Université Paris Dauphine
Département MIDO
Master 1 – Méthodes de Monte Carlo 2018–2019

Méthodes de Monte Carlo
– Projet –

stoehr@ceremade.dauphine.fr

Consignes.

À rendre avant le 08 janvier 2019

- Rapport contenant réponses/commentaires, codes utilisés et sorties (Notebook, Rmarkdown ou LaTeX+ knitr). À défaut, un rapport au format **.pdf** et un script contenant l'ensemble des codes utilisés. Dans ce cas, **il est interdit de copier-coller du code dans le corps du texte**. Une rédaction soignée et concise est attendue.
- **seul le language** R **est autorisé**.
- Les codes doivent :
  - être bien commentés. Il est possible qu'une explication orale vous soit demandée.
  - être optimisés (vectorisés) un minimum pour utiliser les spécificités du language.
  - s'exécuter sans erreurs et permettre de reproduire l'intégralité des résultats présentés dans le rapport.
- Les graphiques doivent être soigneusement annotés et présentés (titre, couleur, légendes, ...).
- **Chaque jour de retard sera pénalisé d'un point.**

Exercice 1 *(Simulation de variables aléatoires).*

Méthode du rejet. On souhaite simuler suivant une densité f de $\mathbb{R}^2$ proportionnelle à

$$\tilde{f}(x,y) = \frac{\cos^2(x) + 2\sin^2(y)\cos^4(x)}{1+4(y-1)^2} \exp\left(-\frac{\{x-2\}^2}{2}\right) \mathbb{1}_{\{x\in[0,4]\}} \mathbb{1}_{\{y[0,2]\}}. \tag{1}$$

1. Justifiez que pour obtenir une réalisation suivant la loi de densité f, on peut appliquer l'algorithme du rejet à $\tilde{f}$.

2. Proposer une constante M et une densité g telles que pour tout $(x,y) \in \mathbb{R}^2$, $\tilde{f}(x,y) \leq Mg(x,y)$.

3. En déduire une méthode de simulation suivant f.

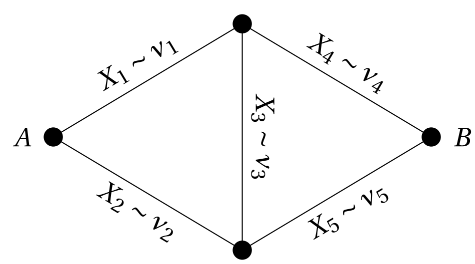
Algorithme de Metropolis–Hastings. L'algorithme de Metropolis–Hastings est un exemple particulier de méthode de Monte Carlo par chaîne de Markov, permetant de générer des chaînes de Markov, $\left(\mathbf{x}^{(t)}\right)_{t\geq 1}$ irréductibles, apériodiques et de loi stationnaire de densité f connue à une constante de normalisation près. Il est fondé sur le choix d'une loi instrumentale de densité g, **partout positive**, et sur le noyau de transition suivant

$$\mathbf{x}^{(t+1)} = \begin{cases} \xi \sim g & \text{avec probabilité} \quad \alpha\left(\mathbf{x}^{(t)}, \xi\right) \\ \mathbf{x}^{(t)} & \text{avec probabilité} \quad 1 - \alpha\left(\mathbf{x}^{(t)}, \xi\right) \end{cases} \qquad \text{avec} \quad \alpha\left(\mathbf{x}^{(t)}, \xi\right) = 1 \wedge \frac{f(\xi)}{f\left(\mathbf{x}^{(t)}\right)} \frac{g\left(\mathbf{x}^{(t)}\right)}{g(\xi)}.$$

4. À l'aide de l'algorithme de Metropolis–Hastings et en utilisant la fonction g obtenue à la question 2, produire une chaîne de Markov $(\mathbf{x}_t)_{t \in \mathbb{N}} = \left(x_1^{(t)}, x_2^{(t)}\right)$ de loi stationnaire f.

5. Comparer la distribution ainsi obtenue avec celle obtenue par la méthode du rejet.

Exercice 2 *(Méthodes de Monte Carlo).* On considère le graphe $\mathscr{G}$ dont chaque arête est munie d'un poids $X_1, \ldots, X_5$. On note $\mathbf{x} = (X_1, \ldots, X_5)$ le vecteur des poids et on suppose que les poids sont indépendants et distribués suivant des lois $\nu_1, \ldots, \nu_n$ précisées ultérieurement.



Un chemin simple entre A et B est défini par une suite d'arêtes consécutives de $\mathscr{G}$ reliant A à B, toutes les arêtes de la suite étant distinctes. La longueur d'un chemin simple entre A et B correspond à la somme des poids des arêtes constituant ce chemin. On note $d(\mathbf{X})$ la longueur du chemin simple le plus court entre A et B. Dans ce projet, on souhaite estimer par des méthodes de Monte Carlo

$$\delta = \mathbb{P}[X_1 + X_2 + X_3 \geq t] \quad \text{et} \quad p = \mathbb{P}[d(\mathbf{X}) \geq t] \quad t \in \mathbb{R}.$$

Les estimateurs doivent être donnés explicitement dans le rapport et les intervalles de confiance seront donnés au niveau 95%

Partie I – Estimation de δ

1. On suppose que $X_1 + X_2 + X_3$ est distribuée suivant une loi de Weibull de paramètre d'échelle $\lambda = 1$ et de paramètre de forme $k = 2$. Pour les applications numériques, on prendra $t = 2$.

   (a) Proposer une estimation $\widehat{\delta}_n$ de δ par la méthode de Monte Carlo classique.

   (b) Proposer une estimation $\widehat{\delta}_n(q_1, \ldots, q_K)$ de δ par la méthode de stratification avec allocation proportionnelle.

   (c) Comparer les performances de ces deux méthodes.

2. On suppose maintenant que X_1, X_2 et X_3 sont indépendantes avec X_1 et X_2 distribuées suivant la loi exponentielle de paramètre $\lambda = 1$ et X_3 distribuée suivant la loi de fonction de répartition

$$F(t) = \frac{t}{4}\mathbb{1}_{\{t\in[0,1[\}} + \left(\frac{t}{4}+\frac{1}{2}\right)\mathbb{1}_{\{t\in[1,2]\}} + \mathbb{1}_{\{t>2\}}.$$

Pour les applications numériques, on prendra $t = 1$.

(a) Proposer une estimation $\widehat{\delta}_n$ de δ par la méthode de Monte Carlo classique.

(b) Montrer que $\delta = \mathbb{E}[1 - F(t - X_1 - X_2)] = \mathbb{E}[1 - G(t - X_3)]$ où G est la fonction de répartition de la loi gamma $\Gamma(2,1)$.

(c) En déduire de nouvelles méthodes d'estimation de δ. Commenter leurs performances.

Partie II – Estimation de p

On considère n réalisations $\mathbf{X}_1,\dots,\mathbf{X}_n$ du vecteur des poids $\mathbf{X}$ supposés indépendants et distribués suivant des lois exponentielles de paramètres respectifs $\lambda_1,\dots,\lambda_5$. On notera $f(\mathbf{x};\lambda_1,\dots,\lambda_5)$ la densité du vecteur $\mathbf{X}$. Pour les applications numériques, on prend $(\lambda_1,\dots,\lambda_5) = (6,7,3,2,1)$ et $t = 2$.

1. Proposer une estimation $\widehat{p}_n$ de p par la méthode de Monte Carlo classique.

Echantillonage préférentiel.

2. **(a)** Monter qu'en choisissant pour loi d'importance

$$g(\mathbf{x}) = \frac{\mathbb{1}_{\{d(\mathbf{x})\geq t\}} f(\mathbf{x}\,;\lambda_1,\dots,\lambda_5)}{p}$$

on obtient un estimateur d'échantillonnage préférentiel de p de variance nulle.

(b) Proposer une méthode de simulation suivant la densité g.

(c) Quelle limitation voyez-vous quant au choix de cette loi d'importance ?

3. Soient $\mathbf{Y}_1,\dots,\mathbf{Y}_n$ des variables aléatoires *i.i.d.* de densité g.

(a) La quantité suivante converge-t-elle vers p,

$$\frac{\sum_{k=1}^n \mathbb{1}_{\{d(\mathbf{Y}_k)\geq t\}} w(\mathbf{Y}_k)}{\sum_{k=1}^n w(\mathbf{Y}_k)}, \quad \text{avec} \quad w(\mathbf{Y}_k) = \frac{f(\mathbf{Y}_k\,;\lambda_1,\dots,\lambda_5)}{g(\mathbf{Y}_k)} \quad ?$$

(b) Peut-on en déduire une méthode d'estimation de p ? Commenter.

Dans la suite, on choisit comme loi d'importance, la densité du vecteur $(Y_1,\dots,Y_5) := \mathbf{Y}$ tel que les variables $Y_1,\dots,Y_5$ soient indépendantes et distribuées suivant les lois exponentielles de paramètres respectifs $\alpha_1,\dots,\alpha_5$. On la notera $h(\mathbf{y};\alpha_1,\dots,\alpha_5)$.

4. Donner l'estimateur d'échantillonnage préférentiel $\widehat{p}_n(\alpha_1,\dots,\alpha_5)$ de p basé sur la loi d'importance h.

On souhaite choisir les paramètres $\alpha_1,\dots,\alpha_5$ de la loi d'importance de sorte que

$$(\alpha_1^\star,\dots,\alpha_5^\star) = \underset{\alpha_1,\dots,\alpha_5}{\operatorname{argmin}}\, \mathbb{E}_g\left[\ln\left\{\frac{g(\mathbf{Z})}{h(\mathbf{Z}\,;\,\alpha_1,\dots,\alpha_5)}\right\}\right],$$

où $\mathbb{E}_g$ est l'espérance par rapport à la densité g.

5. Soient $\alpha_1^0,\dots,\alpha_5^0$ les valeurs choisies initialement par l'utilisateur. Montrer que

$$(\alpha_1^\star,\dots,\alpha_5^\star) = \underset{\alpha_1,\dots,\alpha_5}{\operatorname{argmax}}\, \mathbb{E}_{\alpha_1^0,\dots,\alpha_5^0}\left[\mathbb{1}_{\{d(\mathbf{Y})\geq t\}}\frac{f(\mathbf{Y}\,;\,\lambda_1,\dots,\lambda_5)}{h(\mathbf{Y}\,;\,\alpha_1^0,\dots,\alpha_5^0)}\ln h(\mathbf{Y}\,;\,\alpha_1,\dots,\alpha_5)\right],$$

où $\mathbb{E}_{\alpha_1^0,\dots,\alpha_5^0}$ est l'espérance par rapport à la densité $h(\cdot;\,\alpha_1^0,\dots,\alpha_5^0)$.

6. Étant donné $\mathbf{Y}_k = (Y_{k,1},\dots,Y_{k,5})$, $k=1,\dots,n$, variables aléatoires indépendantes identiquement distribuées de densité $h(\cdot\,;\,\alpha_1^0,\dots,\alpha_5^0)$, en déduire que l'on peut estimer $(\alpha_1^\star,\dots,\alpha_5^\star)$ par

$$\widehat{\alpha}_j^\star = \frac{\displaystyle\sum_{k=1}^n \mathbb{1}_{\{d(\mathbf{Y}_k)\geq t\}}\prod_{i=1}^5\frac{\lambda_i}{\alpha_i^0}\exp\left(-\{\lambda_i-\alpha_i^0\}\,Y_{k,i}\right)}{\displaystyle\sum_{k=1}^n \mathbb{1}_{\{d(\mathbf{Y}_k)\geq t\}}Y_{k,j}\prod_{i=1}^5\frac{\lambda_i}{\alpha_i^0}\exp\left(-\{\lambda_i-\alpha_i^0\}\,Y_{k,i}\right)},\quad j=1,\dots,5.$$

7. Calculer $\widehat{p}_n(\hat{\alpha}_1^\star,\dots,\hat{\alpha}_5^\star)$ et l'intervalle de confiance au niveau 95% correspondant.

Réduction de variance. Soient $U_{1,1},\dots,U_{1,5},\dots,U_{n,1},\dots,U_{n,5}$ des variables aléatoires indépendantes et identiquement distribuées suivant la loi uniforme sur $[0,1]$. On pose, pour tout $i=1,\dots,n$,

$$\begin{aligned}\rho(U_{i,1},\dots,U_{i,5}) = \min\Bigg(&-\frac{1}{\alpha_1^\star}\ln U_{i,1}-\frac{1}{\alpha_4^\star}\ln U_{i,4}\,;\,-\frac{1}{\alpha_1^\star}\ln U_{i,1}-\frac{1}{\alpha_3^\star}\ln U_{i,3}-\frac{1}{\alpha_5^\star}\ln U_{i,5}\,;\\ &-\frac{1}{\alpha_2^\star}\ln U_{i,2}-\frac{1}{\alpha_5^\star}\ln U_{i,5}\,;\,-\frac{1}{\alpha_2^\star}\ln U_{i,2}-\frac{1}{\alpha_3^\star}\ln U_{i,3}-\frac{1}{\alpha_4^\star}\ln U_{i,4}\Bigg).\end{aligned}$$

8. **(a)** Montrer que l'estimateur suivant converge presque-sûrement vers p :

$$\frac{1}{n}\sum_{i=1}^n\prod_{k=1}^5\frac{\lambda_k}{\alpha_k^\star}U_{i,k}^{\lambda_k/\alpha_k^\star-1}\mathbb{1}_{\{\rho(U_{i,1},\dots,U_{i,5})\geq t\}}$$

(b) En déduire à l'aide de la méthode des variables antithétiques un estimateur $\widehat{p}_n^{(1)}$ de p.

(c) Comparez les performances de l'estimateur $\widehat{p}_n^{(1)}$ avec celles des estimateurs $\widehat{p}_n$ et $\widehat{p}_n(\hat{\alpha}_1^\star,\dots,\hat{\alpha}_5^\star)$.

Projet de Méthodes de Monte Carlo

FAURE Thomas et PETIT Guillaume

PSL Université Paris-Dauphine

8 Janvier 2019

Table des matières

Première partie

Exercice 1

1 Question 1

Pour obtenir une réalisation suivant la loi de densité f, on peut faire la méthode de rejet.

Soit $M > 0$. Si $\tilde{f}(x,y) \leq Mg(x,y)$, alors $f(x,y) = C\tilde{f}(x,y) \leq \tilde{M}g(x,y)$.

On aura par conséquent dans l'algorithme du rejet :

$$\mathcal{U} \leq \frac{f(X_n, Y_n)}{\tilde{M}g(X_n, Y_n)} = \frac{\tilde{f}(X_n, Y_n)}{Mg(X_n, Y_n)} \text{ car } \tilde{M} = CM$$

On peut donc appliquer l'algorithme à $\tilde{f}$ pour simuler f.

2 Question 2

On cherche à simuler selon $\tilde{f}$ donnée par :

$$\tilde{f}(x,y) = \frac{\cos^2(x) + 2\sin^2(x)\cos^4(x)}{1 + 4(y-1)^2} \exp\left(-\frac{\{x-2\}^2}{2}\right) \mathbb{1}_{\{x\in[0,4]\}} \mathbb{1}_{\{y\in[0,2]\}}$$

Notons :

$$h_1(x) = \frac{\exp\left(-\frac{\{x-2\}^2}{2}\right)}{\sqrt{2\pi}},\ h_2(x) = \frac{\mathbb{1}_{\{x\in[0,4]\}}}{4},$$

$$h_3(y) = \frac{1}{\frac{\pi}{2}\left(1+\left(\frac{y-1}{1/2}\right)^2\right)},\ h_4(y) = \frac{\mathbb{1}_{\{y\in[0,2]\}}}{2}$$

Nous pouvons alors exhiber quatre méthodes de rejet intéressantes :

2.1 Première méthode

$$\tilde{f}(x,y) = 2\sqrt{2\pi}\frac{(\cos^2(x) + 2\sin^2(x)\cos^4(x))}{1 + 4(y-1)^2} h_1(x)h_4(y)\mathbb{1}_{\{x\in[0,4]\}}$$

Donc

$$\tilde{f}(x,y) \leq M_1 h_1(x)h_4(y) \text{ avec } M_1 = 6\sqrt{2\pi} \text{ car } \begin{cases} \cos^2(x) + 2\sin^2(x)\cos^4(x) \leq 3 \\ 1 + 4(y-1)^2 \geq 1 \\ \mathbb{1}_{\{x\in[0,4]\}} \leq 1 \end{cases}$$

On note que $H(x,y) = h_1(x)h_4(y)$ est une densité.

2.2 Deuxième méthode

(par un même raisonnement)

$$\tilde{f}(x,y) = \pi\frac{1}{2}(\cos^2(x) + 2\sin^2(x)\cos^4(x))h_1(x)h_3(y)\mathbb{1}_{\{x\in[0,4]\}}\mathbb{1}_{\{y\in[0,2]\}}$$

Donc

$$\tilde{f}(x,y) \leq M_2 h_1(x)h_3(y) \text{ avec } M_2 = \frac{3}{2}\pi\sqrt{2\pi}$$

On note que $H(x,y) = h_1(x)h_3(y)$ est une densité.

2.3 Troisième méthode

(par un même raisonnement)

$$\tilde{f}(x,y) = 4 \times 2\frac{\cos^2(x) + 2\sin^2(x)\cos^4(x)}{1 + 4(y-1)^2}\exp\left(-\frac{\{x-2\}^2}{2}\right)h_2(x)h_4(y)$$

Donc

$$\tilde{f}(x,y) \leq M_3 h_2(x)h_4(y) \text{ avec } M_3 = 4\times 2\times 3 = 24 \text{ car } \exp\left(-\frac{\{x-2\}^2}{2}\right) \leq 1\ ,\ \forall x \in \mathbb{R}$$

2.4 Quatrième méthode

(par un même raisonnement)

$$\tilde{f}(x,y) = 4\times\pi\times\frac{1}{2}(\cos^2(x)+2\sin^2(x)\cos^4(x))\exp\left(-\frac{\{x-2\}^2}{2}\right)h_2(x)h_3(y)\mathbb{1}_{\{y\in[0,2]\}}$$

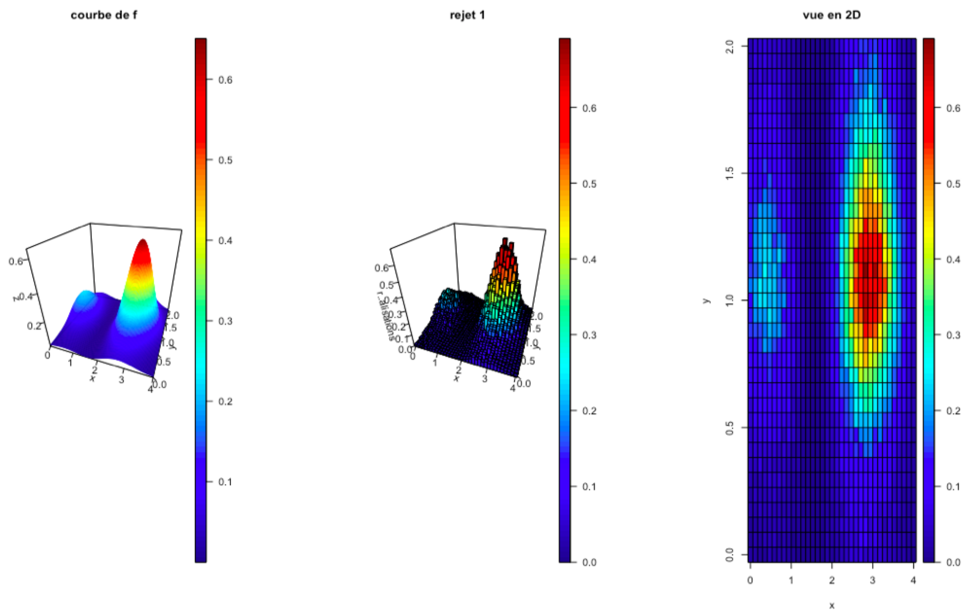
Donc

$$\tilde{f}(x,y) \leq M_4 h_2(x)h_3(y) \text{ avec } M_4 = 6\pi$$

3 Question 3

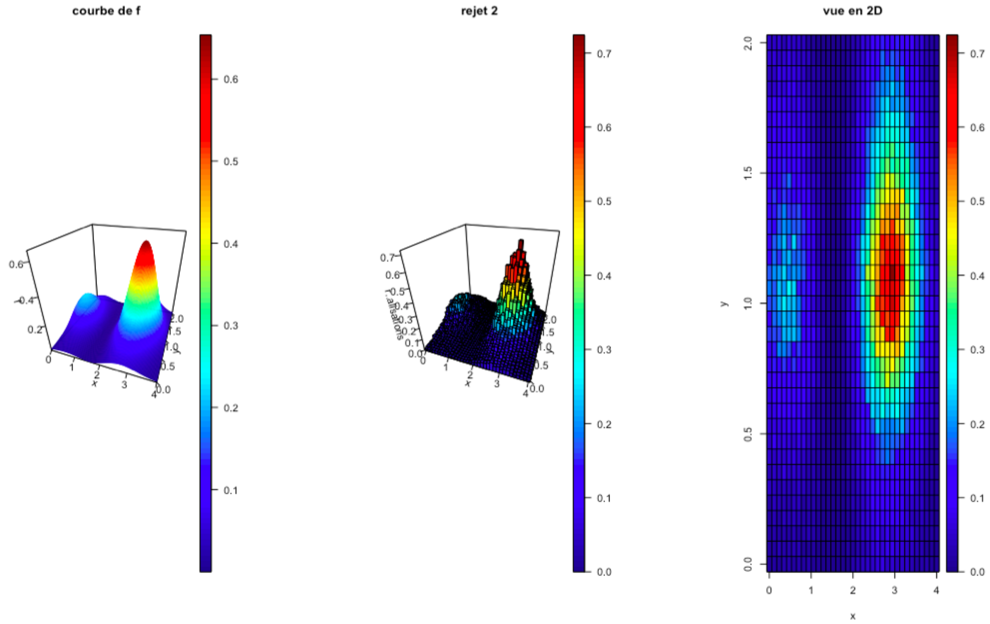
3.1 distributions

Après exécution, on obtient les résultats suivants [1] :

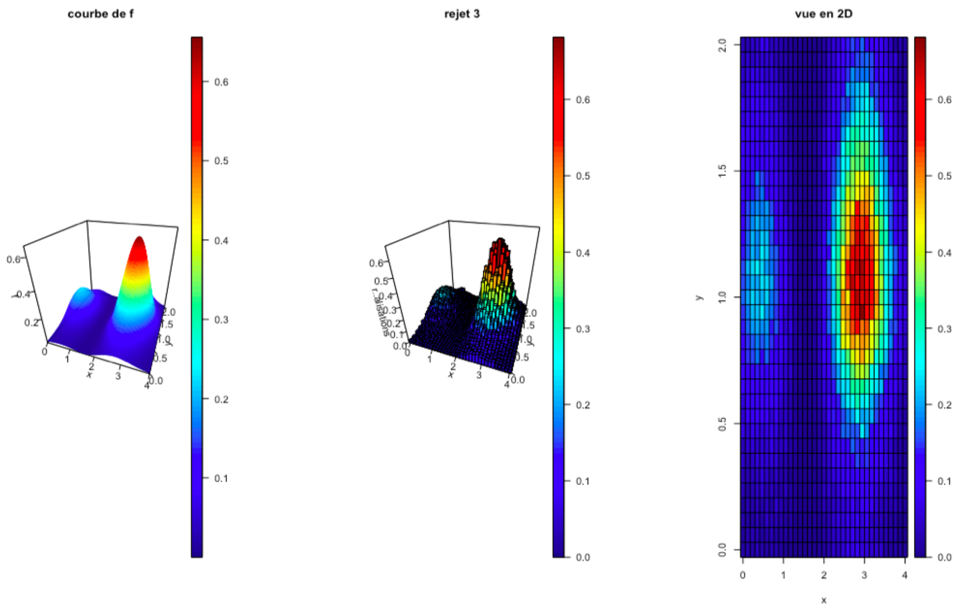


Graphes 3D et 2D pour Rejet 1

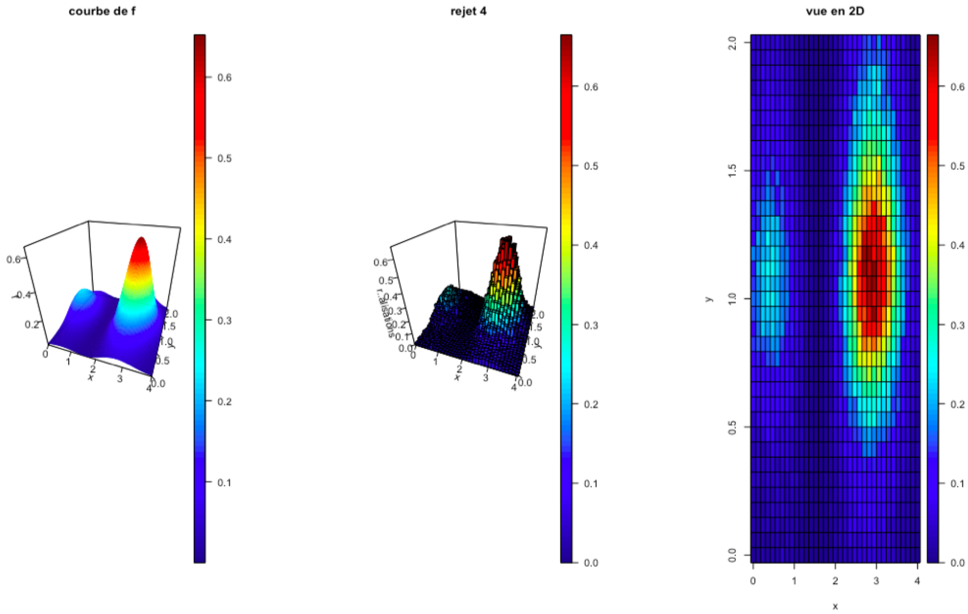
1. On met en annexe comment on a réussi à obtenir un plot de f, ie, calculer la constante de normalisation



Graphes 3D et 2D pour Rejet 2



Graphes 3D et 2D pour Rejet 3



Graphes 3D et 2D pour Rejet 4

3.2 temps d'exécution

Voici leurs temps d'éxécution :

Rejet	nombre de simulations	temps moyen (en ms)
Rejet 1	1.000	26.33
	10.000	103.31
	100.000	1615.97
Rejet 2	1.000	17.31
	10.000	74.03
	100.000	1076.56
Rejet 3	1.000	53.57
	10.000	214.38
	100.000	3394.83
Rejet 4	1.000	36.21
	10.000	149.66
	100.000	2325.20

Les résultats obtenus sont à peu près similaires. On va donc choisir le 2 qui est plus optimal. Ce résultat était prévisible, en effet cette méthode du rejet a la plus petite contante de majoration M. D'après le cours on sait que la le nombre d'essais moyen jusqu'à ce qu'une variable soit acceptée est M.

4 Question 4

L'algorithme de Metropolis Hastings (MH) permet de simuler la densité f, à une constante près. On a :

$$\mathbf{x}^{(t+1)} = \begin{cases} \xi \sim g & \text{avec probabilité} \quad \alpha\left(\mathbf{x}^{(t)}, \xi\right) \\ \mathbf{x}^{(t)} & \text{avec probabilité} \quad 1 - \alpha\left(\mathbf{x}^{(t)}, \xi\right) \end{cases} \quad \text{avec} \quad \alpha\left(\mathbf{x}^{(t)}, \xi\right) = 1 \wedge \frac{f(\xi)}{f\left(\mathbf{x}^{(t)}\right)} \frac{g\left(\mathbf{x}^{(t)}\right)}{g(\xi)}.$$

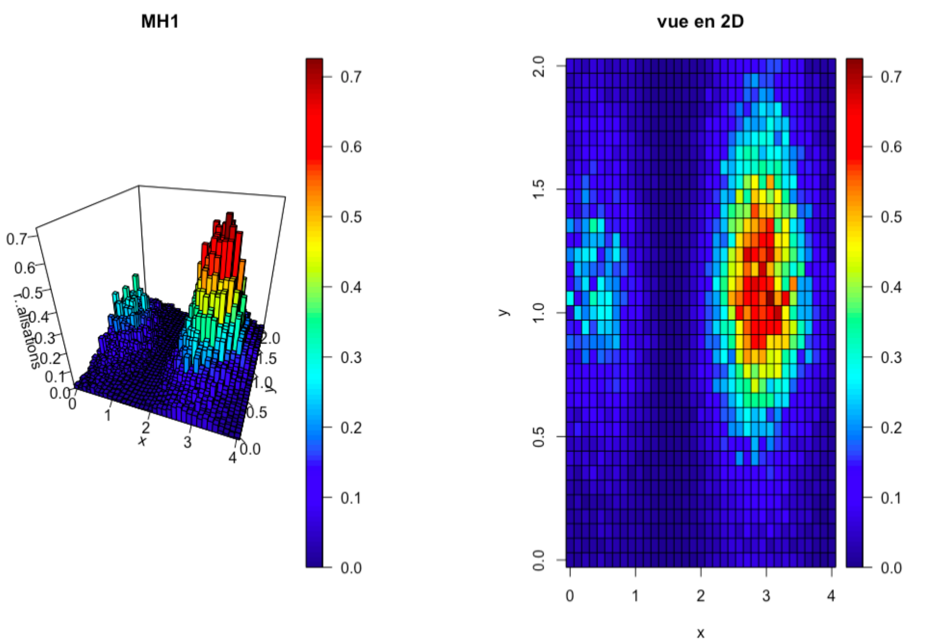
Or,

$$\frac{f(\xi) g\left(\mathbf{x}^{(t)}\right)}{f\left(\mathbf{x}^{(t)}\right) g(\xi)} = \frac{c\tilde{f}(\xi) g\left(\mathbf{x}^{(t)}\right)}{c\tilde{f}\left(\mathbf{x}^{(t)}\right) g(\xi)} = \frac{\tilde{f}(\xi) g\left(\mathbf{x}^{(t)}\right)}{\tilde{f}\left(\mathbf{x}^{(t)}\right) g(\xi)}$$

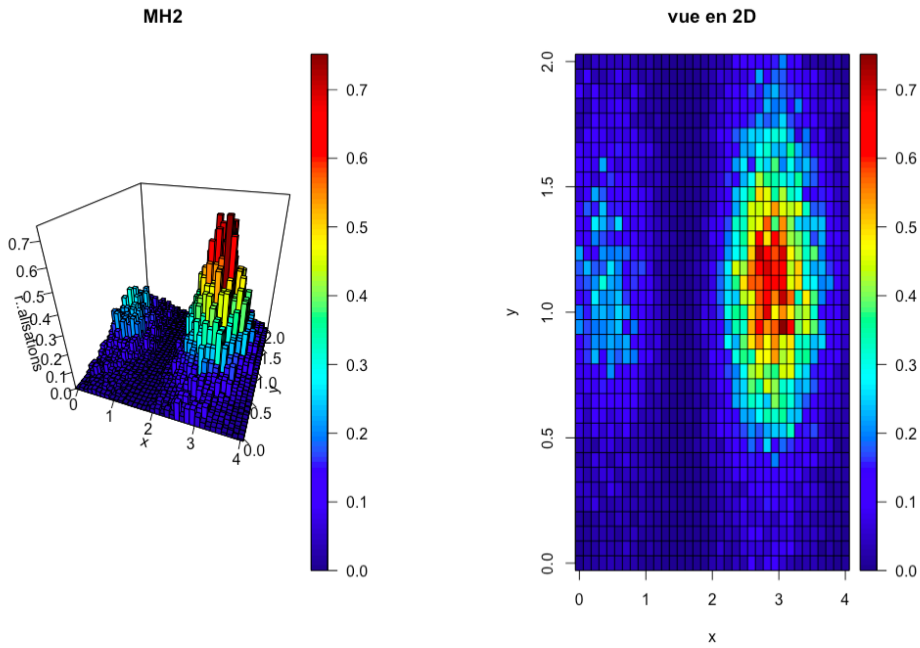
On pourra donc réécrire :

$$\alpha\left(\mathbf{x}^{(t)}, \xi\right) = 1 \wedge \frac{\tilde{f}(\xi) g\left(\mathbf{x}^{(t)}\right)}{\tilde{f}\left(\mathbf{x}^{(t)}\right) g(\xi)}$$

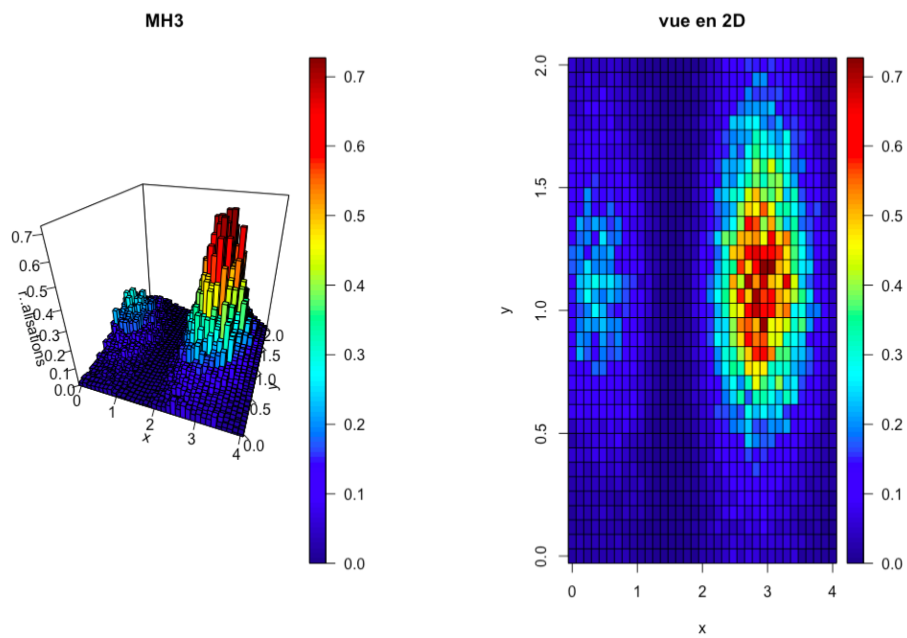
De plus, si $\xi \notin supp(f)$, alors $\alpha(\mathbf{x}^{(t)}, \xi) = 0$ et donc $\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)}$ avec proba 1. Donc $\forall t \in \mathbb{N}$, $\mathrm{x}^{(t)}$ est bien dans le support de f. g_1, g_2, g_3 et g_4 sont des densités, donc positives pour tout couple $(x, y) \in \mathbb{R}^2$. On peut donc appliquer l'algorithme de Métropolis-Hastings et on obtient les résultats suivants :



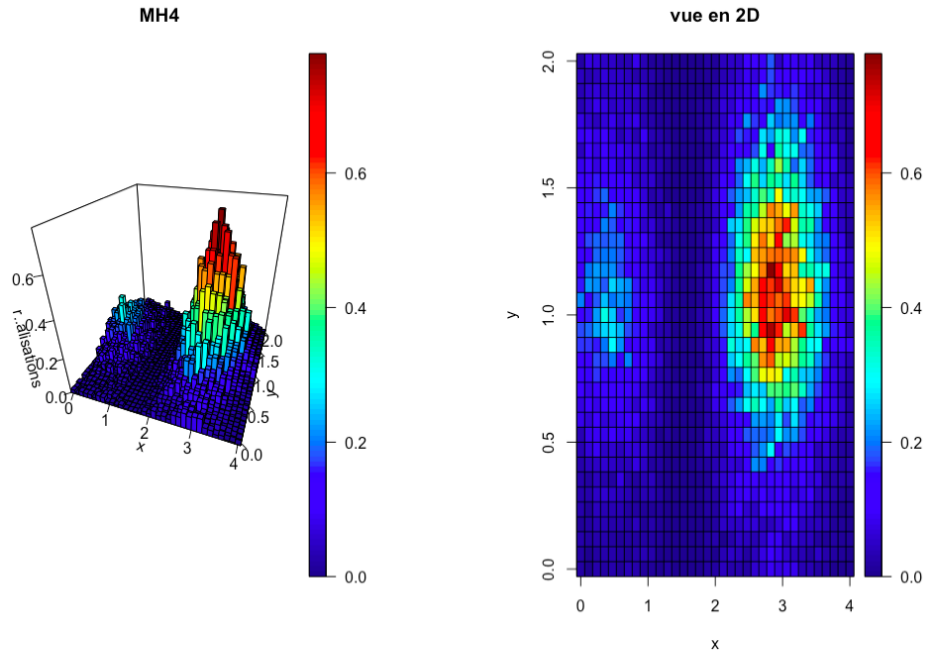
Graphes 3D et 2D pour MH 1



Graphes 3D et 2D pour MH 2

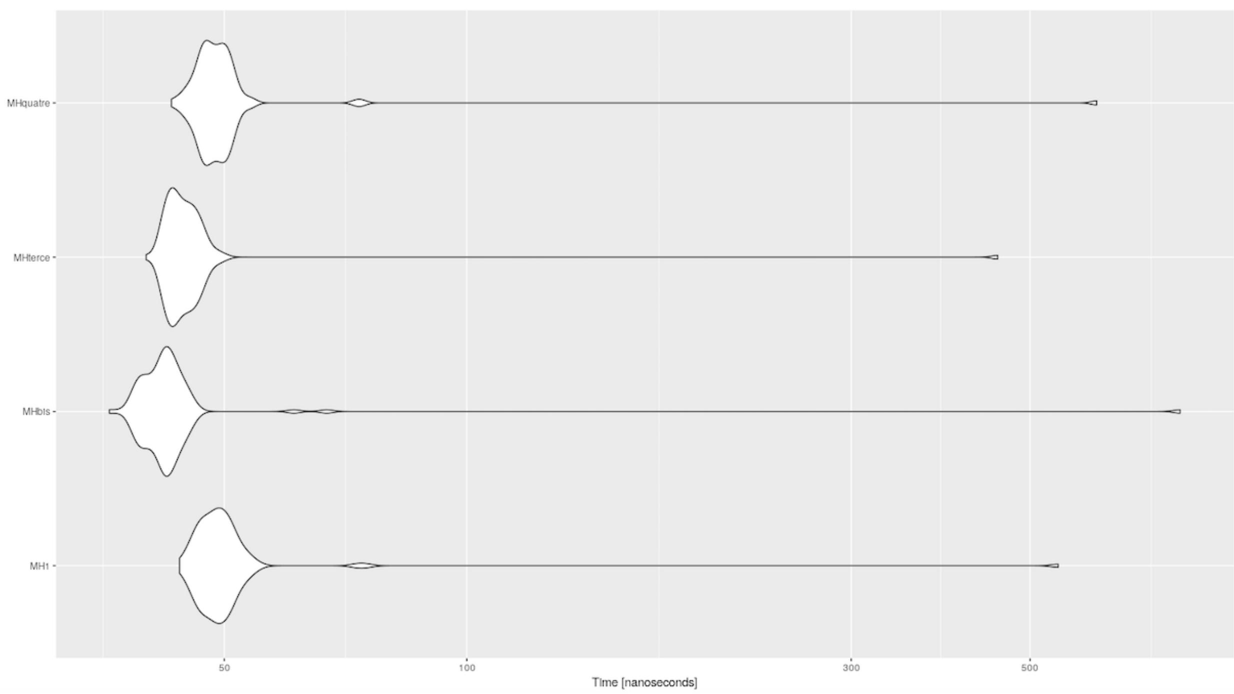


Graphes 3D et 2D pour MH 3



Graphes 3D et 2D pour MH 4

On calcule aussi le temps d'éxécution des fonctions. On obtient cela :

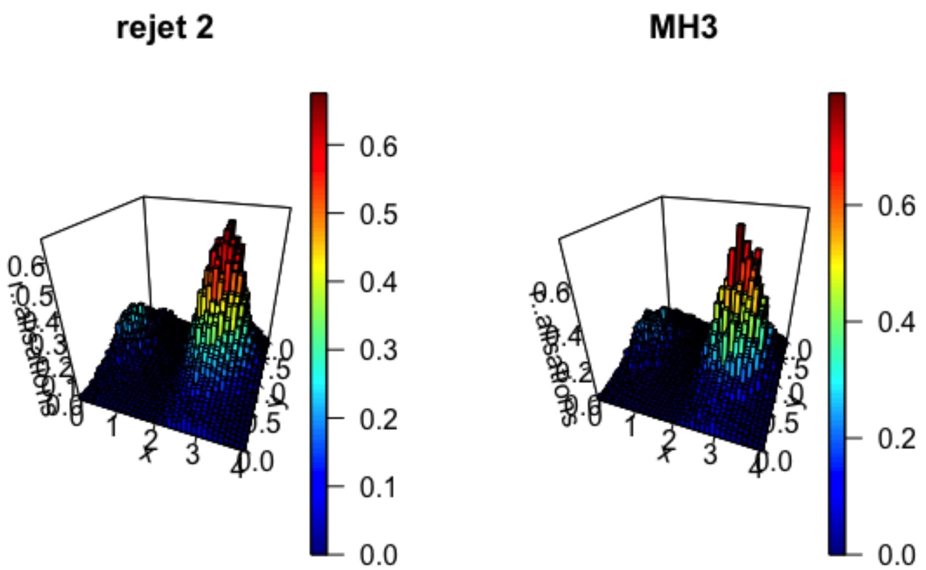


Elles obtiennent toutes des simulations à peu près similaires, on va donc choisir l'algorithme de Métropolis-Hastings numéro 3 qui met le moins de temps.

5 Question 5

Nous allons mettre côte à côte les distributions obtenues et le temps d'exécution des fonctions du rejet 2 et de Metropolis Hastings 3 [2], les fonctions que nous avons gardées.

5.1 graphique



comparaison des distributions entre le rejet 2 et MH

2. On met en annexe les comparaisons pour les autres méthodes pour se faire une idée de la précision de chaque méthode pour la même densité g

5.2 temps d'exécution

fonction étudiée	nombre de simulations	temps moyen (en ms)
rejet 2	1.000	5.29
	10.000	65.85
	100.000	660.18
MH 3	1.000	36.44
	10.000	366.66
	100.000	3767.78

On remarque sur les graphiques ont une meilleure précision pour la méthode du rejet qui est d'ailleurs plus rapide à exécuter quand on compare les temps d'exécution.

Deuxième partie

Exercice 2

6 Partie 1

6.1 Question 1

6.1.1 a

On a :

$$\delta = \mathbb{P}(X_1 + X_2 + X_3 \geq t) = \mathbb{E}[\mathbb{1}_{\{X_1+X_2+X_3 \geq t\}}]$$

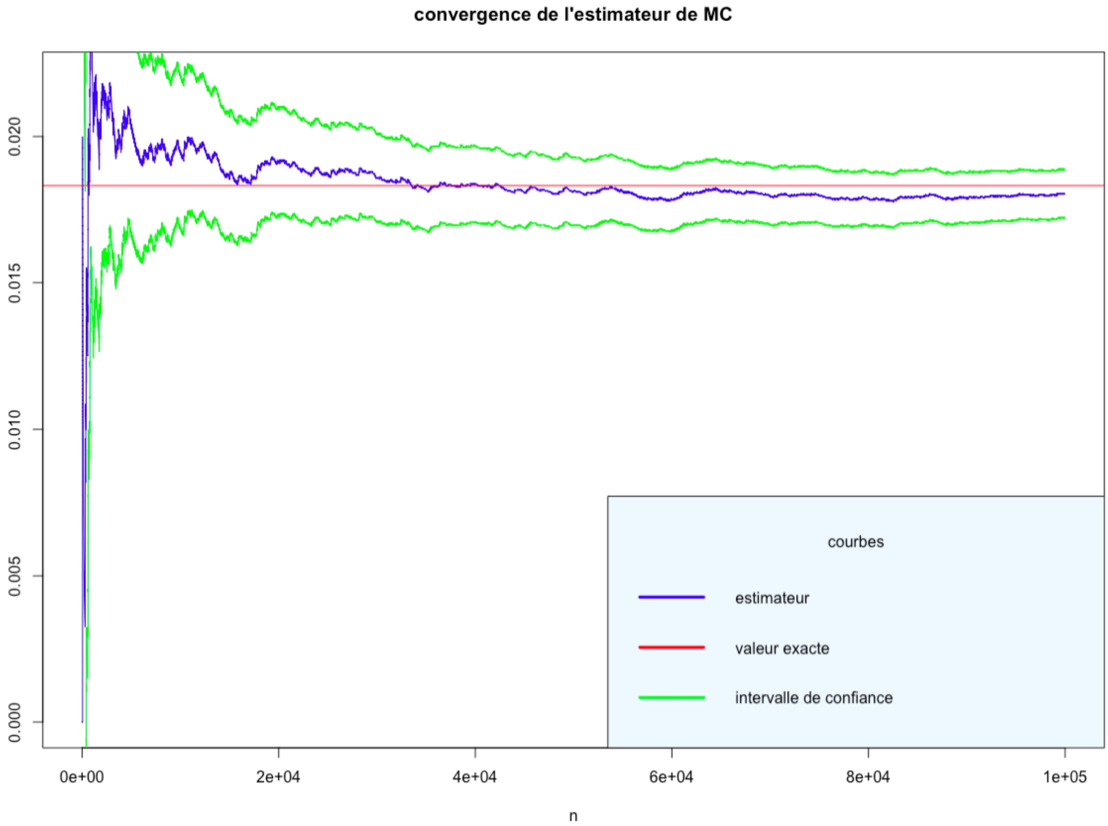
Posons :

$$indicatrice(W) = \mathbb{1}_{\{W \geq t\}}$$

On a alors l'estimateur suivant :

$$\hat{\delta}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{W_i \geq t\}} = \frac{1}{n}\sum_{i=1}^{n} indicatrice(W_i) \text{ où } W_i \sim W(2,1) \text{ iid}$$

Pour 10000 réalisations, $\hat{\delta} = 0.0178$ avec une variance de 0.0175 et $IC_{0.975} = [0.0152, 0.0204]$



6.1.2 b

On veut $p_l = q_l$, ainsi on peut réécrire l'estimateur :

$$\delta_n(p_1, ..., p_L) = \sum_{l=1}^{L} \frac{p_l}{n_l} \sum_{i=1}^{n_l} \mathbb{1}_{\{W_i^{(l)} \geq t\}} = \frac{1}{n} \sum_{l=1}^{L} \sum_{i=1}^{n_l} \mathbb{1}_{\{W_i^{(l)} \geq t\}}$$

Pour créer nos variables $W^{(l)}$ qui suivent la loi $W|W \in D_l$, on pose :

$$W^{(l)} = \Phi^{-1}[\Phi(d_l) + \mathcal{U}[\Phi(d_{l+1}) - \Phi(d_l)]]$$

où Φ et Φ^{-1} sont respectivement la fonction de répartition et la fonction de répartition inverse d'une $Weibull(2, 1)$.

Cependant, l'allocation proportionnelle nous impose $n_l = n\mathbb{P}(W \in D_l)$. On va prendre des D_l tels que $n\mathbb{P}(W \in D_l) \in \mathbb{N}$.

Posons :

$$D_l = \left[\Phi^{-1}\left(\frac{l-1}{L}\right), \Phi^{-1}\left(\frac{l}{L}\right)\right] \text{ pour } l = 1, ..., L$$

On obtient que $n\mathbb{P}(W \in D_l) = \frac{n}{L}$. Il en découle :

1. on prendra $n \equiv O[L]$ et n_l sera constant
2. $W^{(l)} = \Phi^{-1}\left(\frac{l-1}{L} + \frac{u}{L}\right)$
3. $\delta_n(p_1, ..., p_L) = \frac{1}{n}\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{n/L} \mathbb{1}_{\{W_i^{(l)} \geq t\}}$

Pour 10000 réalisations et $L = 1000$, on obtient $\delta_n(p_1, ..., p_L) = 0.183$ avec une variance égale à 0.000212 et $IC_{0.975} = [0.182, 0.184]$. On remarque que plus on stratifie, plus l'intervalle de confiance est petit.

6.1.3 c

Vitesse de convergence

Soit $\epsilon > 0$ et $\alpha \in (0, 1)$. On cherche un rang $n \in \mathbb{N}$ à partir duquel les déviations des estimateurs sont plus petites que ϵ avec probabilité α. Soit ϕ la fonction de répartition de la loi $\mathcal{N}(0, 1)$.

1. **Monte Carlo** :
   Asymptotiquement, $\frac{\sqrt{n}}{\sigma}(\hat{\delta}_n - \delta)$ converge en loi vers une $\mathcal{N}(0, 1)$.
   Donc $\mathbb{P}(|\hat{\delta}_n - \delta| > \epsilon) = \mathbb{P}\left(\frac{\sqrt{n}}{\sigma}|\hat{\delta}_n - \delta| > \frac{\sqrt{n}}{\sigma}\epsilon\right) \underset{n\to+\infty}{\approx} 2\left(1 - \phi\left(\frac{\sqrt{n}}{\sigma}\epsilon\right)\right)$
   et on veut : $\left[2\left(1 - \phi\left(\frac{\sqrt{n}}{\sigma}\epsilon\right)\right) \leq \alpha\right]$ équivalent à $\left[n \geq \frac{\sigma^2\phi^{-1}(1-\alpha/2)}{\epsilon^2}\right]$
2. **Stratification** :
   Asymptotiquement, $\sqrt{n}(\delta_n(p_1, ..., p_n) - \delta)$ converge en loi vers une $\mathcal{N}(0, \tilde{\sigma}^2)$ où $\tilde{\sigma}^2 = \sum\limits_{l=1}^{L}\sigma_l^2$ et $\sigma_l^2 = var\left(\mathbb{1}_{\{w_1^{(l)} \leq t\}}\right)$ Par un raisonnement analogue, on a $n \geq \frac{\tilde{\sigma}^2\phi^{-1}(1-\alpha/2)}{\epsilon^2}$.

Or, on sait que :

$$var(\delta_n(p_1, ..., p_n)) = \frac{1}{n}\tilde{\sigma}^2 \leq var(\hat{\delta}_n) = \frac{1}{n}\sigma^2$$

d'où

$$\tilde{\sigma}^2 \leq \sigma^2$$

Temps de calcul

Il faut néanmoins prendre en compte le temps de calcul. Nous obtenons :

fonction étudiée	nombre de simulations	nombre de strats	temps moyen (en ms)
Monte Carlo	10.000	aucune	1.45
	100.000	aucune	21
Stratification	10.000	1.000	40.068
	100.000	1.000	64.86
	100.000	10.000	787.93

Conclusion

Donc, pour obtenir une même précision, la stratification nécessite moins d'itérations. De plus, l'estimateur stratifié à un intervalle de confiance plus petit. Cependant, il y a un coût en calcul à payer.

6.2 Question 2

6.2.1 a

Montrons que $X_1 + X_2 \sim \Gamma(2,1)$.

On a que X_1 et X_2 sont indépendantes et suivent la loi exponentielle de paramètre $\lambda = 1$.

Posons $W = X_1 + X_2$ et $U = X_1$, on a :

$$f_{(U,W)} = \int_{(\mathbb{R}_*^+)^2} e^{-(x_1+x_2)} dx_1 dx_2 = \int_{(\mathbb{R}_*^+)^2} e^{-w} \mathbb{1}_{w>u} dw du$$

Donc :

$$f_W(w) = e^{-w} \mathbb{1}_{w>0} \int_0^w du = we^{-w} \mathbb{1}_{w>0}$$

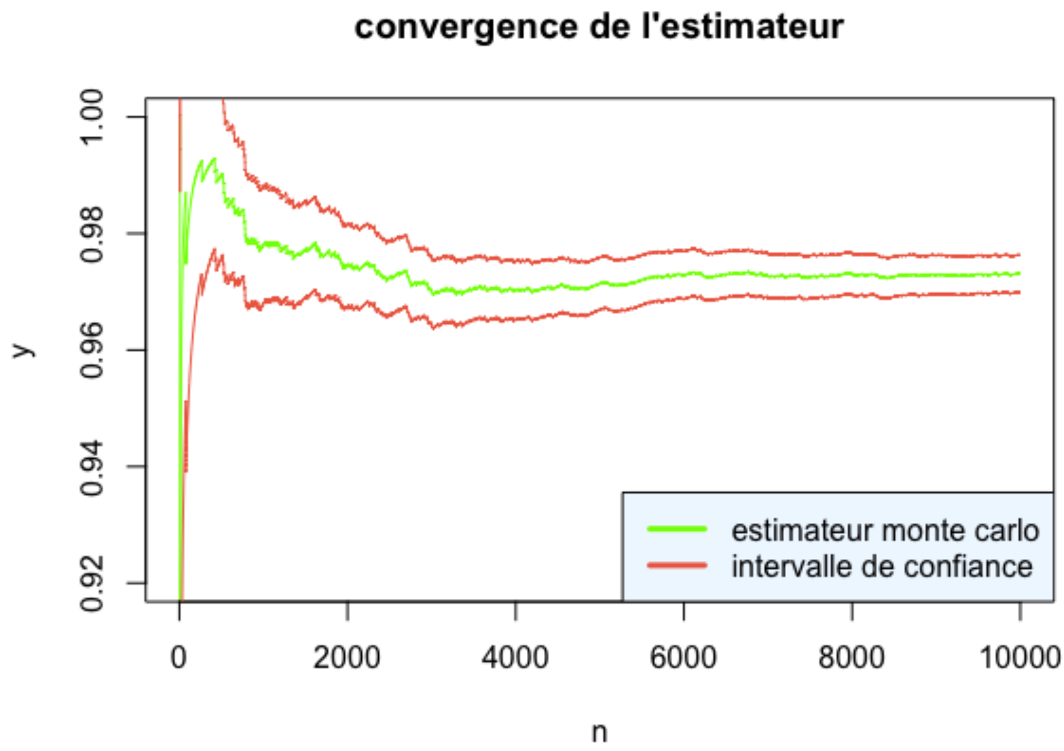
Ainsi :

$$W \sim \Gamma(2,1)$$

On pose l'estimateur de Monte Carlo suivant :

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{W_i + X_{3,i} \geq t\}}$$

On obtient que $\hat{\delta}_{MC} = 0,976$ avec une variance égale à 0.0255 et $IC = [0,973; 0,979]$. On peut observer la convergence de l'estimateur et de son intervalle de confiance sur le graphe ci dessous :



6.2.2 b

$$\delta = 1 - \mathbb{P}(X_3 \leq t - W) = 1 - \mathbb{E}[\mathbb{1}_{\{X_3 \leq t-W\}}]$$

on note $\varphi(x)$ la densité de X_3

$$\mathbb{E}[\mathbb{1}_{\{X_3 \leq t-W\}}] = \int_{(\mathbb{R}_*^+)^2} \varphi(x)g(w)\mathbb{1}_{\{x \leq t-w\}}dxdw = \int_{\mathbb{R}_*^+} g(w) \left(\int_0^{t-w} \varphi(x)dx \right) dw$$

$$= \int_{\mathbb{R}_*^+} g(w)F(t-w)dw = \mathbb{E}[F(t-W)]$$

$$\delta = 1 - \mathbb{E}[F(t-W)] = \mathbb{E}[1 - F(t-W)]$$

De la même manière on trouve,

$$\mathbb{E}[\mathbb{1}_{\{W\leq t-X_3\}}] = \int_{(\mathbb{R}_*^+)^2} \varphi(x)g(w)\mathbb{1}_{\{w\leq t-x\}}dxdw = \int_{\mathbb{R}_*^+} \varphi(x)\left(\int_0^{t-x} g(w)dw\right)dx$$

$$= \int_{\mathbb{R}_*^+} \varphi(x)G(t-x)dx = \mathbb{E}[G(t-X_3)]$$

$$\delta = 1 - \mathbb{E}[G(t-X_3)] = \mathbb{E}[1-G(t-X_3)]$$

Finalement,

$$\delta = \mathbb{E}[1-F(t-X_1-X_2)] = \mathbb{E}[1-G(t-X_3)]$$

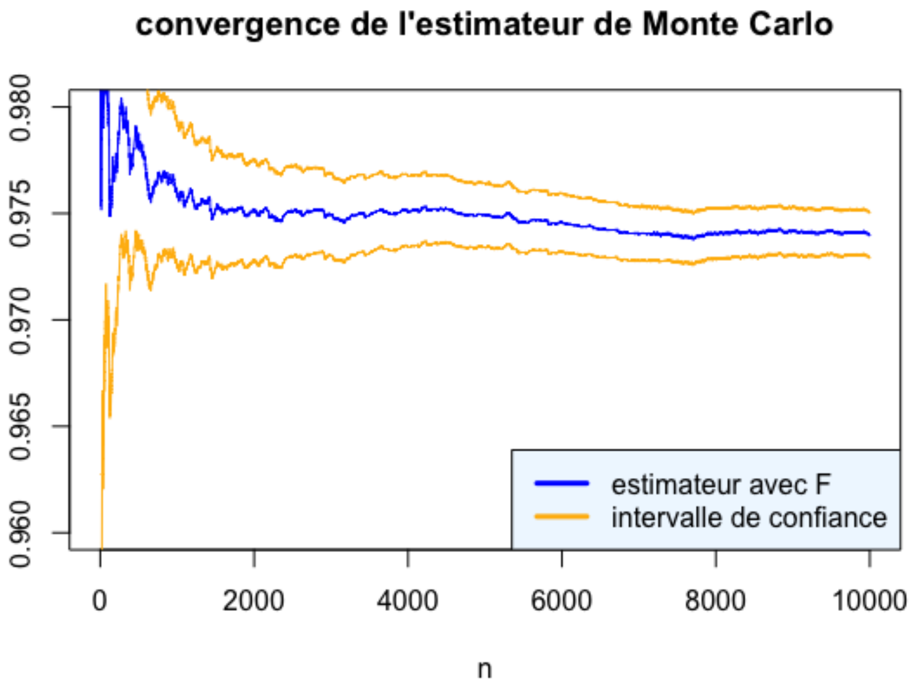
6.2.3 c

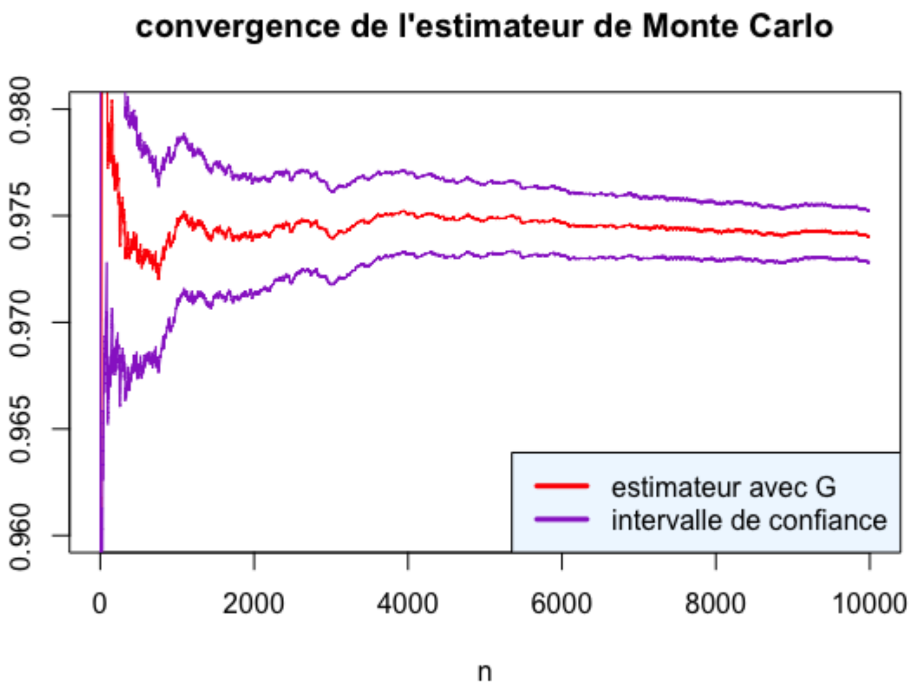
Méhode de Monte Carlo

Finalement, nous allons calculer un estimateur de Monte Carlo pour chacune des espérances précédentes.
1er cas : $\delta = \mathbb{E}[1-F(t-X_1-X_2)]$

On pose $\hat{\delta}_F = \frac{1}{n}\sum_{i=1}^{n} 1 - F(t-W_i)$. Pour 10.000 réalisation, $\hat{\delta}_F = 0,974$ avec une variance égale à 0.00276 et l'intervalle de confiance $IC = [0,973; 0,975]$.
2ème cas : $\delta = \mathbb{E}[1-G(t-X_3)]$

On pose $\hat{\delta}_G = \frac{1}{n}\sum_{i=1}^{n} 1 - G(t-X_{3,i})$ Pour 10.000 réalisation, $\hat{\delta}_G = 0,974$ avec une variance égale à 0.00367 et l'intervalle de confiance $IC = [0,973; 0,975]$.
On peut d'ailleurs remarquer visualiser leur convergence sur ces graphiques :


convergence de l'estimateur de Monte Carlo
0.960
0.965
0.970
0.975
0.980
0
2000
4000
6000
8000
10000
n
estimateur avec F
intervalle de confiance


convergence de l'estimateur de Monte Carlo
0.960
0.965
0.970
0.975
0.980
0
2000
4000
6000
8000
10000
n
estimateur avec G
intervalle de confiance

Stratification

On va aussi stratifier l'une des deux espérances. On choisit celle avec F. On effectue exactement le même raisonnement que pour la question 1.b) en prenant $D_l = \left[F^{-1}\left(\frac{l-1}{L}\right), F^{-1}\left(\frac{l}{L}\right)\right]$ avec $l = 1, ...L$ et F^{-1} la fonction de répartition inverse de X_3 et n multiple de L. On pose l'estimateur suivant :

$$\delta_n(p_1, ..., p_L) = \frac{1}{n}\sum_{l=1}^{L}\sum_{i=1}^{n/L} 1 - F(t - W_i) \text{ où } W_i \overset{iid}{\sim} \Gamma(2, 1)$$

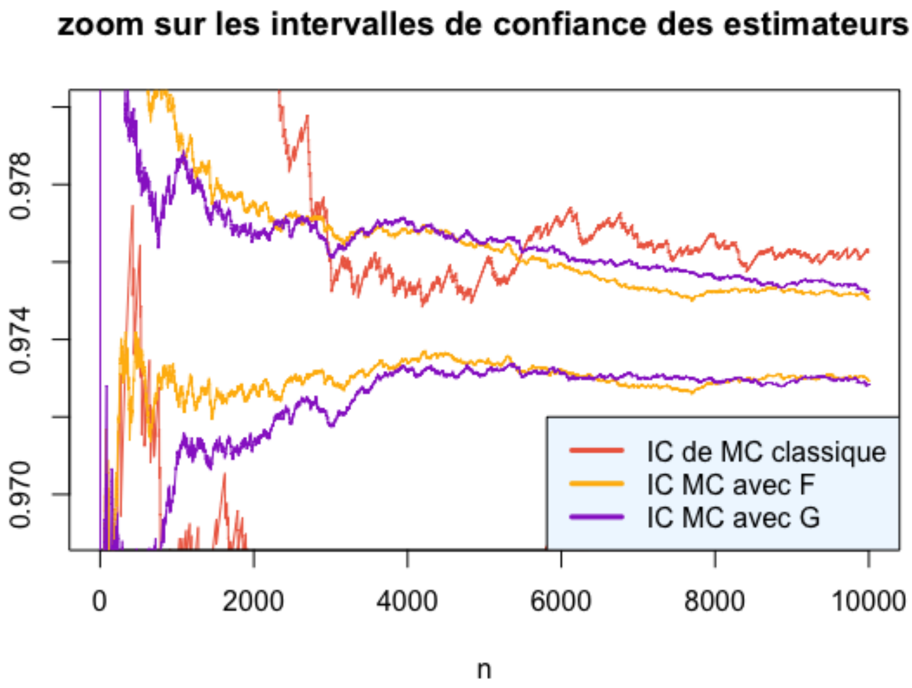
Pour 10.000 réalisations et $L = 1.000$, $\delta_n(p_1, ..., p_L) = 0.9740879$ avec une variance égale à 0.000233 et l'intervalle de confiance $IC_{0.975} = [0.9740846, 0.9740913]$

Conclusion

On remarque que l'estimateur stratifié est bien plus précis que l'estimateur de Monte Carlo pour un même nombre de réalisations. Cependant, comme on peut le voir sur le tableau ci-dessous, il y a coût de calcul bien plus important pour la stratification.

fonction étudiée	nombre de simulations	nombre de strats	temps moyen (en ms)
Monte Carlo avec F	10.000	aucune	3.48
	100.000	aucune	65.23
Monte Carlo avec G	10.000	aucune	2.9
	100.000	aucune	51.38
Monte Carlo simple	10.000	aucune	4.48
	100.000	aucune	75.14
Stratification avec F	10.000	1.000	57.60
	100.000	1.000	157.92
	100.000	10.000	967.81

Il est intéressant de noter que $\hat{\delta}_G$ et $\hat{\delta}_F$ sont plus rapides que $\hat{\delta}_n$. On va donc essayer de voir la précision des intervalles de confiance.



On peut donc conclure que $\hat{\delta}_G$ et $\hat{\delta}_F$ ont des IC plus précis que $\hat{\delta}_n$. Ainsi, si on veut privilégier la précision, on choisira la stratification, tandis que si on veut privilégier la rapidité, on choisira $\hat{\delta}_G$.

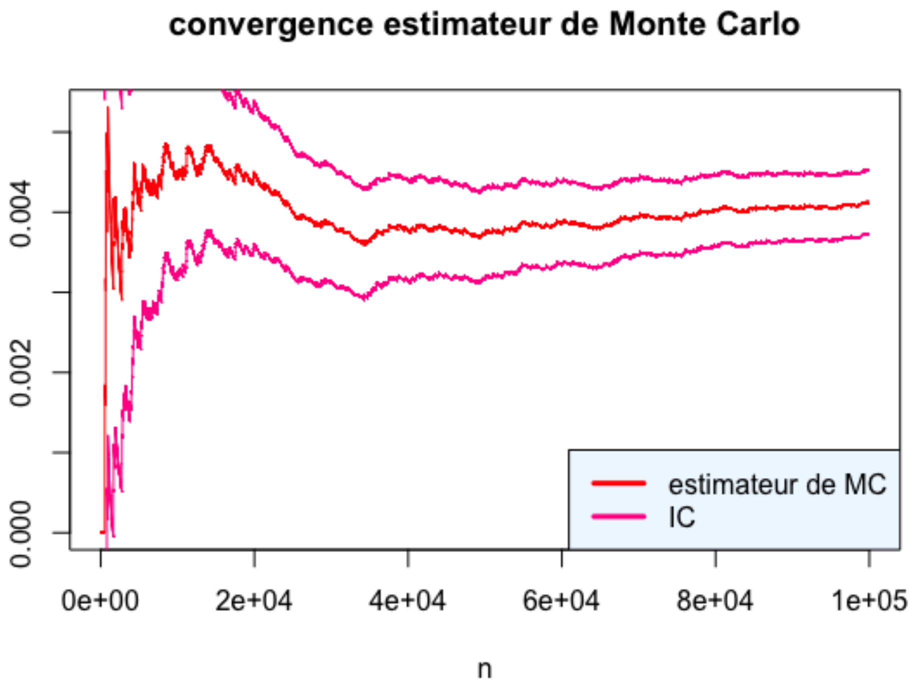
7 Partie II

7.1 Question 1

Par le cours, on pose

$$\hat{p}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{d(X_k)\geq t\}} \text{ où } X_k \overset{iid}{\sim} f \text{ et } X_k = (X_{1,k}, X_{2,k}, X_{3,k}, X_{4,k}, X_{5,k})$$

Pour 10.000 réalisations, on obtient que $\hat{p}_n \approx 0,00412$ avec une variance égale à 0.00411 et $IC_{0,975} = [0,00372; 0,00452]$.



7.2 Question 2

7.2.1 a

$$\text{Soit } g(x) = \frac{\mathbb{1}_{\{d(x)\geq t\}} f(x;\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)}{p} \text{ pour loi d'importance.}$$

Pour commencer, on va montrer que g est une densité, puis que l'estimateur d'échantillonnage préférentielle à une variance nulle.

1. **Densité** :

$$\int_{\mathbb{R}^+} g(x)dx = \frac{1}{p}\int_{\mathbb{R}^+} \mathbb{1}_{\{d(x)\geq t\}} f(x;\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)dx = \frac{1}{p}\mathbb{E}_f[\mathbb{1}_{\{d(x)\geq t\}}] = 1$$

2. **Variance** :
On obtient l'estimateur $\hat{\delta}_n(g) = \frac{1}{n}\sum_{i=1}^{n} \frac{f(X_i;\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)}{g(X_i)}\mathbb{1}_{\{d(X_i)\geq t\}}$
où $X_i \overset{iid}{\backsim} g$

Or, $\dfrac{f(x;\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)}{g(x)}\mathbb{1}_{\{d(x)\geq t\}} = p$

Donc $\mathbb{V}(\hat{\delta}_n(g)) = \mathbb{V}(\frac{1}{n}\sum_{i=1}^{n} p) = 0$

7.2.2 b

On va simuler g par rejet. On peut appliquer la même méthode qu'à l'exercice 1 car on connait g à une constante $(1/p)$ près. On a que

$$g(x) = \frac{\mathbb{1}_{\{d(x)\geq t\}} f(x;\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)}{p} \leq \frac{f(x;\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)}{p}$$

et donc on obtiendra comme condition $\mathcal{U} \leq \mathbb{1}_{d(Y_k)\geq t}$ où $Y_k \sim g$

7.2.3 c

g dépend de p, qui est la quantité à chercher. De plus, $\frac{f(x)}{g(x)} = \frac{p}{\mathbf{1}_{\{d(X)\geq t\}}}$ n'est pas bien définie car l'indicatrice peut être nulle. Ce problème vient du fait que le support de f n'est pas inclut dans celui de g. On ne peut donc pas utiliser la méthode de l'échantillonage préférentiel car la loi d'importance n'est pas bien posée.

7.3 Question 3

7.3.1 a

$$\begin{aligned} \frac{\sum_{k=1}^{n} \mathbf{1}_{\{d(Y_k)\geq t\}} w(Y_k)}{\sum_{k=1}^{n} w(Y_k)} &= \frac{np}{\sum_{k=1}^{n} w(Y_k)} && (1) \\ &= \frac{p}{\frac{1}{n}\sum_{k=1}^{n} w(Y_k)} && (2) \end{aligned}$$

or (Y_k) iid et intégrables, donc

$$\frac{1}{n}\sum_{k=1}^{n} w(Y_k) \xrightarrow[n\to+\infty]{loi} \mathbf{E}_g[w(Y_k)]$$

par la loi des grands nombres et $\mathbf{E}_g[w(Y_k)] = \int_{supp(g)} f(y)dy \neq 1$ car $supp(f) \nsubseteq supp(g)$. Donc la quantité ne converge pas vers p.

7.3.2 b

Dans l'exercice 1, on connaissait f à une constante près et on a réussi à estimer la constante dans le rejet.
Ici, on connait g à la constante $\frac{1}{p}$ près. On peut estimer $\frac{1}{p}$ et ainsi obtenir une estimation de p. On obtient que l'estimation de p vaut à peu près 0,0038 avec comme intervalle de confiance $IC_{0.975} = [0.00381, 0.00397]$

7.4 Question 4

on appelle : $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4; \lambda_5)$
$\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$
$\alpha^0 = (\alpha_1^0, \alpha_2^0, \alpha_3^0, \alpha_4^0, \alpha_5^0)$
0n a

$$h(Y, \alpha) = \prod_{i=1}^{5} \alpha_i exp(-\alpha_i y_i)$$

D'après le cours, l'estimateur d'échantillonage préférentiel $\hat{p}_n(\alpha)$ est donné par :

$$\hat{p}_n(\alpha) = (1/n) \sum_{k=1}^{n} \frac{f(Y^k, \lambda)}{h(Y^k, \alpha)} \mathbf{1}_{\{d(Y^k) \geq t\}}$$

avec $Y^k = (Y_1^k, Y_2^k, Y_3^k, Y_4^k, Y_5^k)$ un vecteur de variables aléatoires simulé selon la loi de densité $h(Y^k, \alpha)$
Après plusieurs essais, on obtient les résultats suivants :

$(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$	estimateur	variance	intervalle de confiance
$(2, 7, 4, 4, 2)$	0.00206	0.242	$[-0.001, 0.005]$
$(10, 4, 8, 3, 3)$	0	0	$[0, 0]$
$(3, 5, 2, 2, 8)$	1.7×10^{-11}	2.96×10^{-11}	$[-1.65 \times 10^{-8}, 5.9 \times 10^{-8}]$
$(10, 5, 5, 2, 1)$	0.00405	0.012	$[-0.00337, 0.000473]$

On remarque qu'en fonction des $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$, les résultats peuvent être différents.

7.5 Question 5

on veut choisir les α de sorte que
$(\alpha_1^\star, \alpha_2^\star, \alpha_3^\star, \alpha_4^\star, \alpha_5^\star) = \underset{\alpha}{argmin} E_g[ln(\frac{g(Z)}{h(Z,\alpha)})]$
On note $(\alpha_1^0, \alpha_2^0, \alpha_3^0, \alpha_4^0, \alpha_5^0)$ les valeurs rentrées par l'utilisateur. On a :

$$\underset{\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5}{argmin} \; E_g[ln(\frac{g(Z)}{h(Z, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)})]$$

$$= \underset{\alpha}{argmin} E_{\alpha^0}[ln(\frac{g(Z)}{h(Z,\alpha)})\frac{g(Z)}{h(Z,\alpha^0)}]$$

$$= \underset{\alpha}{argmin} E_{\alpha^0}[-ln(\frac{h(Z,\alpha)}{g(Z)})\frac{g(Z)}{h(Z,\alpha^0)}]$$

$$= \underset{\alpha}{argmin}(E_{\alpha^0}[-ln(h(Z,\alpha))\frac{g(Z)}{h(Z,\alpha^0)}] + E_{\alpha^0}[ln(g(Z))\frac{g(Z)}{h(Z,\alpha^0)}])$$

le second membre ne dépend pas des $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ car g(Z) n'en dépend pas (il dépend des $\lambda_1, \lambda_2, \lambda_3, \lambda_4; \lambda_5$ comme fonction de f)

$$= \underset{\alpha}{argmin} E_{\alpha^0}[-ln(h(Z,\alpha))\frac{g(Z)}{h(Z,\alpha^0)}]$$

$$= \underset{\alpha}{argmin} E_{\alpha^0}[-ln(h(Z,\alpha)))\frac{\mathbf{1}_{\{d(Y)\geq t\}}f(Z,\lambda)}{ph(Z,\alpha^0)}]$$

$$= \underset{\alpha}{argmax} E_{\alpha^0}[\mathbf{1}_{\{d(Z)\geq t\}}ln(h(Z,\alpha))\frac{f(Z,\lambda)}{h(Z,\alpha^0)}]$$

car

$$\underset{f}{argmin} - f = \underset{f}{argmax} f$$

et ca ne depend pas de la constante p
finalement on a bien :

$$(\alpha_1^\star, \alpha_2^\star, \alpha_3^\star, \alpha_4^\star, \alpha_5^\star) = \underset{\alpha}{argmax} E_{\alpha^0}[\mathbf{1}_{\{d(Y)\geq t\}}ln(h(Y,\alpha))\frac{f(Y,\lambda)}{h(Y,\alpha^0)}]$$

7.6 Question 6

On se donne $Y_k = (Y_{k,1}, Y_{k,2}, Y_{k,3}, Y_{k,4}, Y_{k,5})$ pour $k = 1, ..., n$ variables aléatoires indépendantes et identiquement distribuées de densité h(.,α^0) et on va chercher à estimer $(\alpha_1^\star, \alpha_2^\star, \alpha_3^\star, \alpha_4^\star, \alpha_5^\star)$.
on va commencer par estimer l'espérance par :

$$(1/n)\sum_{k=1}^{n} \mathbf{1}_{\{d(Y^k)\geq t\}}\frac{f(Y_k,\lambda)}{h(Y_k,\alpha^0)}ln(h(Y_k,\alpha)$$

qui se réécrit :

$$(1/n)\sum_{k=1}^{n} \mathbf{1}_{\{d(Y^k)\geq t\}}[\prod_{i=1}^{5}\frac{\lambda_i}{\alpha_i^0}exp(-(\lambda_i - \alpha_i^0)Y_{k,i}]ln(h(Y_k,\alpha))$$

On note :
$C_k = \mathbf{1}_{\{d(Y^k)\geq t\}}[\prod_{i=1}^{5} \frac{\lambda_i}{\alpha_i^0} exp(-(\lambda_i - \alpha_i^0)Y_{k,i}]$
Calculons

$$\frac{\partial}{\partial \alpha_j}((1/n)\sum_{k=1}^{n} C_k ln(h(Y_k,\alpha)))$$

$$= (1/n)\sum_{k=1}^{n} C_k[\frac{\partial}{\partial \alpha_j}\prod_{i=1}^{5} \alpha_i exp(-\alpha_i Y_{k,i})]\frac{1}{\prod_{i=1}^{5} \alpha_i exp(-\alpha_i Y_{k,i})}$$

$$= (1/n)\sum_{k=1}^{n} C_k \prod_{i\neq j} \alpha_i exp(-\alpha_i Y_{k,i})\frac{[exp(-\alpha_j Y_{k,j}) - \alpha_j Y_{k,j} exp(-\alpha_i Y_{k,j})]}{\prod_{i=1}^{5} \alpha_i exp(-\alpha_i Y_{k,i})}$$

$$= (1/n)\sum_{k=1}^{n} C_k[\frac{1}{\alpha_j} - Y_{k,j}]$$

le $\hat{\alpha}_j^\star$ que nous cherchons est celui qui annule cette quantité, c'est-à-dire

$$\hat{\alpha}_j^\star = \frac{\sum_{k=1}^{n} \mathbf{1}_{\{d(Y^k)\geq t\}}[\prod_{i=1}^{5} \frac{\lambda_i}{\alpha_i^0} exp(-(\lambda_i - \alpha_i^0)Y_{k,i}]}{\sum_{k=1}^{n} \mathbf{1}_{\{d(Y^k)\geq t\}} Y_{k,j}[\prod_{i=1}^{5} \frac{\lambda_i}{\alpha_i^0} exp(-(\lambda_i - \alpha_i^0)Y_{k,i}]}$$

c'est bien un maximum car en calculant la dérivée seconde on trouve :

$$\frac{\partial^2}{\partial \alpha_j^2}((1/n)\sum_{k=1}^{n} C_k ln(h(Y_k,\alpha))) = -(1/n)\sum_{k=1}^{n} \frac{C_k}{\alpha_j^2}$$

qui est négative car les C_k sont positifs. Finalement la dérivée seconde est bien négative et les $\hat{\alpha}_j^\star$ sont bien des maximum.
On peut remarquer que les $\hat{\alpha}_j^\star$ dépendent des α_i^0 saisis par l'utilisateur. On remarquera également que peu importe les α_i^0 saisis la méthode reste plus optimale que la méthode de Monte Carlo classique.

7.7 Question 7

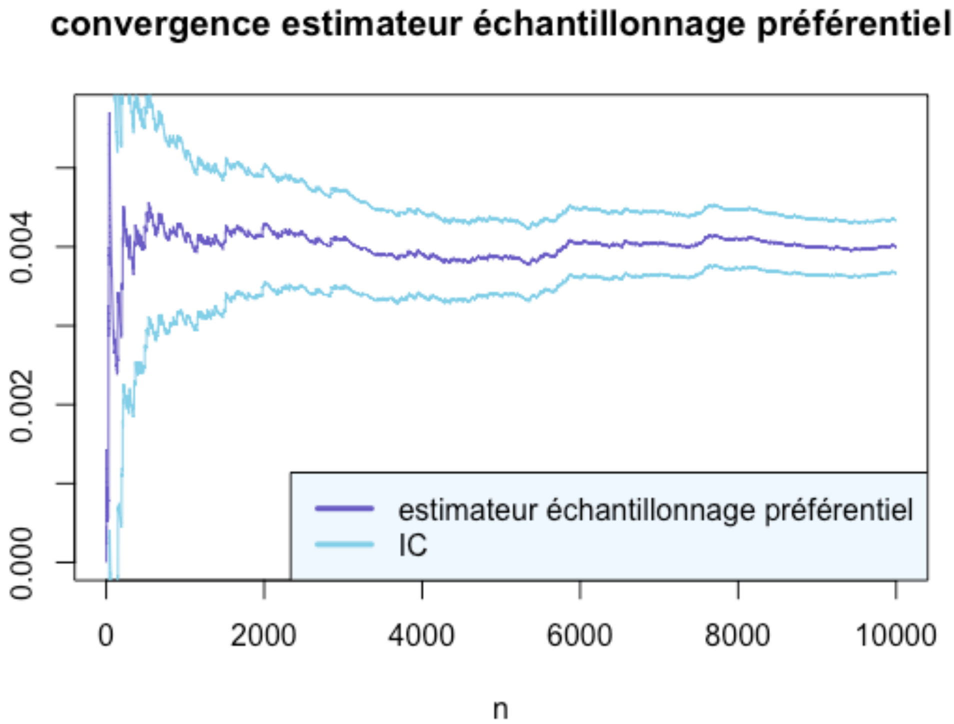
On reprend l'estimateur d'échantillonnage préférentiel de la question 4

$$\hat{p}_n(\hat{\alpha}^\star) = (1/n)\sum_{k=1}^{n} \frac{f(Y^k,\lambda)}{h(Y^k,\hat{\alpha}^\star)}\mathbf{1}_{\{d(Y^k)\geq t\}}$$

en posant cette fois-ci :

1. $\hat{\alpha}^\star = (\hat{\alpha}_1^\star, \hat{\alpha}_2^\star, \hat{\alpha}_3^\star, \hat{\alpha}_4^\star, \hat{\alpha}_5^\star)$
2. $Y^k = (Y_1^k, Y_2^k, Y_3^k, Y_4^k, Y_5^k)$ un vecteur de variables aléatoires simulées selon la loi de densité $h(Y^k, \hat{\alpha}^\star)$

Après 10.000 réalisations, $\hat{p}_n(\hat{\alpha}^\star) = 0.00413$ avec une variance égale à 0.000302 et $IC_{0.975} = [0.00379, 0.00447]$.



7.8 Question 8

7.8.1 a

Soient $U_{1,1}, ..., U_{1,5}, U_{2,1}, ..., U_{n,5}$ des variables aléatoires indépendantes et identiquement distribuées suivant la loi uniforme sur [0,1].
On pose :

$$\rho(U_{i,1}, ..., U_{i,5}) = min(\frac{1}{\hat{\alpha}_1^\star} ln(U_{i,1}) - \frac{1}{\hat{\alpha}_4^\star} ln(U_{i,4}); \frac{1}{\hat{\alpha}_{i\,1}^\star} ln(U_{i,1}) - \frac{1}{\hat{\alpha}_3^\star} ln(U_{i,3}) - \frac{1}{\hat{\alpha}_5^\star} ln(U_{i,5});$$

$$\frac{1}{\hat{\alpha}_2^\star} ln(U_{i,2}) - \frac{1}{\hat{\alpha}_5^\star} ln(U_{i,5}); \frac{1}{\hat{\alpha}_2^\star} ln(U_{i,2}) - \frac{1}{\hat{\alpha}_3^\star} ln(U_{i,3}) - \frac{1}{\hat{\alpha}_4^\star} ln(U_{i,4}))$$

On a :

$$\frac{1}{n}\sum_{k=1}^{n}\prod_{i=1}^{5}\frac{\lambda_i}{\hat{\alpha}_i^\star}U_{i,k}^{\frac{\lambda_i}{\hat{\alpha}_i^\star}-1}\mathbf{1}_{\{\rho(U_{i,1},\ldots,U_{i,5})\geq t\}} \xrightarrow[n\to+\infty]{ps} \mathbf{E}_U[\prod_{i=1}^{5}\frac{\lambda_i}{\hat{\alpha}_i^\star}U_i^{\frac{\lambda_i}{\hat{\alpha}_i^\star}-1}\mathbf{1}_{\{\rho(U_1,\ldots,U_5)\geq t\}}]$$

on effectue le changement de variable $U_i = exp(-\alpha^\star Y_i)$

$$\mathbf{E}_U[\prod_{i=1}^{5}\frac{\lambda_i}{\hat{\alpha}_i^\star}U_i^{\frac{\lambda_i}{\hat{\alpha}_i^\star}-1}\mathbf{1}_{\{\rho(U_i,\ldots,U_i)\geq t\}}] = \int_{[0,1]^5}\prod_{i=1}^{5}\frac{\lambda_i}{\hat{\alpha}_i^\star}u_i^{\frac{\lambda_i}{\hat{\alpha}_i^\star}-1}\mathbf{1}_{\{\rho(u_1,\ldots,u_5)\geq t\}}du_1du_2du_3du_4du_5$$

qui devient aprés le changement de variable (en notant $y = (y_1, y_2, y_3, y_4, y_5)$) :

$$\int_{\mathbb{R}_\star^5}\prod_{i=1}^{5}\lambda_i exp(-\lambda_i y_i)\mathbf{1}_{\{d(y)\geq t\}}dy$$

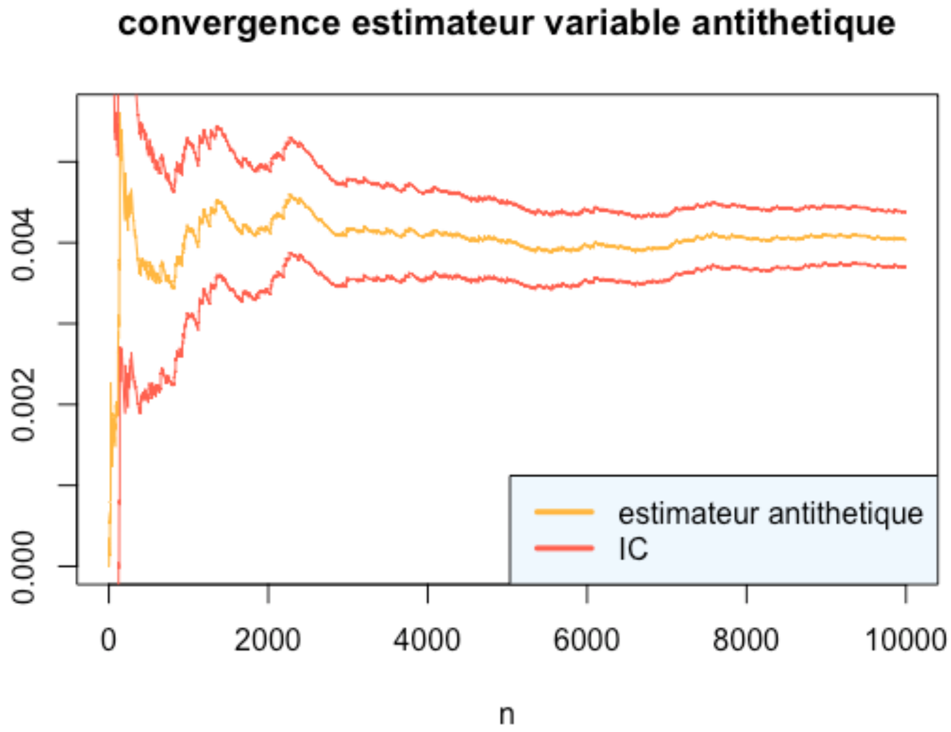
$$= \int_{\mathbb{R}_\star^5} f(y,\lambda)\mathbf{1}_{\{d(y)\geq t\}}dy = \mathbf{E}_f[\mathbf{1}_{\{d(y)\geq t\}}] = p$$

7.8.2 b

On va choisir comme variable antithétique 1-U avec U qui suit une loi uniforme sur [0,1] (on a vu dans le cours que cette variable suit bien la même loi que U). Alors on a un estimateur donné par :

$$\hat{p}_n^{(1)} = (1/n)\sum_{k=1}^{n}\frac{\prod_{i=1}^{5}\frac{\lambda_i}{\hat{\alpha}_i^\star}U_{i,k}^{\frac{\lambda_i}{\hat{\alpha}_i^\star}-1}\mathbf{1}_{\{\rho(U_{i,1},\ldots,U_{i,5})\geq t\}} + \prod_{i=1}^{5}\frac{\lambda_i}{\hat{\alpha}_i^\star}(1-U_{i,k})^{\frac{\lambda_i}{\hat{\alpha}_i^\star}-1}\mathbf{1}_{\{\rho(1-U_{i,1},\ldots,1-U_{i,5})\geq t\}}}{2}$$

Après 10.000 réalisations, $\hat{p}_n^{(1)} = 0.0039$ avec une variance égale à 0.000133 et $IC_{0.975} = [0.00357, 0.00424]$



7.8.3 c

En comparant les simulations des questions précédentes, Dans un premier temps on remarque que :

$$\mathbb{V}(\hat{p}_n^{(1)}) = 0.000133 < \mathbb{V}(\hat{p}_n(\hat{\alpha}^\star)) = 0.000302 < \mathbb{V}(\hat{p}_n) = 0.00411$$

La méthode de la variable antithétique est de meilleurs qualité que les deux autres, en effet il divise la variance par 2 par rapport à l'échantillonnage préférentiel et par 3 par apport à MC classique. Estimer par la variable antithétique permet dan ce cas la de réduire l'erreur d'approximation. On va analyser les temps d'exécution pour voir si le coup en calcul à payer est important pour obtenir une meilleure précision.

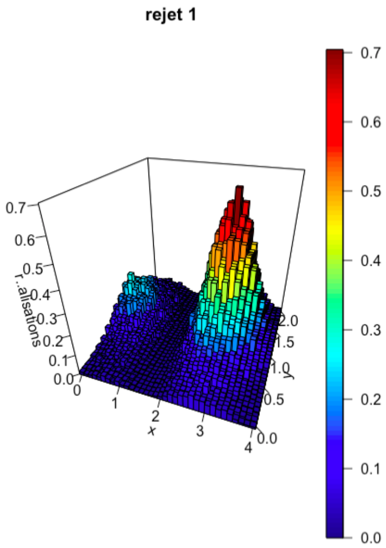
fonction étudiée	nombre de simulations	temps moyen (en ms)
estimateur de Monte Carlo	1.000	1.8
	10.000	18.59
	100.000	246.41
Echantillonnage préférentiel	1.000	2.04
	10.000	21.85
	100.000	280.3
Antithétique	1.000	3.93
	10.000	40.33
	100.000	542.34

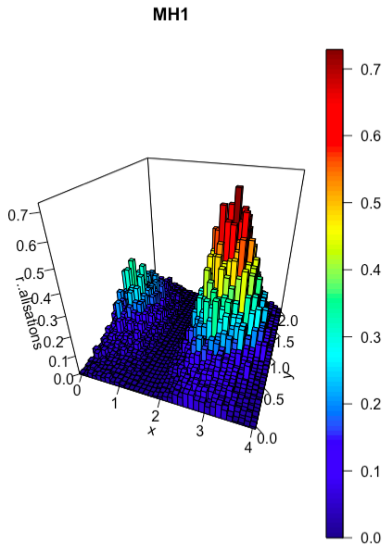
On remarque le temps de calcul pour l'échantillonnage préférentiel est très proche de celui de Monte Carlo, pour une bien meilleure précision. Cependant, celui de l'estimateur de la variable antithétique est plus important. Pas étonnant sachant que la méthode de la variable antithétique fait deux fois plus de calcule que la méthode de Monte Carlo classique.

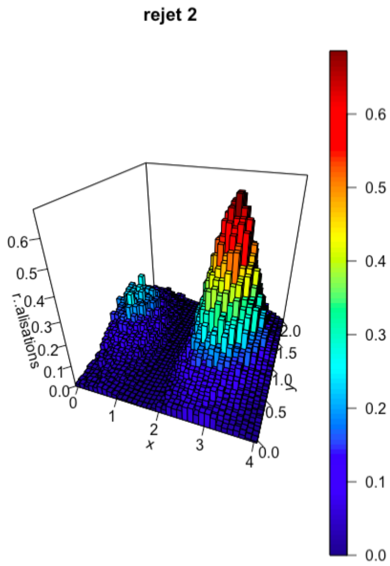
Troisième partie
ANNEXE

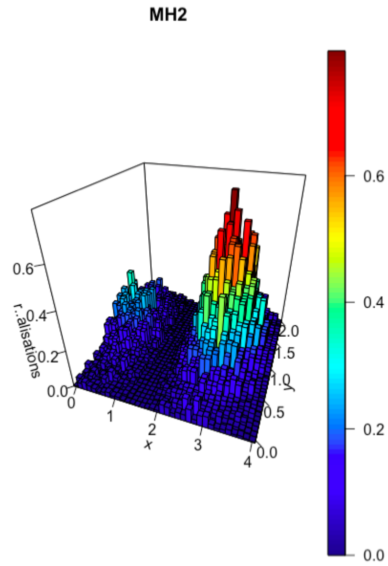
8 Exercice 1, question 5.

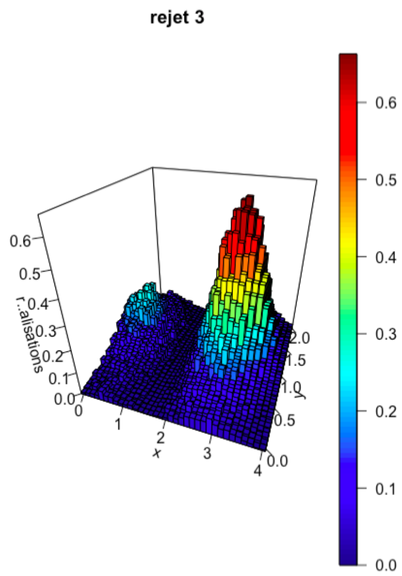
Nous allons mettre côte à côte les distributions obtenues

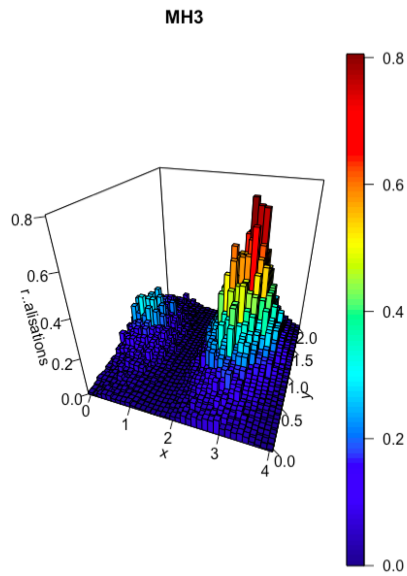


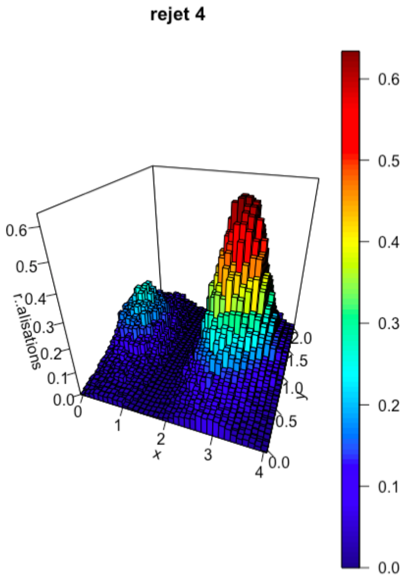


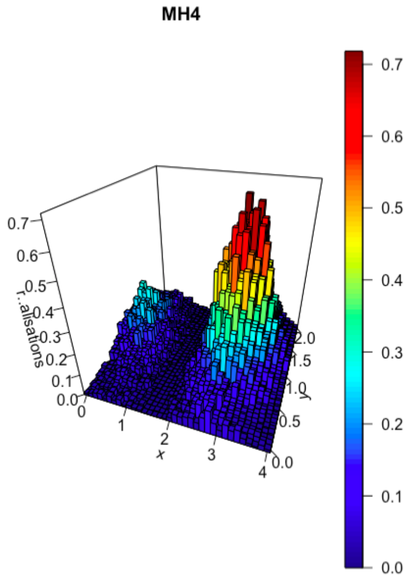

rejet 2
r..alisations
x
y
0.0
0.1
0.2
0.3
0.4
0.5
0.6


MH2
r..alisations
x
y
0.0
0.2
0.4
0.6


rejet 3
r..alisations
x
y
0.0
0.1
0.2
0.3
0.4
0.5
0.6


MH3
r..alisations
x
y
0.0
0.2
0.4
0.6
0.8





On remarque sur les graphiques une meilleure précision pour la méthode du rejet qui est d'aileurs plus rapide à exécuter quand on compare les valeurs des questions 3 et 4.

9 constante de normalisation pour f

On note c la constante de normalisation. On sait que $f = c\tilde{f}$ et que $\tilde{f} \leq Mg$. Donc

$$f = c\tilde{f} \leq cMg = \tilde{M}g \text{ où } \tilde{M} = cM$$

Donc $c = \dfrac{\tilde{M}}{M} = \dfrac{\mathbb{E}[T]}{M}$
Or, on ne connait pas l'espérance de T, on va donc l'estimer par la méthode de Monte Carlo pour obtenir une estimation de c. On obtient :

$$\sqrt{n}(\mathbb{E}\hat{[T]} - \mathbb{E}[T]) \xrightarrow[n\to+\infty]{loi} \mathcal{N}(O, \mathbb{V}(T))$$

On obtient donc pour c :

$$\sqrt{n}(\hat{c} - c) \xrightarrow[n\to+\infty]{loi} \mathcal{N}(O, \mathbb{V}(T)/M^2)$$

On ne connait pas la variance de T. Cependant on sait que T suit une loi géométrique de 1/p, où p est la probabilité d'acceptation de la méthode du rejet et on sait estimer cette valeur dans le rejet. Donc on sait estimer la variance qui vaut $\mathbb{V}\hat{(T)} = \dfrac{1-\hat{p}}{\hat{p}^2}$. On peut en conclure par le TCL et Slutsky :

$$IC_{0.975}(c) = \left[\hat{c} - \sqrt{\frac{\mathbb{V}\hat{(T)}}{n}}, \hat{c} + \sqrt{\frac{\mathbb{V}\hat{(T)}}{n}}\right]$$