpha)\gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1}}{2.666744}$$

Define $X = y$

3. si no, regresar a paso 1.

5 Bayesian Inference

Accept y Metropolis Hasting (topico de la proxima clase) surgen naturlament en estadisticas Bayesianas. En el Analisis Bayesiano ademas de espexifizar el modelo de los datos observados $X = x_1, x_2, \dots, x_n$ dado un vector de parametros desconocidos θ definido por $f(x|\theta)$, se define θ como una variable aleatoria que tiene una distribucion prior $\pi(\theta)$. El Conocimiento de θ se actualiza con el conocimiento que se obtiene de los datos basados en la inferencia con respecto a θ . en la distribucion posterior definida por

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int_{\Theta} f(x|\theta)\pi(\theta)d\theta} \text{ Teorema de Bayes}$$

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta)$$

5.1 Ejemplo

x_1, x_2, \dots, x_n iid $\text{Bernulli}(\theta)$

5.1.1 Frecuentista

IC de $(1 - \alpha)100\%$ para θ

$$\hat{\theta} \pm z_{\alpha/2} \sqrt{\frac{\hat{\theta}(1 - \hat{\theta})}{n}}$$

5.1.2 Bayesiana

$$f(x_1, x_2, \dots, x_n | \theta) = \prod_{i=1}^n \theta^{x_i} (1 - \theta)^{1-x_i} = \theta^{\sum x_i} (1 - \theta)^{n - \sum x_i}$$
$$\pi(\theta | x_1, x_2, \dots, x_n) = \frac{\theta^{\sum x_i} (1 - \theta)^{n - \sum x_i} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}}{\int_0^1 \theta^{\sum x_i} (1 - \theta)^{n - \sum x_i} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1} d\theta}$$

En este caso la distribución a priori es conjugada con $f(x|\theta)$ porque $\pi(\theta|x)$ pertenece a la misma familia de la distribución pariori.

5.1.3 Calculation

$$E[\theta | x_1, x_2, \dots, x_n] = \frac{\sum x_i + \alpha}{\sum x_i + \alpha + n - \sum x_i + \beta}$$
$$= \frac{\sum x_i + \alpha}{n + \alpha + \beta}$$

Si n es mucho mayor que $\alpha + \beta$ el promedio posterior se inclina hacia el promedio muestral

5.2 Assignment

Considere $X_1, X_2, \dots, X_n; \sigma^2$ conocido y la distribución a priori

$$\pi(\theta) = N(\mu_0, \tau^2)$$

Demuestre que

$$\pi(\theta | x_1, x_2, \dots, x_n) \propto N\left(\frac{n\hat{x} + \frac{\mu}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}, \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}\right)$$

6 Markor Chain Monte Calro (MCMC)

Definition 1. Un metodo MCMC para simular de una distribución F es cualquier metodo que produzca una cadena de MARKOV erogodica X_N cuya distribución estacionari es F

Elejemplos dde estos metodos:

1. Metropolis-Hastings
2. Gibbs Sampling
3. Data Augmentation Algorithm