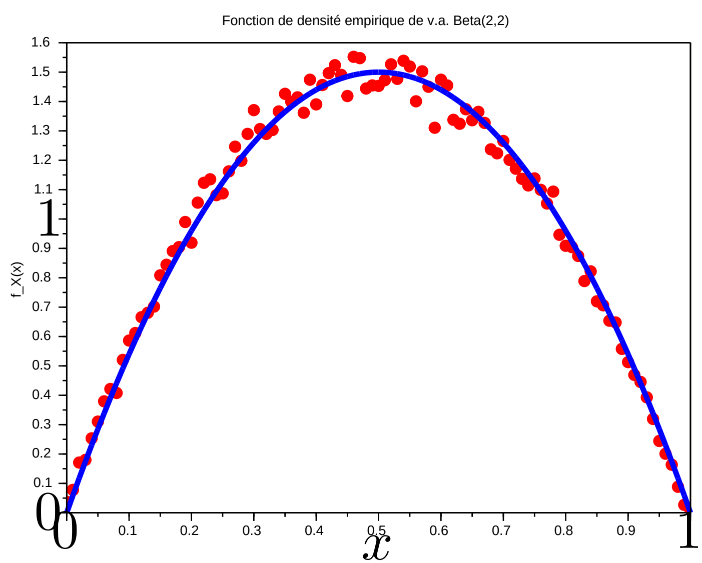
Markov Chain Monte Carlo

TP 4: Méthode de Rejet. Version 19-20.



Théorie

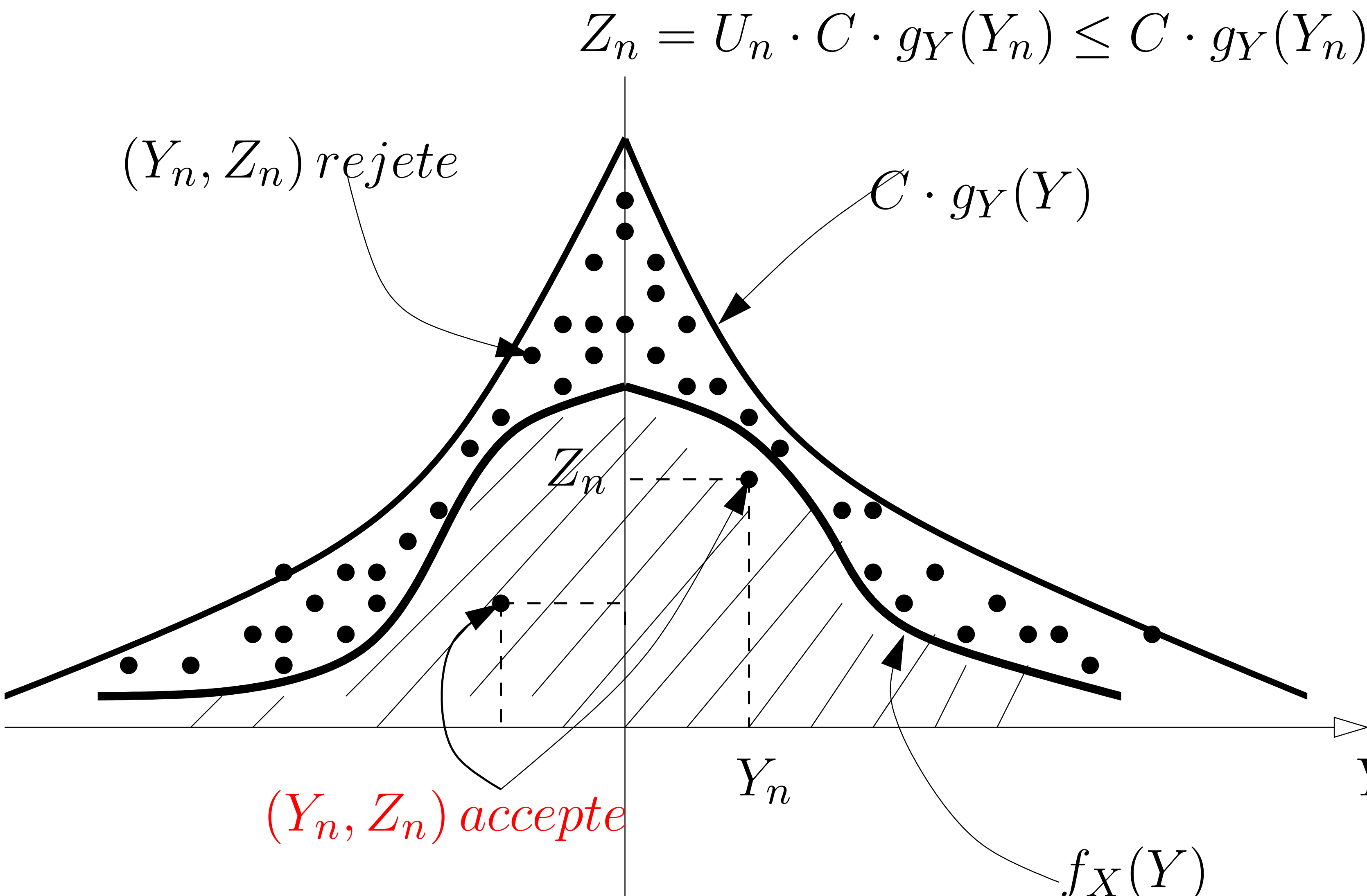
- Idées de Méthode de Rejet
  - • On sait simuler la v.a. $Y \Rightarrow Y = \{Y_1, Y_2, Y_3, ..., Y_{mc}\}$
  - • On connaît la fonction de densité $g_Y(y)$ de v.a. Y
  - • On connaît la fonction de densité $f_X(x)$ de v.a. X
  - • On utilise quelques réalisations de Y_i pour former X.

- Théorème 3
  - • X est la variable aléatoire à valeurs dans R avec la fonction de densité f_X
  - • Y est la variable aléatoire à valeurs dans R avec la fonction de densité g_Y
  - • Il existe une constante $C(\geq 1)$ satisfaisant

$$C \cdot g_Y(x) \geq f_X(x) \quad x \in \mathbb{R}$$

  - • U une variable aléatoire de loi uniforme sur [0, 1] indépendante de Y .
  - • Alors la loi de Y sachant $\{U < \frac{f_X(Y)}{C \cdot g_Y(Y)}\}$ est la loi de X.

Visualisation de Rejet



$Z_n = U_n \cdot C \cdot g_Y(Y_n) \leq f_X(Y_n) \Rightarrow$ on accepte Y_n et $X = Y_n$

Théorie

- Il faut montrer que la fonction de densité de v.a. X est bien $f_X(x)$:

$$\mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(x)dx.$$

Plus précisément
en posant l'évènement

$$A = \{U \cdot C \cdot g_Y(Y) < f_X(Y)\}$$

il faut montrer que

$$\mathbb{P}(Y \leq x | A) = \int_{-\infty}^{x} f_X(x)dx.$$

Théorie

- En posant l'évènement

$$A = \{U \cdot C \cdot g_Y(Y) < f_X(Y)\}$$

il faut montrer que

$$\mathbb{P}(Y \leq x | A) = \int_{-\infty}^{x} f_X(x)dx.$$

$\bullet$ Partons de la définition de la probabilité conditionnelle:

$$\mathbb{P}(Y \leq x | A) = \frac{\mathbb{P}((Y \in A) \bigcap A)}{\mathbb{P}(A)}$$

Demonstration

- $$\color{red}{A = \{U < f_X(\color{blue}{Y}\color{red}{)/(C \cdot g_Y(}\color{blue}{Y}\color{red}{))\}}}$$

  $$\mathbb{P}((Y \leq x) \bigcap A) =$$

  $$\int_0^1 \int_{-\infty}^{+\infty} 1_{\{(y \leq x),\; u < f_X(\color{blue}{y}\color{black}{)/(C \cdot g_Y(}\color{blue}{y}\color{black}{))\}}} \quad f_U(u) \cdot g_Y(y) du dy =$$

  $$\int_0^{\frac{f_X(Y)}{C \cdot g_Y(Y)}} \int_{-\infty}^{x} g_Y(y) dy du =$$

- On utilise l'independence des variables aléatoires U et Y

  $$f_{U,Y}(u, y) = f_U(u) \cdot g_Y(y)$$

Demonstration

- $$\int_{-\infty}^{x} g_Y(y)dy \int_0^{\frac{f_X(y)}{Cg_Y(y)}} du = \int_{-\infty}^{x} \frac{f_X(y)}{Cg_Y(y)} g_Y(y)dy =$$

$$\int_{-\infty}^{x} \frac{1}{C} f_X(y)dy = \frac{F_X(x)}{C}.$$

Demonstration

Si $(x = +\infty)$ on peut calculer en répétant le raisonnement précédant que

$$\mathbb{P}(A) = \mathbb{P}(U < \frac{f_X(Y)}{C g_Y(Y)}) = \int_{-\infty}^{+\infty} \frac{1}{C} f_X(y) dy = \frac{1}{C}$$

Donc

$$\mathbb{P}(Y \leq x | A) = \int_{-\infty}^{x} f_X(x) dx = F_X(x)$$

ce qu'il fallait démontrer.

Algorithme de Rejet

- Quelles realisations de Y choisir, les quelles rejeter?
  - function[X]=$Reject()$
  - ∘ $k = 1$;
  - ∘ for n=1: N_{mc}
  - ∘ Simuler v.a. Y
  - ∘ Simuler $U[0,1]$
  - ∘ if $U \leq \frac{f_X(Y)}{C \cdot g_Y(Y)}$
  - ∘ $X(k) = Y$; $\quad k = k + 1;$
  - ∘ endif
  - ∘ endfor
  - endfunction
- La constante C vérifie la condition:

$$\forall x \in [a,b], \quad \frac{f_X(x)}{g_Y(x)} \leq C$$

Simulation de v.a. Normale par Rejet

- On veut simuler v.a. Normale de densité $f_X(y) = \frac{e^{-y^2/2}}{\sqrt{2\pi}}$
- On sait simuler le loi de Laplace de densité $g_Y(y) = \frac{1}{2} e^{-|y|}$
- On calcule la constante $C = \sqrt{\frac{2e}{\pi}}$:

$$\frac{f_X(y)}{g_Y(y)} \leq \sqrt{\frac{2e}{\pi}} \quad \forall y \in \mathbb{R}$$

Simulation de v.a. Laplace

- V. a. Laplace est définie par sa fonction de densité ou de repartition:

$$g_Y(y) = \begin{cases} \frac{1}{2}e^{-y}, & y \geq 0 \\ \frac{1}{2}e^{y}, & y < 0 \end{cases} \qquad F_Y(y) = \begin{cases} 1 - \frac{1}{2}e^{-y} & y \geq 0 \\ \frac{1}{2}e^{y} & y < 0 \end{cases}$$

- Inversion de $F_Y(y)$:

$$\begin{cases} u < 1/2 \\ y = ln(2u) \end{cases} \qquad \begin{cases} u \geq 1/2 \\ y = -ln(2(1-u)) \end{cases}$$

Simulation de v.a. Laplace

- Simulation par l'inversion:

$$
\begin{cases}
U = rand() \\
if \quad U < 1/2 \\
Y = ln(2U) \\
else \\
Y = -ln(2(1-U))
\end{cases}
$$

- On remarque que $2U \in [0,1]$ et $Y = ln(2U)$ est une v.a. exponentielle multipliée par (-1), on remarque aussi que $2(1-U) \in [0,1]$, $\quad Y = -ln(2(1-U))$ suit la loi exponentielle. Donc on peut simuler Y par l'algorithme suivant

$$
\begin{cases}
if \quad rand() < 1/2 \\
Y = -V.A.exponentielle \quad (Y = ln(rand())) \\
else \\
Y = V.A.exponetielle \quad (Y = -ln(rand()))
\end{cases}
$$

Simulation 1 de v.a. Normale par Rejet

- Simulation d'un échantillon de v.a. Normale $X \in \mathbb{R}$ à l'aide de v.a. Laplace Y.
  - function[X]=Rejet_Normale_1( )
  - ∘ $C = \sqrt{\frac{2e}{\pi}}$
  - ∘ $k = 1;$
  - ∘ for n=1: N_{mc}
  - ∘ U=rand()
  - ∘ if $U < 1/2$

    $Y = -ln(2U)$

    else $Y = -ln(2(1-U))$
  - ∘ endif
  - ∘ Simuler U[0,1] (indépendante)
  - ∘ $f = (\frac{e^{-Y^2/2}}{\sqrt{2\pi}}); \quad g = \frac{1}{2}e^{-|Y|}$
  - ∘ if $U \leq \frac{f}{(C \cdot g)}$
  - ∘ $X(k) = Y; \quad k = k + 1;$
  - ∘ endif
  - ∘ endfor
  - endfunction

Simulation 2 de v.a. Normale par Rejet

- Simulation d'un échantillon de v.a. Normale $X \in \mathbb{R}$ à l'aide de v.a. Laplace Y.
  - function[X]=Rejet_Normale_2(   )
  - ◦ $C = \sqrt{\frac{2e}{\pi}}$
  - ◦ $k = 1;$
  - ◦ for n=1: N_{mc}
  - ◦ if rand() <1/2
    - Y= - V_A_Exp(1)
    - else Y= V_A_Exp(1)
  - ◦ endif
  - ◦ Simuler U[0,1] (indépendante)
  - ◦ $f = (\frac{e^{-Y^2/2}}{\sqrt{2\pi}});\quad g = \frac{1}{2}e^{-|Y|}$
  - ◦ if $U \leq \frac{f}{(C \cdot g)}$
  - ◦ $X(k) = Y; \quad k = k + 1;$
  - ◦ endif
  - ◦ endfor
  - endfunction

Travail à faire pour v.a. Normale

- Montrer que $C = \sqrt{\frac{2e}{\pi}}$

  Pour cela considerer deux $y \geq 0$ et $y < 0$.
  Pour $y \geq 0$ introduire la fonction

$$h(y) = \frac{f_X(y)}{g_Y(y)} = (\frac{2e^{-y^2/2}}{\sqrt{2\pi}})/(e^{-y}),$$

  deriver cette fonction par rapport à y, trouver un point de max et montrer que $C = h(y_{max})$.

- Verifier les simulations
  - Soient $N_{mc} = 10000$
  - Tracer les fonctions de repartition de X:
    $[a, b] = [-5, 5], \Delta = 0.1, N_x = 100$
  - Tracer les fonctions de densité de X :
    $[a, b] = [-5, 5], \Delta = 0.1, N_x = 100$

Simulation de la loi Beta (α, β)

- Une variable aléatoire de loi Beta $B(\alpha, \beta)$ (avec $\alpha > 0$ et $\beta > 0$) a pour densité

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} 1_{0 \leq x \leq 1}$$

- On utilise la méthode de rejet avec $Y = U([0,1])$ (loi uniforme) dont la fonction de densité est egale $g_Y(x) = 1$.
- La constante de rejet vaut

$$C = sup_{0 \leq x \leq 1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

- En dérivant la fonction $f_X(x)$ pour trouver le point de max x_0 on montre

$$x_0 = \frac{\alpha - 1}{\alpha + \beta - 2}$$

et

$$C = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_0^{\alpha-1}(1-x_0)^{\beta-1}.$$

Simulation de la loi Beta (α, β)

- On forme le vecteur X sans les 'zeros' par l'introduction d'un nouveau indice k.
  - function[X]=Rejet_Beta(α, β)
  - ∘ Calculer $x_0 = \frac{\alpha-1}{\alpha+\beta-2}$,
  - ∘ Calculer C=Beta(x_0, α, β)
  - ∘ $k = 1;$
  - ∘ for $n = 1 : N_{mc}$
  - ∘ Y=rand();
  - ∘ Simuler $U[0, 1]$ indépendant de Y
  - ∘ Simuler f=Beta(Y, α, β);
  - ∘ if $U \leq \frac{f}{C}$
  - ∘ $X(k) = Y; \quad k = k + 1;$
  - ∘ endif
  - ∘ endfor
  - endfunction

  - function[f]=Beta(x, α, β)

  $f = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} \cdot x.^{(\alpha-1)} . \cdot (1-x).^{(\beta-1)}$

  - endfunction

Travail à faire pour v.a. Beta (α, β)

- Simuler la v.a. Beta $(2, 2)$, Beta $(2, 5)$, Beta $(1.5, 3.5)$,
- Soient $N_{mc} = 10000$
- Tracer les fonctions de repartition de X:
  $[a, b] = [0, 1], \Delta = 0.01, N_x = 100$
- Tracer les fonctions de densité de X :
  $[a, b] = [0, 1], \Delta = 0.01, N_x = 100$
- Vous confirmez que v.a. $X \in [0, 1]$