

Université Paris-Dauphine

M1 Mathématiques Appliquées

Projet : Méthodes de Monte-Carlo

Cohen Ethan, Bedoucha Gary

Janvier 2019

Table des matières :

1. Exercice 1 : Simulation de variables aléatoires	3
2. Exercice 2 : Méthode de Monte-Carlo	16
3. Annexe : Code R	39

Exercice 1 : Simulation de variables aléatoires.

Question 1.

On souhaite montrer qu'appliquer l'algorithme de rejet à la fonction \tilde{f} qui est proportionnelle à f permet de simuler suivant la densité f .

En effet, appliquons la méthode de rejet à \tilde{f} : il existe une densité instrumentale g et $M \geq 1$ tel que :

$$\forall x \in \mathbb{R}^2, \tilde{f}(x) \leq M g(x)$$

$$\Leftrightarrow \frac{f(x)}{c} \leq M g(x)$$

$$\Leftrightarrow f(x) \leq c M g(x)$$

où c est la constante de normalisation telle que $f(x) = c \tilde{f}(x)$.

En effet, f étant une densité on a $c = \frac{1}{\int_{\mathbb{R}^2} \tilde{f}(x,y) dx dy}$. On pose alors $\tilde{M} = cM$.

On prend maintenant $U \sim U(0,1)$ et u une de ses réalisations.

On a l'équivalence suivante :

$$u \leq \frac{f(x)}{\tilde{M} g(x)} \Leftrightarrow u \leq \frac{\tilde{f}(x)}{M g(x)}$$

On conclut alors, qu'appliquer la méthode de rejet à f avec \tilde{M} et g est équivalent à l'appliquer à \tilde{f} avec M et g . On va alors, pour les questions suivantes, utiliser la fonction \tilde{f} et lui appliquer la méthode de rejet ce qui permettra d'obtenir des réalisations suivant la loi de densité f .

Remarque 1 : On peut calculer empiriquement la constante de normalisation c de 2 différentes manières : (on aura besoin de la valeur approximative de c pour la suite)

Méthode 1.a : Comme $\int_{\mathbb{R}^2} \tilde{f}(x, y) dx dy = \frac{1}{c}$, on calcule l'intégrale de \tilde{f} par une méthode de Monte-Carlo classique en l'interprétant comme une espérance sous une loi $U(0,4) \times U(0,2)$ par exemple.

Autrement dit :

$$\begin{aligned} & \int_{\mathbb{R}^2} \tilde{f}(x, y) dx dy \\ &= \int_{\mathbb{R}^2} \frac{\cos^2(x) + 2\sin^2(y) \cos^4(x)}{1 + 4(y-1)^2} e^{-\frac{(x-2)^2}{2}} 1_{\{0 \leq x \leq 4\}} 1_{\{0 \leq y \leq 2\}} dx dy \\ &= \int_{\mathbb{R}^2} \frac{\cos^2(x) + 2\sin^2(y) \cos^4(x)}{1 + 4(y-1)^2} e^{-\frac{(x-2)^2}{2}} \frac{1}{4} \cdot 4 1_{\{0 \leq x \leq 4\}} \frac{1}{2} \cdot 2 1_{\{0 \leq y \leq 2\}} dx dy \\ &= E_g [h(x, y)] \end{aligned}$$

$$\text{où } h(x, y) = 8 \cdot \frac{\cos^2(x) + 2\sin^2(y) \cos^4(x)}{1 + 4(y-1)^2} e^{-\frac{(x-2)^2}{2}}$$

g est la densité de la loi uniforme $(0,4) \times$ uniforme $(0,2)$.

L'estimateur obtenu par la méthode de Monte-Carlo est donc :

$$\tilde{f}^{(n)} = \frac{1}{n} \sum_{k=1}^n h(X_1^{(k)}, X_2^{(k)})$$

où les $X_1^{(k)}, X_2^{(k)}$ sont respectivement des réalisations de lois uniforme $U(0,4)$ et $U(0,2)$.

Finalement, $c = \frac{1}{\tilde{f}^{(n)}}$

[1] 0.4363217

Méthode 1.b : On peut également faire de même en interprétant cette fois l'intégrale de \tilde{f} comme étant une espérance sous une loi $N(2,1) \times U(0,2)$.

L'estimateur obtenu par la méthode de Monte-Carlo est donc :

$$\tilde{f}^{(n)} = \frac{1}{n} \sum_{k=1}^n h(X_1^{(k)}, X_2^{(k)})$$

où cette fois $h(x,y) = 2\sqrt{2\pi} \frac{\cos^2(x) + 2\sin^2(y) \cos^4(x)}{1 + 4(y-1)^2} 1_{\{0 \leq x \leq 4\}}$

et les $X_1^{(k)}, X_2^{(k)}$ sont respectivement des réalisations de lois uniforme $N(2,1)$ et $U(0,2)$.

De même, $c = \frac{1}{\tilde{f}^{(n)}}$

[1] 0.4340545

Méthode 2 : On utilise le fait que : $p = \mathbb{P}(\text{Acceptation del'algoderejet}) = \frac{1}{\bar{M}} = \frac{1}{cM}$

En pratique, $p = \frac{n}{K+n} = \frac{n}{N}$

Où :

- n est le nombre de fois où l'on accepte l'algorithme de rejet (ce n est choisi par l'utilisateur car celui-ci veut créer un vecteur constitué de n réalisations suivant la loi de densité f pour « recréer » la densité de la loi).

- K est le nombre de fois où l'on rejette l'algorithme de rejet c'est à dire le nombre de fois que l'on "rentre" dans la boucle $while U(Y_k) \geq \frac{\tilde{f}(Y_k)}{Mg(Y_k)}$.

- N est le nombre total de fois où on compare U et $\frac{\tilde{f}}{Mg}$ dans l'algorithme de rejet = n + K

Ainsi, $p = \frac{1}{cM} \Rightarrow c = \frac{1}{pM}$

[1] 0.4332513

Questions 2 et 3 : Méthode de rejet pour obtenir des réalisations suivant la densité f.

Ne sachant pas simuler des réalisations suivant la loi de densité f, on va utiliser l'algorithme de rejet à cet effet.

Rappel de cours :

Soient $(U_n)_{n \geq 1}$ une suite de variables *i.i.d.* de loi $\mathcal{U}([0, 1])$ et $(\mathbf{Y}_n)_{n \geq 1} = (Y_{1,n}, \dots, Y_{d,n})_{n \geq 1}$ une suite de variables *i.i.d.* de loi de densité g telles que ces deux suites soient indépendantes. Alors, pour T défini par

$$T := \inf\{n \geq 1 : U_n \leq \alpha(\mathbf{Y}_n)\}, \quad \text{où } \alpha(\mathbf{Y}_n) := \frac{f(\mathbf{Y}_n)}{Mg(\mathbf{Y}_n)},$$

\mathbf{Y}_T suit la loi de densité f. Autrement dit pour simuler $\mathbf{X} \sim f$, il suffit de simuler

$$\mathbf{Y} \sim g \quad \text{et} \quad U | \mathbf{Y} = \mathbf{y} \sim \mathcal{U}([0, Mg(\mathbf{y})]),$$

jusqu'à ce que $u < f(\mathbf{y})$.

Source : Polycopié de cours de Mr. Stoehr.

De ce fait, connaître la densité instrumentale g et "sa" valeur M associée détermine entièrement l'algorithme de rejet.

On va donc utiliser 5 différentes méthodes de rejet pour simuler suivant f en faisant apparaître à chaque fois un g et un M associé différent.

Méthode 1 : *Rejet avec une densité instrumentale g suivant une loi uniforme (0,2) x uniforme (0,4).*

$$\text{On a } \forall (x, y) \in \mathbb{R}^2, \cos^2(x) + 2 \sin^2(y) \cos^4(x) \leq 3$$

$$\cdot \frac{1}{1+4(y-1)^2} \leq 1$$

$$\cdot e^{-\frac{(x-2)^2}{2}} \leq 1$$

$$\text{D'où on obtient : } \tilde{f}(x, y) \leq 3 \times 4 \times \frac{1}{4} 1_{\{x \in [0,4]\}} \times \frac{1}{2} \times 2 \times 1_{\{y \in [0,2]\}}$$

$$= 24 \times \frac{1}{4} 1_{\{x \in [0,4]\}} \times \frac{1}{2} 1_{\{y \in [0,2]\}}$$

$$= M \times g(x, y)$$

Ainsi, $\tilde{f}(x, y) \leq M \times g(x, y)$

où :

. $g(x, y) = \frac{1}{4} 1_{\{x \in [0, 4]\}} \times \frac{1}{2} 1_{\{y \in [0, 2]\}} = \text{densité de la loi } U(0, 4) \times U(0, 2)$

. $M=24$

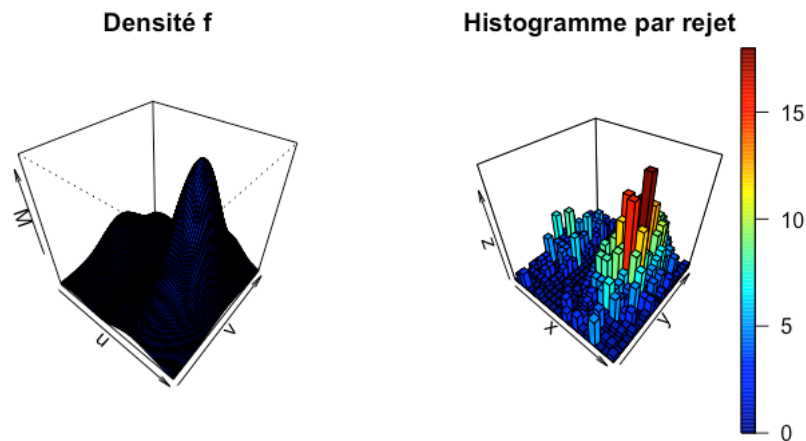


Figure 1 : Densité de f et histogramme pour la méthode rejet 1.

Remarque 2 : Pour pouvoir tracer la densité f, nous avons eu besoin d'avoir accès à la valeur de c qu'on a calculé précédemment dans la Remarque 1 (nous avons tracé la fonction $\tilde{f}(x, y) \times c$).

Dans les 4 méthodes suivantes, le raisonnement est quasi-identique : ces méthodes diffèrent simplement quant au choix de la densité instrumentale g et de la valeur M associée.

C'est pourquoi nous indiquerons désormais pour chaque méthode seulement la densité instrumentale g et la valeur M.

Méthode 2 : Rejet avec une densité instrumentale g suivant une loi Normale (2,1) x Uniforme (0,2).

$$\tilde{f}(x, y) \leq M \times g(x, y)$$

où :

. $g(x, y) = \text{densité de la loi } N(2, 1) \times U(0, 2)$

. $M = 6\sqrt{2\pi}$

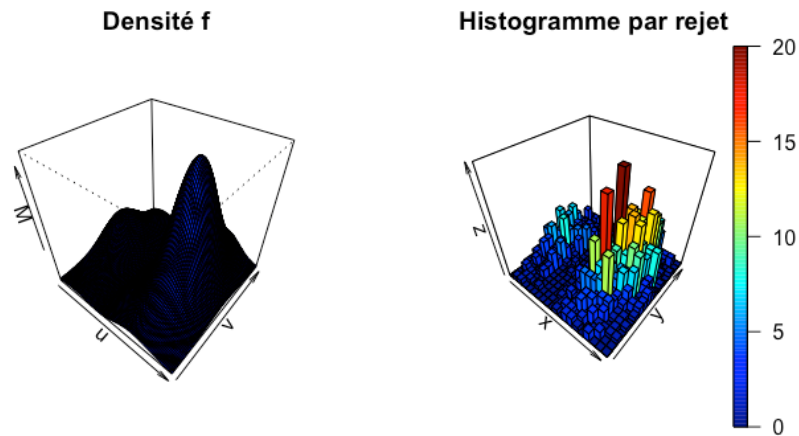


Figure 2 : Densité de f et histogramme pour la méthode rejet 2.

Méthode 3 : Rejet avec une densité instrumentale g suivant une loi Uniforme (0,4) x Cauchy $(1, \frac{1}{2})$.

$$\tilde{f}(x, y) \leq M \times g(x, y)$$

où :

. $g(x, y)$ = densité de la loi $U(0,4) \times C\left(1, \frac{1}{2}\right)$

. $M = 6\pi$

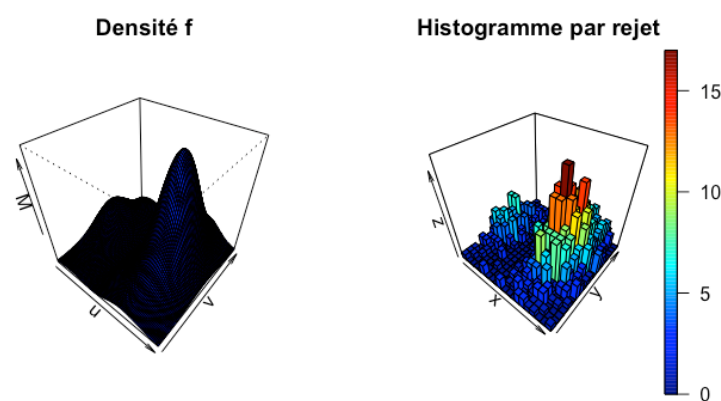


Figure 3 : Densité de f et histogramme pour la méthode rejet 3.

Méthode 4 : Rejet avec une densité instrumentale g suivant une loi Normale (2,1) x Cauchy $(1, \frac{1}{2})$.

$$\tilde{f}(x, y) \leq M \times g(x, y)$$

où :

. $g(x, y)$ = densité de la loi $N(2, 1) \times C\left(1, \frac{1}{2}\right)$

$$. M = \frac{3}{\sqrt{2}} \pi^{3/2}$$

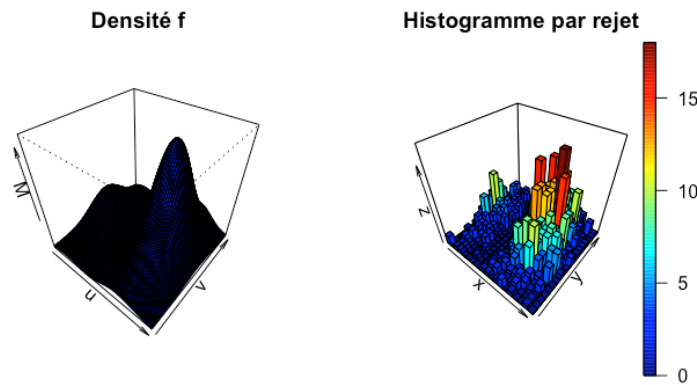


Figure 4 : Densité de f et histogramme pour la méthode rejet 4.

Méthode 5 : Rejet avec une densité instrumentale g suivant une loi Normale (2,1) tronquée [0,4] (noté NT (2,1)) x Cauchy $(1, \frac{1}{2})$ tronquée [0,2] (noté CT $(1, \frac{1}{2})$).

$$\tilde{f}(x, y) \leq M \times g(x, y)$$

où :

. $g(x, y)$ = densité de la loi $NT(2, 1) \times CT\left(1, \frac{1}{2}\right)$

$$. M = \frac{3}{\sqrt{2}} \pi^{3/2} \times [FC(2) - FC(0)] \times [FN(4) - FN(0)]$$

avec : FC = Fonction de répartition de la loi de Cauchy $(1, \frac{1}{2})$

FN = Fonction de répartition de la loi de Normale (2, 1)

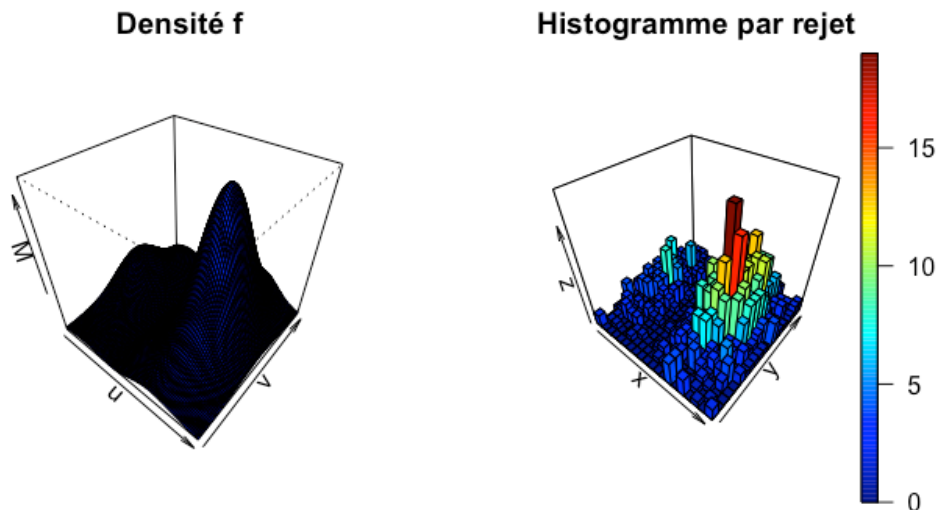


Figure 5 : Densité de f et histogramme pour la méthode rejet 5.

Remarque 3 : Comparaison des temps d'exécution des 5 différentes méthodes de l'algorithme de rejet.

Dans le code présent en annexe, nous avons calculé le temps d'exécution correspondant à chaque méthode dans le but de comparer leur performance.

En théorie, la constante M correspond au nombre d'essai moyen avant l'acceptation d'une réalisation venant de f. De ce fait, plus la valeur de M est petite, plus le temps d'exécution de l'algorithme est petit et la méthode de rejet est d'autant plus performante.

Vérification en pratique :

Voici ci-dessous les constantes M (ici noté C_i) correspondantes à la méthode i.

```
C=c(C1,C2,C3,C4,C5)
```

```
C
```

```
[1] 24.000000 15.039770 18.849556 11.812207 7.946812
```

Voici ci-dessous les temps d'exécution T (ici noté Tdiff i) correspondant à la méthode i.

```
T=c(Tdiff1,Tdiff2,Tdiff3,Tdiff4,Tdiff5)
T
```

```
Time differences in secs
[1] 1.7049630 1.1850860 1.3064158 0.9762361 0.7200470
```

Finalement, la pratique confirme bien la théorie, plus la valeur M est petite, plus le temps d'exécution T est petit.

On peut donc ordonner les 5 différentes méthodes de rejet du plus performant au moins performant :

NT(2,1) x CT(1,0.5) > N(2,1) x C(1,0.5) > N(2,1) x U(0,2) > U(0,4) x C(1,0.5) > U(0,4) x U(0,2)

Algorithme de Metropolis-Hastings :

On peut observer dans un premier temps que l'algorithme de Metropolis-Hastings peut s'utiliser en négligeant les constantes de la densité correspondante à la loi dont on souhaite simuler des réalisations.

En effet, en gardant les notations de l'énoncé et en observant la forme de la fonction α , si on a $c \in \mathbb{R}$ tel que $f = c\tilde{f}$ alors :

$$\min\left(1, \frac{f(x)g(y)}{f(y)g(x)}\right) = \min\left(1, \frac{c\tilde{f}(x)g(y)}{c\tilde{f}(y)g(x)}\right) = \min\left(1, \frac{\tilde{f}(x)g(y)}{\tilde{f}(y)g(x)}\right)$$

De ce fait, appliquer l'algorithme avec f ou \tilde{f} revient au même et nous permet de générer des chaînes de Markov $(X^{(t)})_{t \geq 1}$ irréductibles, apériodiques et de loi stationnaire de densité f.

a. Metropolis-Hastings 1

Dans un premier temps, nous avons implémenté l'algorithme de Metropolis-Hastings 1 naïvement et on remarque que la chaîne de Markov $(X^{(t)})_{t \geq 1}$ pour n=1000 termes, nous permet d'obtenir un histogramme qui "ressemble" peu (ou plutôt pas assez) à la densité f.

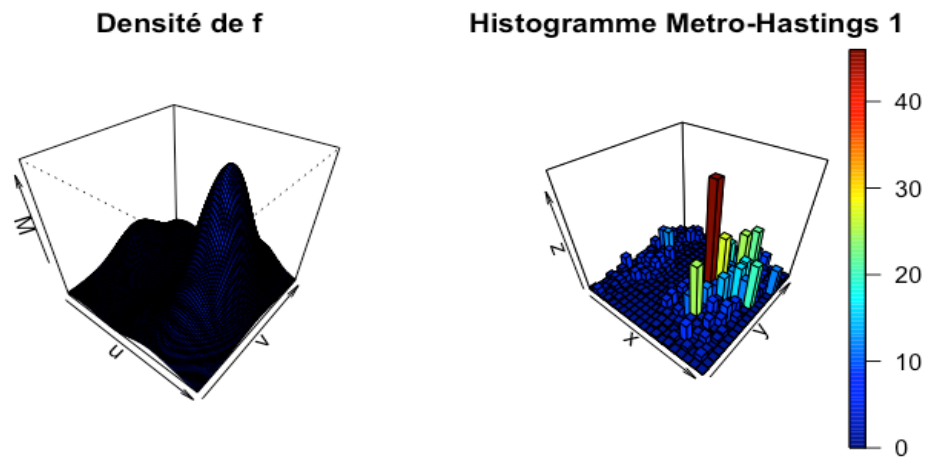


Figure 6 : Densité de f et histogramme pour la méthode de Metropolis-Hastings.

En fait, une des explications possibles de ce phénomène est que l'algorithme de Metropolis-Hastings ne donne pas assez de réalisations distinctes sûrement dû à la petite valeur de α (on reste donc longtemps "bloqué" à $X^{(t+1)} = X^{(t)}$ ce qui arrive avec probabilité $1-\alpha$ qui est proche de 1 quand α petit).

D'autre part, compte tenu du fait que la chaîne de Markov produite, est de loi stationnaire de densité f (c'est à dire qu'à partir d'un certain rang, les valeurs de $(X^{(t)})_{t \geq 1}$ correspondent à des réalisations de la loi dont la densité est f), il est préférable de ne pas prendre en compte les premières valeurs de la suite.

On est donc amené à implémenter un algorithme de Metropolis-Hastings 2 qui corrige ces problèmes en vue de nous permettre d'obtenir un histogramme qui coïncide avec la densité f .

b. Metropolis-Hastings 2

Une solution aux problèmes mentionnés ci-dessus pourrait être de faire tourner n fois l'algorithme et de prendre à chaque fois la millièème réalisation $X^{(1000)} = (X_1^{(1000)}, X_2^{(1000)})$.

On a ainsi créé une matrice (de taille $n \times 2$) constituée des " n " millièmes observations ...

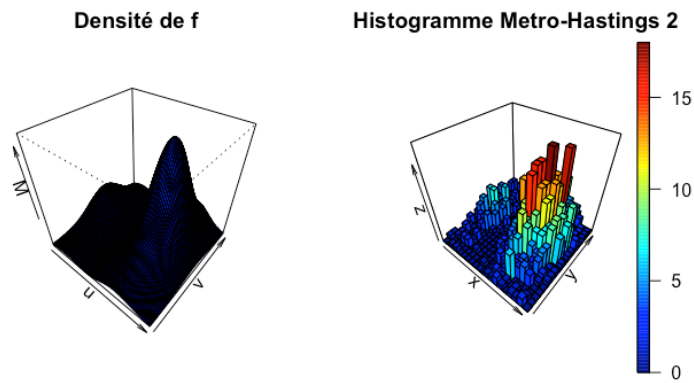


Figure 7 : Densité de f et histogramme pour la méthode de Metropolis-Hasting Version 2

Cette fois, on obtient bien un histogramme qui coïncide plus avec la densité f.

Question 5 : Comparaison méthode de rejet et méthode de Metropolis-Hastings 2

Remarque 4 : Pour comparer les 2 méthodes, on a choisi de prendre l'algorithme de rejet le plus performant, c'est à dire celui ayant pour densité instrumentale g suivant une loi NT (2,1) x CT $(1, \frac{1}{2})$.

a. Comparaison par histogramme

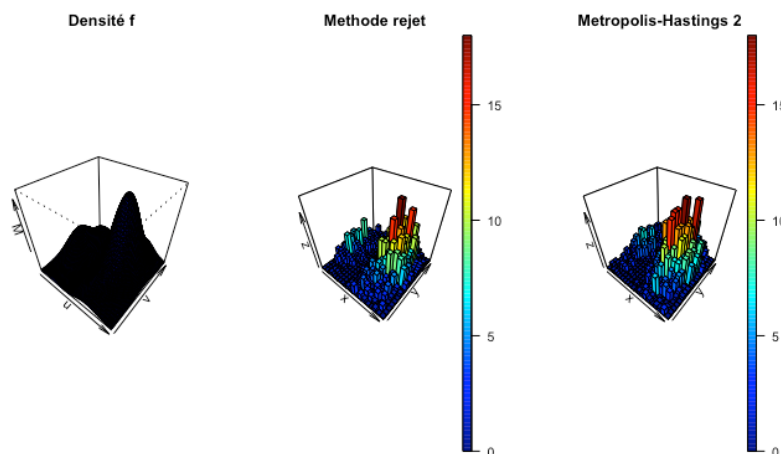


Figure 8 : Densité de f, histogramme par la méthode de Metropolis-Hasting 2 et la "meilleure" méthode de rejet (méthode 5).

⇒ On peut difficilement comparer à l'aide des histogrammes.

b. Comparaison par Q-Q Plot

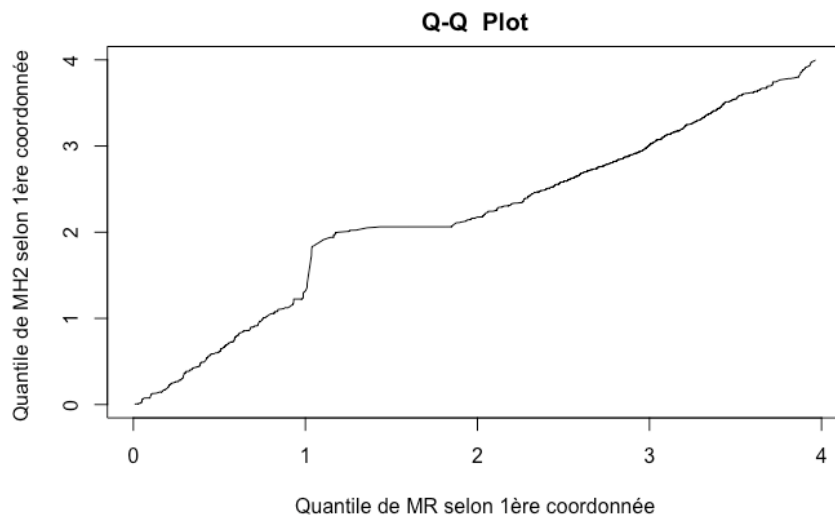


Figure 9 : QQ-Plot de la méthode de rejet 5 et Metropolis-Hasting selon la 1ere coordonnée.

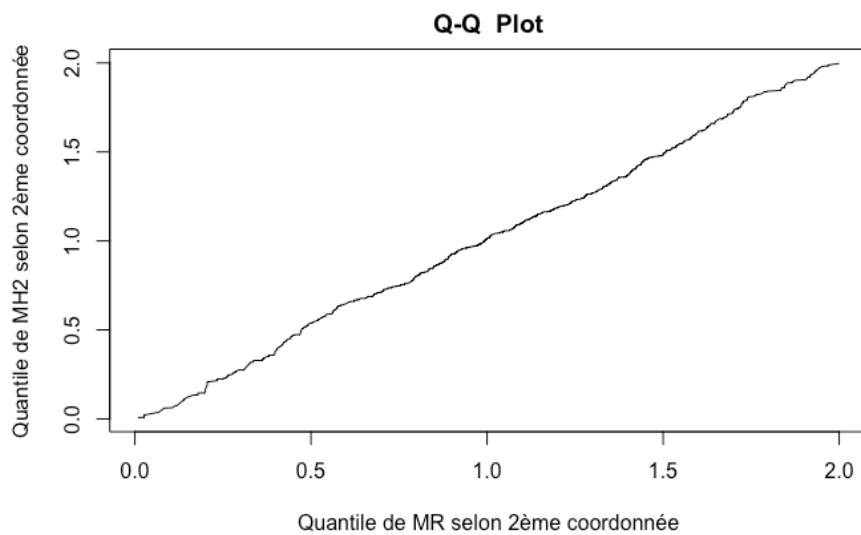


Figure 10 : QQ-Plot de la méthode de rejet 5 et Metropolis-Hasting selon la 2ème coordonnée.

On observe que selon la 2^{ème} coordonnée, les Q-Q Plot des deux méthodes sont pratiquement identiques.

Néanmoins, comme on peut le voir dans le Q-Q Plot selon la 1^{ère} coordonnée, les deux méthodes diffèrent dans leur vitesse de convergence de leur 1^{ère} coordonnée vers des réalisations de la loi dont la densité est la 1^{ère} composante de f : on n'obtient pas une droite ce qui signifie que les quantiles obtenus par ces 2 méthodes sont différents.

Or les quantiles caractérisant la loi, les lois simulées respectivement par les 2 méthodes sont différentes.

D'autre part, on a démontré mathématiquement en cours que Y_T suit la loi dont la densité est f , on déduit que le problème vient de l'algorithme de Metropolis-Hastings qui est donc moins performant.

Conclusion :

Bien que la méthode de Metropolis-Hasting soit beaucoup plus efficace du point de vue computationnel que la méthode d'acceptation-rejet, sa vitesse de convergence vers la fonction à simuler est directement liée au choix de la fonction instrumentale g .

Dans ce cas-ci, l'algorithme de Metropolis-Hastings est moins performant que l'algorithme du rejet.

Exercice 2 : Méthode de Monte-Carlo.

Partie 1 : Estimation de δ

Question 1.

La fonction de densité d'une variable de Weibull de paramètre (λ, k) est :

$$f_{\lambda,k}(x) = \left(\frac{k}{\lambda}\right) \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}$$

On prend dans cette question $k=2$ et $\lambda=1$. On obtient ainsi la fonction de densité suivante :

$$f(x) = 2xe^{-2x}$$

1.a On veut estimer $\delta = \mathbb{P}[X_1 + X_2 + X_3 \geq t] = \mathbb{P}[Z \geq t]$ par une méthode de Monte-Carlo classique avec $Z \sim W(2,1)$.

On réécrit $\delta = E_f[1_{\{Z \geq t\}}] = E_f[h(Z)]$ où $h(z) = 1_{\{z \geq t\}}$.

Soit $(Z_k)_{k \geq 1}$ une suite de variables aléatoires i.i.d suivant une loi de Weibull $(2,1)$.

On estime δ par l'estimateur de Monte-Carlo classique :

$$\hat{\delta}_n = \frac{1}{n} \sum_{k=1}^n h(Z_k)$$

L'intervalle de confiance de niveau $1-\alpha$ avec $\alpha=5\%$ est alors de la forme :

$$\text{IC}(1-\alpha) = \left[\hat{\delta}_n \pm q_{1-\frac{\alpha}{2}} \sqrt{\frac{V(h(Z))}{n}} \right]$$

Où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1-\frac{\alpha}{2}$ d'une $N(0,1)$.

Remarque : par la loi forte des grands nombres, on a bien

$$\hat{\delta}_n \xrightarrow{ps} E_f[h(z)] = \delta$$

En effet, les conditions pour appliquer la LFGN sont vérifiées :

- $(h(Z_k))_{k \geq 0}$ est une suite de variables aléatoires dans L^1 .
- $(h(Z_k))_{k \geq 0}$ est une suite de variables aléatoires identiques.
- $(h(Z_k))_{k \geq 0}$ est une suite de variables aléatoires indépendante par transformation mesurables de variables aléatoires indépendantes.

Dans la suite de l'exercice, on appliquera plusieurs fois ce même procédé sans le mentionner.

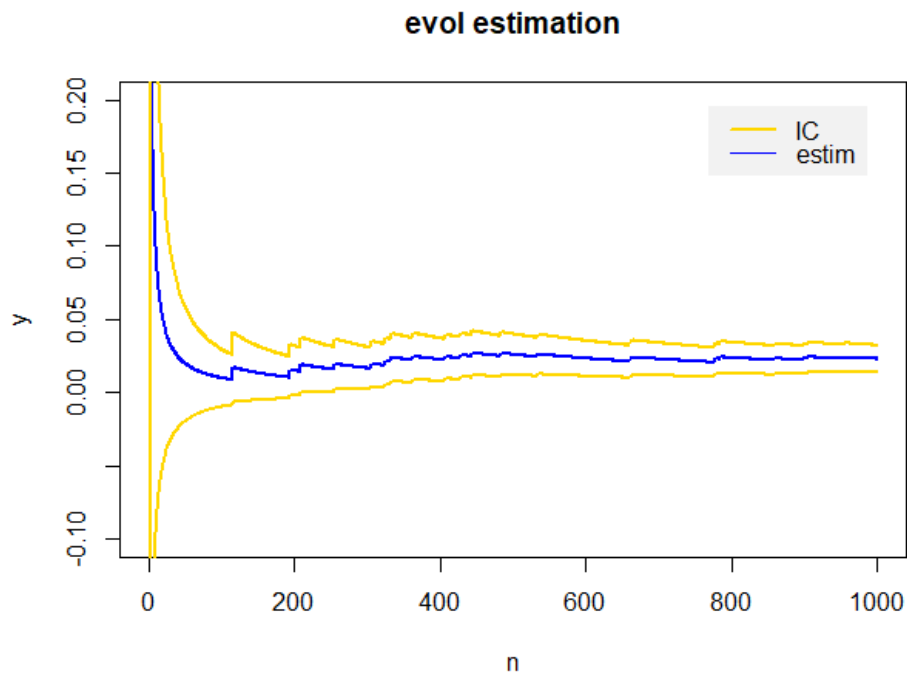


Figure 11 : évolution de l'estimateur ainsi que l'intervalle de confiance.

```

value      var born_inf_IC born_sup_IC
10000 0.019 0.01864086 0.01632403 0.02167597

```

Figure 12 : valeurs de l'estimateur, de la variance et de l'IC pour MC classique.

1.b Par la méthode de la fonction inverse on a qu'en prenant $U \sim U[0,1]$, on a :

$$F_z^{-1}(U) \sim Z = W(2,1).$$

Or $F_z(x) = 1 - \exp(-(\lambda x)^\alpha)$

On résout cette équation en x et on trouve $x = \frac{1}{\lambda}(-\ln(1-y))^{1/\alpha}$ et en prenant $\lambda=2$ et $\alpha=1$, on obtient :

$$F^{-1}(U) = \frac{1}{2}(-\ln(1-U)) \sim Z = W(2,1)$$

$$\begin{aligned} \text{Ainsi } \mathbb{P}[Z \geq t] &= \mathbb{P}[F^{-1}(U) \geq t] = E_U \left[1_{\{\frac{1}{2}(-\ln(1-U)) \geq t\}} \right] \\ &= E_U[h(U)] \end{aligned}$$

où $h(x) = 1_{\{F^{-1}(x) \geq t\}} = 1_{\{\frac{1}{2}(-\ln(1-x)) \geq t\}}$

D'où $\mathbb{P}[Z \geq t] = E_U[h(U)]$ où $U \sim U[0,1]$.

On reprend les résultats de l'exercice 1 du TD4 :

On pose :

. $\forall k \in \{1, \dots, N\}, D_k = \left[\frac{k-1}{N}, \frac{k}{N} \right)$ une partition de $[0,1]$.

. U la variable de stratification.

. $\frac{k-1+U_i}{N} = U_i^{(k)}$ où $U_i^{(k)} \sim U / U \in D_k$ et $U_i \sim U[0,1]$.

. $n_k = N \cdot \mathbb{P}[U \in D_k] = N \cdot \frac{1}{N} = 1$ (nous choisissons un n_k comme ceci car nous réalisons ici une méthode de stratification avec allocation proportionnelle).

Ainsi l'estimateur obtenu par stratification avec allocation proportionnelle est :

$$\hat{\delta}_n(n_1, \dots, n_N) = \frac{1}{N} \sum_{k=1}^N h\left(\frac{k-1+U_k}{N}\right) \text{ où } U_k \sim U([0,1])$$

	value	var	born_inf_IC	born_sup_IC
10000	0.0183	0.01796691	0.01567285	0.02092715

Figure 13 : valeurs de l'estimateur, de la variance et de l'IC pour MC stratifié.

1.c On observe dans les valeurs obtenues que les estimateurs ont l'air de converg  vers la m me valeur   savoir $\approx 0,018$.

De plus, on observe que la variance (et donc la longueur de l'intervalle de confiance) de l'estimateur par m thode de stratification est l g rement inf rieure   celle de l'estimateur MC classique.

On peut alors supposer que l'estimateur par m thode de stratification avec allocation proportionnelle est l g rement plus performant que l'estimateur de Monte-Carlo classique (comme on l'a vu en cours).

Question 2.

2.a On a $\delta = \mathbb{P}[X_1 + X_2 + X_3 \geq t]$ et $X_1 + X_2$ suit une loi $\Gamma(2,1)$ par somme d'exponentielle ind pendantes de m me param tre.

On pose $X_1 + X_2 = Y$ et donc $\delta = E[1_{\{Y+X_3 \geq t\}}] = E[\varphi(Y, X_3)]$

o  φ d finit par $\varphi(x, y) = 1_{\{x+y \geq t\}}$.

L'estimateur par m thode de Monte-Carlo classique de δ est donc le suivant :

$$\hat{\delta}_n = \frac{1}{n} \sum_{k=1}^n \varphi(Y^{(k)}, X_3^{(k)})$$

o  :

- $(Y^{(k)})_{k \geq 1} \sim \Gamma(2,1)$ est une suite de variables al atoires identiques et ind pendantes.

- $(X_3^{(k)})_{k \geq 1}$ est une suite de variables al atoires identiques et ind pendantes qui ont pour fonction de r partition $F(t) = \frac{t}{4} 1_{\{0 \leq t \leq 1\}} + (\frac{t}{4} + \frac{1}{2}) 1_{\{1 \leq t \leq 2\}} + 1_{\{t \geq 2\}}$.

Ayant accès à la fonction de répartition de X_3 , on peut simuler la suite $(X_3^{(k)})_{k \geq 1}$ (qui sont des réalisations de la loi X_3), par la méthode de la fonction inverse.

De même, la suite $(Y^{(k)})_{k \geq 1}$ est simulée suivant une loi $\Gamma(2,1)$.

$IC(1-\alpha) = [\hat{\delta}_n \pm q_{1-\frac{\alpha}{2}} \sqrt{\frac{V(\varphi(Y, X_3))}{n}}]$ où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ d'une loi $N(0,1)$.

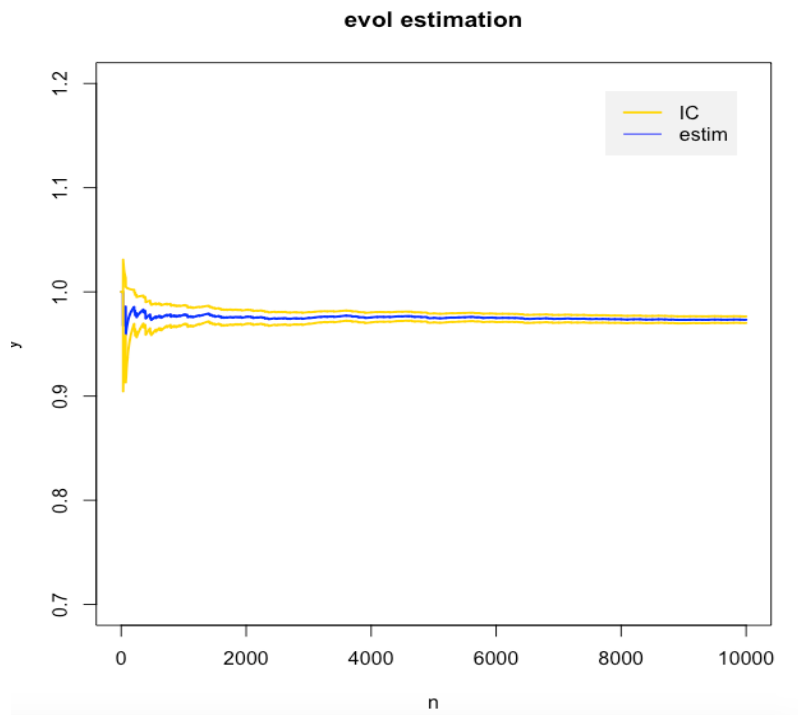


Figure 14 : Evolution de l'estimateur et de l'IC.

```

value      var born_inf_IC born_sup_IC
10000 0.9733 0.02598971  0.9701403  0.9764597

```

Figure 15 : Valeurs obtenues pour MC classique.

2.b On rappelle la propriété de l'espérance conditionnelle suivante :

Si X, Y deux variables aléatoires indépendantes (X à valeur dans E et Y à valeur dans F) alors, $\forall h$ fonctions mesurables (h dans L^1) allant de $E \times F$ dans \mathbb{R}^+ on a :

$$E[h(X, Y)|Y] = \int_E h(x, Y) f_X(x) dx$$

En prenant $h(x, y) = 1_{\{x \leq y\}}$ la propriété précédente nous donne :

$$\mathbb{P}[X \leq Y|Y] = F_X(Y)$$

et puis en passant à l'espérance on obtient :

$$\begin{aligned} E[\mathbb{P}[X \leq Y|Y]] &= E[F_X(Y)] \\ &= E\left[E[1_{\{X \leq Y\}}|Y]\right] \\ &= E[1_{X \leq Y}] \\ &= \mathbb{P}[X \leq Y] \end{aligned}$$

c'est à dire $\mathbb{P}[X \leq Y] = E[F_X(Y)]$

Donc : $\delta = 1 - \mathbb{P}[X_1 + X_2 + X_3 \leq t]$

$$= 1 - \mathbb{P}[X_3 \leq t - (X_1 + X_2)] \text{ avec } X_3 \perp (X_1 + X_2).$$

Or par la propriété précédente appliquée à $X = X_3$ et $Y = X_1 + X_2$, on obtient :

$$\mathbb{P}[X_3 \leq t - (X_1 + X_2)] = E[F(t - (X_1 + X_2))]$$

où F est la fonction de répartition de X_3 .

D'où :

$$\delta = 1 - E[F(t - (X_1 + X_2))].$$

De même, $\delta = 1 - \mathbb{P}[X_1 + X_2 + X_3 \leq t]$

$$= 1 - \mathbb{P}[X_1 + X_2 \leq t - X_3] , X_3 \perp (X_1 + X_2)$$

Or par la propriété précédente à $Y = X_3$ et $X = X_1 + X_2$, on obtient :

$$\mathbb{P}[X_1 + X_2 \leq t - X_3] = E[G(t - (X_3))]$$

où G est la fonction de répartition de la loi $\Gamma(2,1)$.

D'où :

$$\delta = 1 - E[G(t - (X_3))].$$

2.c On peut alors estimer δ par 2 nouvelles méthodes de Monte-Carlo classique :

Méthode 1 : Avec la fonction de répartition G de la loi $\Gamma(2,1)$ et en simulant $(X_3^{(k)})_{k \geq 1}$ i.i.d avec la méthode de la fonction inverse.

Notre estimateur par la méthode de Monte-Carlo classique sera alors :

$$\frac{1}{n} \sum_{k=1}^n g(X_3^{(k)}) \text{ où } g(x) = 1 - G(t - x)$$

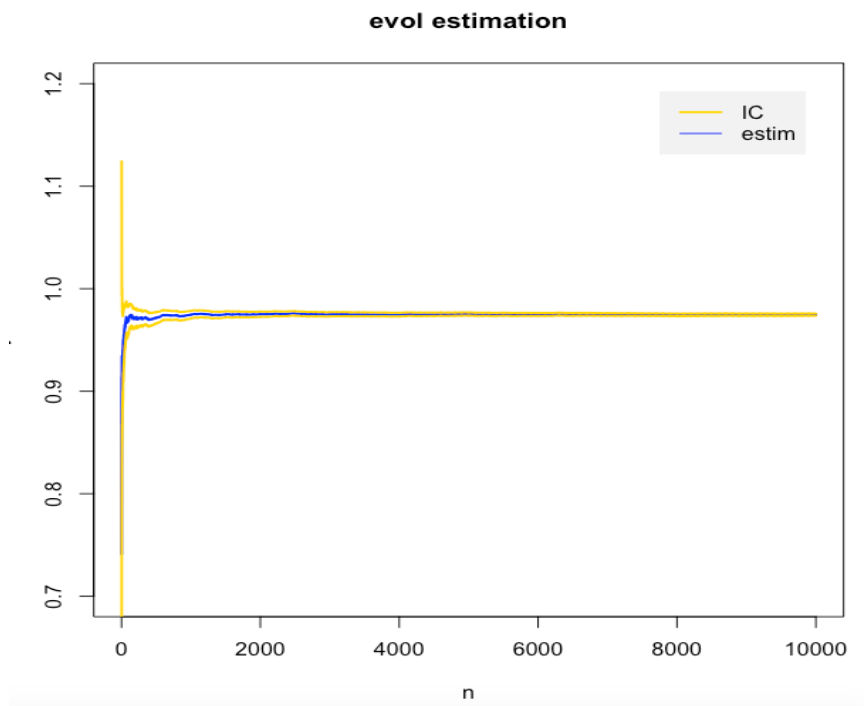


Figure 16 : Evolution de l'estimateurs en utilisant MC avec G.

	value	var	born_inf_IC	born_sup_IC
10000	0.9747559	0.003574545	0.9735841	0.9759277

Figure 17 : Valeurs obtenus pour MC classique avec G.

Méthode 2 : Avec la fonction de répartition F et en simulant $(X_1^{(k)} + X_2)_{k \geq 1}$:
 $= (Y^{(k)})_{k \geq 1}$ i.i.d suivant une loi $\Gamma(2,1)$.

Notre estimateur par méthode de Monte-Carlo classique sera alors :

$$\frac{1}{n} \sum_{k=1}^n \theta(Y^{(k)}) \text{ où } \theta(y) = 1 - F(t - y)$$

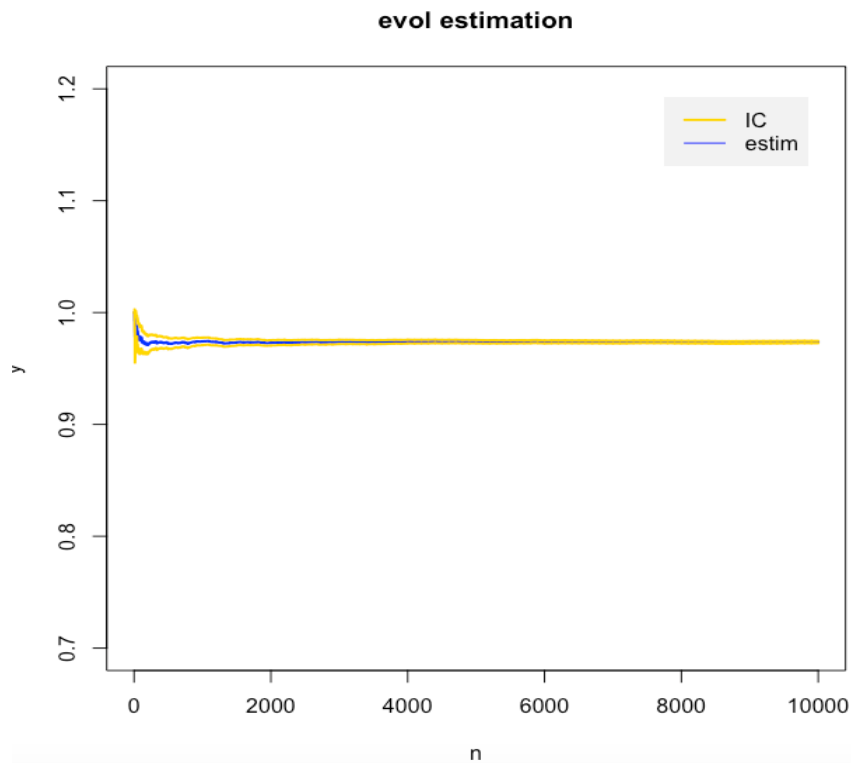


Figure 18 : Evolution de l'estimateurs en utilisant MC avec F.

```
value      var born_inf_IC born_sup_IC
10000 0.9736494 0.002928411 0.9725888 0.9747101
```

Figure 19 : Valeurs obtenues pour MC classique avec F.

COMPARAISON : On trace le plot associé à l'évolution des 3 estimateurs.

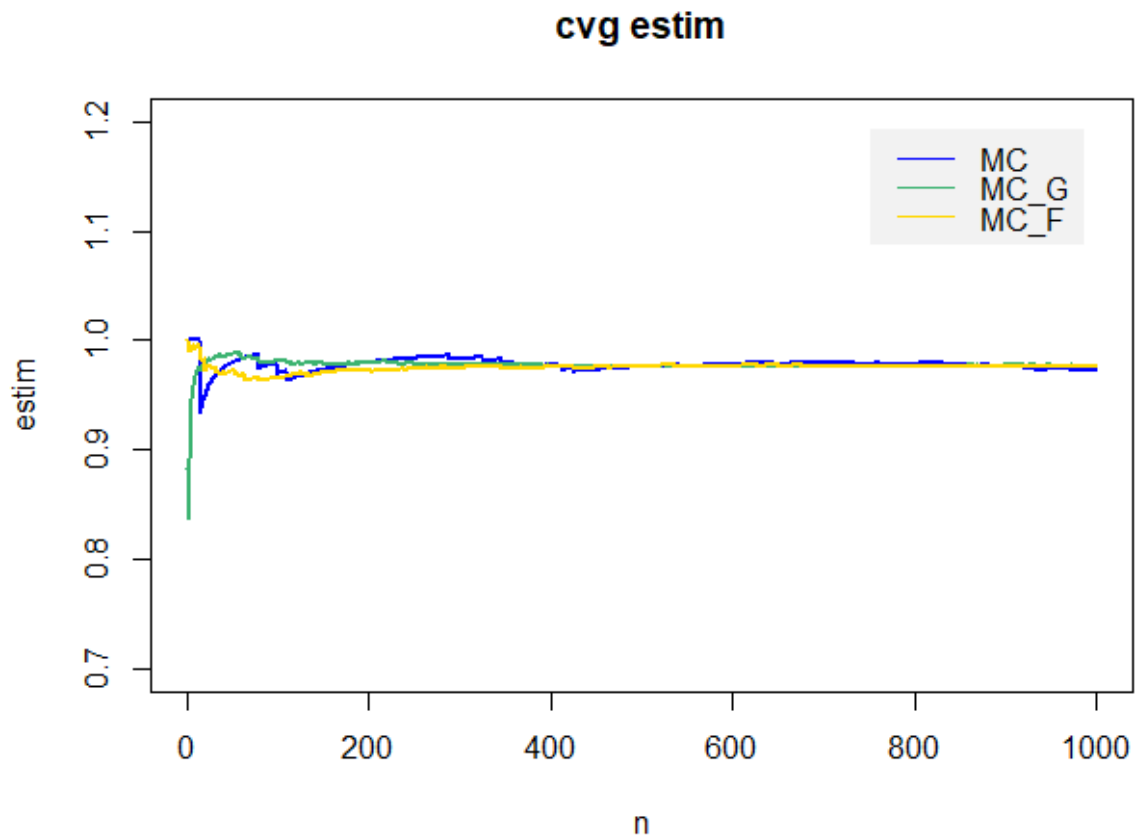


Figure 20 : Comparaison de l'évolution des 3 estimateurs.

On remarque tout d'abord que dans le graphe des convergences, les 3 méthodes convergent vers la même valeur à savoir $\approx 0,97$.

De plus, dans les valeurs obtenues, on observe et ceci est confirmé par la vitesse des convergences dans le graphique, que les estimateurs de Monte-Carlo avec F et G sont plus performants.

En effet, on observe une variance d'environ 0.026 (*ordre de grandeur* 10^{-2}) pour l'estimateur par MC classique, tandis que pour les estimateurs obtenus avec F et G ont des variances d'ordre de grandeur 10^{-3} (0.0035 pour l'estimateur obtenu avec G et 0.0029 pour l'estimateur obtenu avec F).

Ainsi, la variance est plus petite d'un dixième pour ces 2 derniers estimateurs comparativement à celle par MC classique : On peut alors grandement supposer que ces estimateurs sont plus performants.

Concernant la comparaison entre les estimateurs F et G : la variance de l'estimateur obtenu avec F est de 3.5×10^{-3} tandis que celle obtenue avec G est de 2.9×10^{-3} .

Pour résumer, on peut ainsi dire que "l'estimateur F" est légèrement plus performant que "l'estimateur G".

Remarque : comparer les variances pour déterminer l'estimateur le plus performant, revient à comparer la longueur de l'intervalle de confiance. En effet, l'estimateur est d'autant plus performant que la variance et donc que la longueur de l'IC sont petites.

Partie II : Estimation de p

Question 1.

On simule un vecteur $(X_k)_{k \geq 1} = (X_1^{(k)}, X_2^{(k)}, X_3^{(k)}, X_4^{(k)}, X_5^{(k)}, k \geq 1)$ i.i.d avec les $(X_k)_{k \geq 1}$ de densité $f(\mathbf{x}, \lambda_1, \dots, \lambda_5)$ qui est la densité de produit d'exponentielle indépendants $\varepsilon(\lambda_i), 1 \leq i \leq 5$.

Il existe, d'après le graphe fourni dans l'énoncé, 4 différents chemins pour aller de A à B.

Ainsi $d(X) = \min(X_1 + X_4, X_2 + X_5, X_1 + X_3 + X_5, X_2 + X_3 + X_4)$

On veut alors estimer $p = \mathbb{P}[d(X) \geq t] = E_f[1_{\{d(X) \geq t\}}] = E_f[g(X)]$

où $g(x) = 1_{[d(x) \geq t]}$, avec x à valeur dans \mathbb{R}^5 .

On a alors, l'estimateur par méthode de Monte-Carlo classique :

$$\hat{p}_n = \frac{1}{n} \sum_{k=1}^n g(X^{(k)})$$

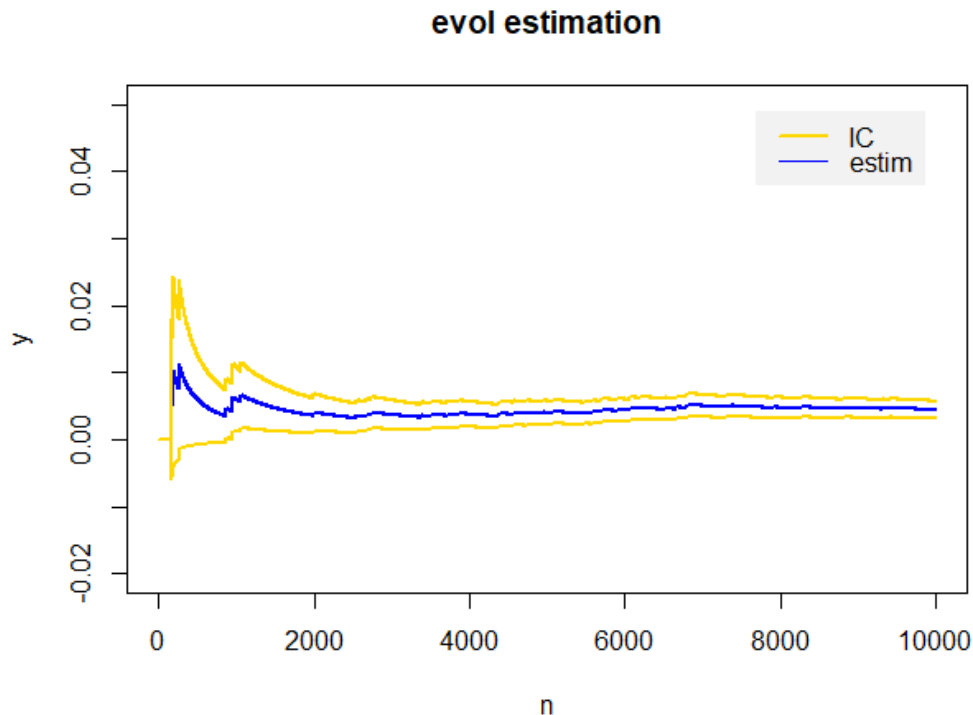


Figure 21 : Evolution de l'estimateur de MC classique.

value	var	born_inf_IC	born_sup_IC
10000	0.004	0.003984398	0.00276283 0.00523717

Figure 22 : Valeurs de l'estimateurs de la variance et de l'IC.

Question 2.

$$\begin{aligned}
 \text{2.a } E_f[1_{\{d(x) \geq t\}}] &= \int_{\mathbb{R}^5} 1_{\{d(x) \geq t\}} f(x, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) dx_1 dx_2 dx_3 dx_4 dx_5 \\
 &= \int_{\mathbb{R}^5} \frac{1_{\{d(x) \geq t\}} f(x, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)}{g(x)} g(x) dx_1 dx_2 dx_3 dx_4 dx_5 \\
 \text{Or } \frac{1_{\{d(x) \geq t\}} f(x, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)}{g(x)} &= p \\
 &= \int_{\mathbb{R}^5} p g(x) dx_1 dx_2 dx_3 dx_4 dx_5 \\
 &= E_g[p] \\
 &= p
 \end{aligned}$$

On a alors l'estimateur d'échantillonnage préférentiel suivant :

$$\hat{p}_n = \frac{1}{n} \sum_{k=1}^n p = p = \text{constante}$$

donc $\text{var}(\hat{p}_n) = \text{var}(p) = 0$.

2.b On va utiliser une méthode de rejet pour simuler selon \tilde{g} tel que $g = c\tilde{g}$

où : $c = \frac{1}{p}$ et $\tilde{g} = 1_{\{d(x) \geq t\}} f(x, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$.

Ici $f(x, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$ est la densité instrumentale.

On a ainsi l'équivalence entre appliquer la méthode de rejet à g avec $\tilde{M} = cM$ et f et l'appliquer à \tilde{g} avec M et f . (cf exercice 1, question 1)

On a $1_{\{d(x) \geq t\}} \leq 1$ donc $\tilde{g}(x) \leq f(x, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$

$$= M f(x, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$$

avec $M=1$ (remarque : on a bien la condition requise $M \geq 1$)

On applique ainsi la méthode de rejet avec pour densité instrumentale f (qui est le produit d'exponentielle indépendantes) et pour ce choix M : on obtient ainsi des réalisations suivant la loi dont la densité est g .

2.c On constate qu'en choisissant cette loi d'importance g (quand bien même si cette loi est optimale comme vu en cours), l'estimateur par échantillonnage préférentiel $\hat{p}_n = p$. Or p est justement la quantité que l'on souhaite estimée.

On ne peut donc pas appliquer cette méthode d'échantillonnage préférentiel avec comme loi d'importance g : g en tant que loi d'importance est inexploitable.

Question 3.

3.a Soit $(y_k)_k$ une suite de variables aléatoires i.i.d de densité f .

On a :

$$\frac{1}{n} \sum_{k=1}^n 1_{\{d(Y_k) \geq t\}} w(Y_k) \xrightarrow{ps} E_g[1_{\{d(Y) \geq t\}} w(Y)]$$

car $(Z_k)_{k \geq 0} = 1_{\{d(Y_k) \geq t\}} w(Y_k)$ est une suite de variables aléatoires i.i.d et dans L^1 .

$$\text{Or } E_g[1_{\{d(Y) \geq t\}} w(Y)] = \int_{\text{support } g} 1_{\{d(y) \geq t\}} w(y) g(y) dy$$

$$= \int_{\text{support } g} 1_{\{d(y) \geq t\}} \frac{f(y, \lambda_1, \dots, \lambda_5)}{g(y)} g(y) dy$$

$$= \int_{\text{support } f} (1_{\{d(y) \geq t\}}) f(y, \lambda) dy$$

$$= E_f[1_{\{d(Y) \geq t\}}]$$

$$= \mathbb{P}_f[d(X) \geq t] = p$$

De même, $\frac{1}{n} \sum_{k=1}^n w(Y_k) \xrightarrow{ps} E_g[w(Y)]$ car les $(w(Y_k))_{k \geq 0}$ sont i.i.d et dans L^1 .

$$\begin{aligned} \text{Or : } E_g[w(Y)] &= \int_{\text{support } g} g(y) \frac{f(y)}{g(y)} dy \text{ avec } \text{support } g = \{x \in (\mathbb{R}^+)^5; d(x) \geq t\} \\ &= \int_{(\mathbb{R}^+)^5} f(y) (1_{d(y) \geq t}) dy \\ &= E_f[1_{\{d(Y) \geq t\}}] \\ &= \mathbb{P}[d(Y) \geq t] \text{ où } Y \text{ suit la loi dont la densité est } f. \\ &= p \text{ (par définition).} \end{aligned}$$

Ainsi la quantité converge presque sûrement vers $\frac{p}{p} = 1$ qui est différent de p .

Remarque : $w(Y_k)$ est bien défini $\rightarrow g(y_k) \neq 0$ ps quand $y_k \sim g$.

En effet $y_k \sim g$ implique que ps $y_k \in \text{support } g = \{x \in (\mathbb{R}^+)^5; d(x) \geq t\}$ donc

$$g(Y_k) = \frac{f(y_k, \lambda_1, \dots, \lambda_5)}{p} \text{ ps car } 1_{\{d(Y_k) \geq t\}} = 1 \text{ ps.}$$

Ainsi on a un résultat encore plus fort, la quantité vaut presque sûrement 1 et il n'est donc pas étonnant que la quantité converge presque sûrement vers 1.

3.b On ne peut donc pas en déduire une méthode d'estimation de p à l'aide de la quantité précédente : celle-ci ne converge pas vers p . Ainsi, g est inexploitable tout court (et pas seulement en tant que loi d'importance).

Question 4.

$p = \mathbb{P}[d(X) \geq t] = E_f[1_{\{d(X) \geq t\}}] = E_f[U(X)]$ où $U(x) = 1_{\{d(x) \geq t\}}$ et X suit une loi dont la densité est f .

D'où en prenant h comme loi d'importance (on peut le faire car $\text{support}(f) \subset \text{support}(h)$), ils sont même tous 2 égaux à $(\mathbb{R}^+)^5$:

$$p = E_h \left[U(X) \frac{f(X)}{h(X)} \right]$$

On en déduit alors l'estimateur d'échantillonnage préférentiel suivant :

$$\hat{p}_n(\alpha_1, \dots, \alpha_5) = \frac{1}{n} \sum_{k=1}^n U(X_k) \frac{f(X_k)}{h(X_k)}$$

où :

. $(X_k)_{k \geq 0} \sim$ loi dont la densité est f .

. h est de densité $\otimes \varepsilon(\alpha_i)_{1 \leq i \leq 5}$ et f de densité $\otimes \varepsilon(\lambda_i)_{1 \leq i \leq 5}$

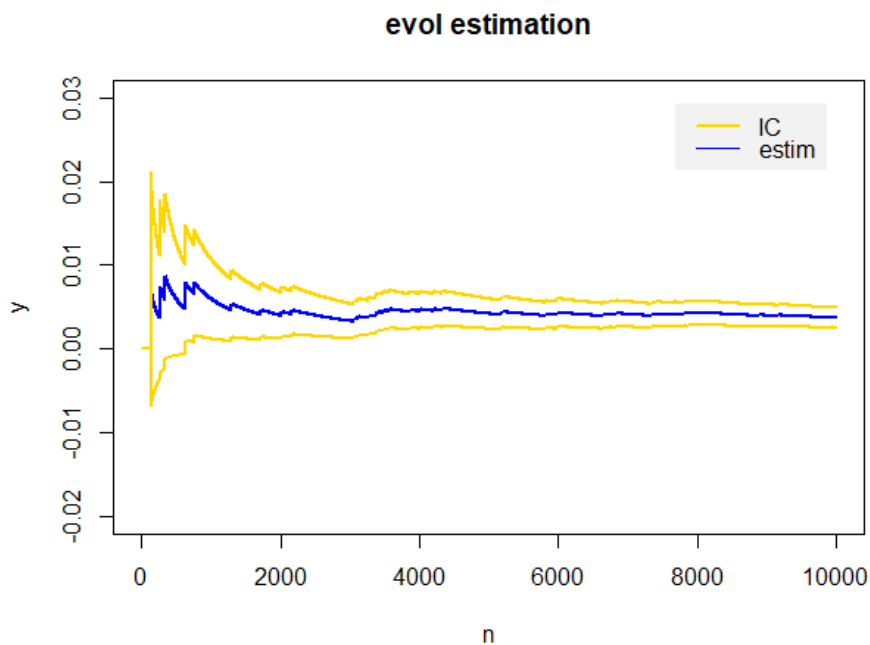


Figure 23 : Evolution de l'estimateur de MC avec échantillonnage préférentiel et des α rentrées par l'utilisateur (ici $\alpha=(6,7,3,2,1)$ par exemple).

```
$deltachap
[1] 0.0037

$variance
[1] 0.003686679

$IC
[1] 0.002509949 0.004890051
```

Figure 24 : Valeurs de l'estimateurs de la variance et de l'IC.

Question 5.

On pose $\alpha = (\alpha_1, \dots, \alpha_5)$ et $\lambda = (\lambda_1, \dots, \lambda_5)$ pour alléger les notations dans les calculs qui vont suivre.

On a :

$$\begin{aligned} & \arg \min \left(\alpha, E_g \left[\ln \left(\frac{g(z)}{h(z, \alpha)} \right) \right] \right) \\ &= \arg \min \left(\alpha, \int_{\text{sup } g} \ln \left(\frac{g(z)}{h(z, \alpha)} \right) g(z) dz \right) \\ &= \arg \min \left(\alpha, \int_{\text{sup } g} [\ln(g(z)) - \ln(h(z, \alpha))] (1_{\{d(z) > t\}}) f(z, \lambda) dz \right) \end{aligned}$$

Par définition de $g(z)$ et du fait que p est constante (donc n'influe en rien sur l'argmin).

De plus α n'apparaît pas dans l'expression de g (donc g n'influe en rien sur l'argmin) donc on a :

$$= \arg \min \left(\alpha, \int_{\text{sup } g} -\ln(h(z, \alpha)) (1_{\{d(z) > t\}}) f(z, \lambda) \frac{h(z, \alpha^0)}{h(z, \alpha^0)} dz \right)$$

Or $\arg \min(f) = \arg \max(-f)$

$$= \arg \max \left(\alpha, \int_{\text{sup } g} \ln(h(z, \alpha)) (1_{\{d(z) > t\}}) f(z, \lambda) \frac{h(z, \alpha^0)}{h(z, \alpha^0)} dz \right)$$

On change de mesure : on peut le faire car $\{x \in (\mathbb{R}^+)^5; d(x) \geq t\} = \text{supp}(g) \subset \text{supp}(h^0) = \{x \in \mathbb{R}^5\}$

Ceci implique donc :

$$= \arg \max \left(\alpha, E_{h(\cdot, \alpha^0)} \left[\frac{\ln(h(Z, \alpha)) (1_{\{d(Z) > t\}}) f(Z, \lambda)}{h(Z, \alpha^0)} \right] \right)$$

D'où le résultat voulu.

Question 6.

Par la question 5,

$$\text{On a } (\alpha_1^*, \dots, \alpha_5^*) = \arg \max \left(\alpha, E_{h(\cdot, \alpha^0)} \left[\frac{\ln(h(Y, \alpha)) (1_{\{d(Y) > t\}}) f(Y, \lambda)}{h(Y, \alpha^0)} \right] \right)$$

on note $I = E_{h(\cdot, \alpha^0)} \left[\frac{\ln(h(Y, \alpha)) (1_{\{d(Y) \geq t\}}) f(Y, \lambda)}{h(Y, \alpha^0)} \right]$

On pose $S = \frac{1}{n} \sum_{k=1}^n 1_{\{d(Y_k) \geq t\}} W(Y_k, \lambda_i, \alpha_i^0) \ln(h(Y_k, \alpha))$

où $W(Y, \lambda_i, \alpha_i^0) = \frac{f(Y, \lambda)}{h(Y, \alpha^0)} = \frac{\prod_{i=1}^5 \lambda_i \exp(-\lambda_i Y_i)}{\prod_{i=1}^5 \alpha_i^0 \exp(-\alpha_i^0 Y_i)} = \prod_{i=1}^5 \frac{\lambda_i}{\alpha_i^0} \exp(-Y_i(\lambda_i - \alpha_i^0))$

(Remarque : On a $S \xrightarrow{ps} I$ par la LGFN)

On a alors : $\frac{\partial S}{\partial \alpha_j} = \frac{1}{n} \sum_{k=1}^n 1_{\{d(Y_k) \geq t\}} W(Y_k, \lambda_i, \alpha_i^0) \frac{\partial}{\partial \alpha_j} \ln(h(Y_k, \alpha))$

Or $h(Y_k, \alpha) = \prod_{k=1}^5 \alpha_k \exp(-\alpha_k Y_k) \Rightarrow \ln h(Y_k, \alpha) = \sum_{k=1}^5 [\ln(\alpha_k) - \alpha_k Y_k]$

D'où $\frac{\partial}{\partial \alpha_j} \ln(h(Y_k, \alpha)) = \frac{1}{\alpha_j} - Y_j$

CPO :

On veut $\frac{\partial S}{\partial \alpha_j} = 0$

$\Leftrightarrow \sum_{k=1}^n 1_{\{d(Y_k) \geq t\}} W(Y_k, \lambda_i, \alpha_i^0) \left(\frac{1}{\alpha_j} - Y_j \right) = 0$

Et en posant $\bar{W} := 1_{\{d(Y_k) \geq t\}} W(Y_k, \lambda_i, \alpha_i^0)$ il vient :

$\Leftrightarrow \sum_{k=1}^n \bar{W} \frac{1}{\alpha_j} = \sum_{k=1}^n \bar{W} Y_j$

$\Leftrightarrow \hat{\alpha}_j^* = \frac{\sum_{k=1}^n \bar{W}}{\sum_{k=1}^n \bar{W} Y_j}$

$\Leftrightarrow \hat{\alpha}_j^* = \frac{\sum_{k=1}^n 1_{\{d(Y_k) \geq t\}} W(Y_k, \lambda_i, \alpha_i^0)}{\sum_{k=1}^n 1_{\{d(Y_k) \geq t\}} W(Y_k, \lambda_i, \alpha_i^0) Y_{k,j}}$

avec $W(Y_k, \lambda_i, \alpha_i^0) = \prod_{i=1}^5 \frac{\lambda_i}{\alpha_i^0} \exp(-Y_{k,i}(\lambda_i - \alpha_i^0))$

D'où finalement :

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

Question 7.

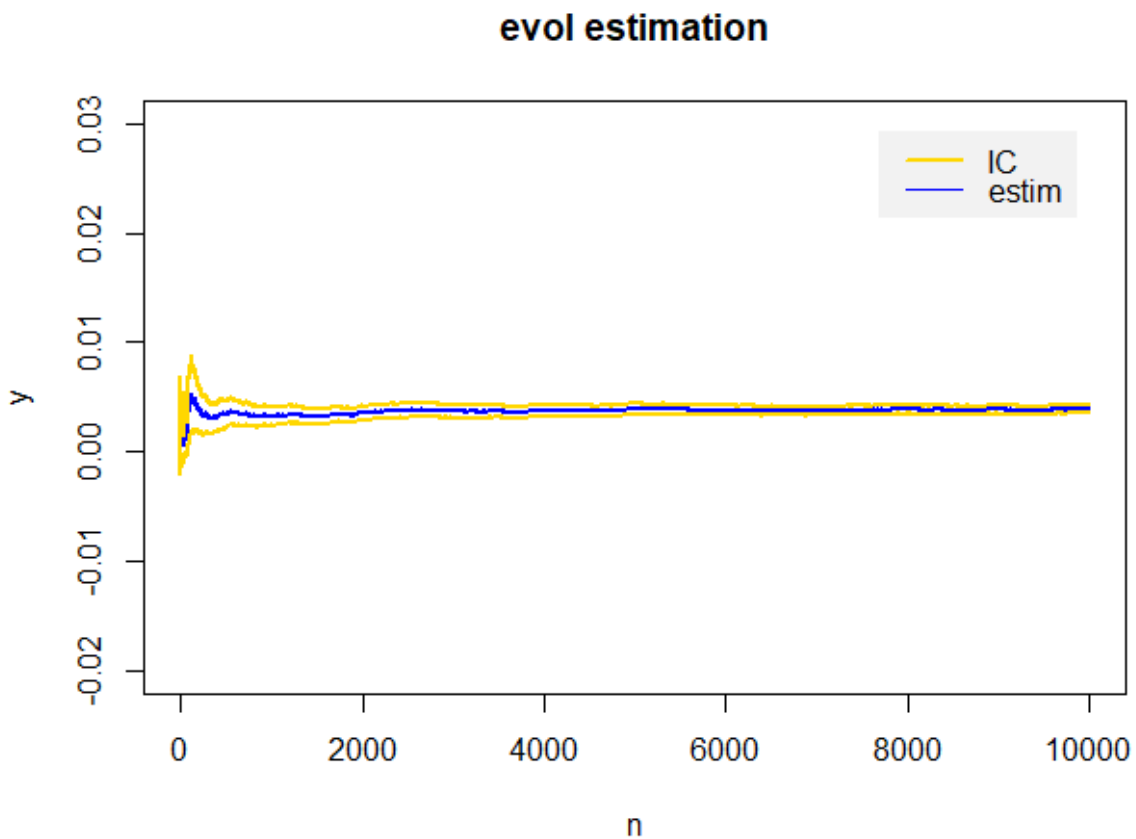


Figure 26 : Evolution de l'estimateur de MC avec échantillonnage préférentielle et α^* optimal.

```

$deltachap
[1] 0.003879888

$variance
[1] 0.0003010444

$IC
$IC$IC
[1] 0.003539822 0.004219954

$IC$LongueurIC
[1] 0.0006801322

```

Figure 25 : Valeurs de l'estimateurs de la variance et de l'IC avec α^* .

COMPARAISON AVEC LA QUESTION 4 :

On observe, qu'en utilisant α^* , on obtient une variance d'environ 0.0003 (ordre de grandeur de 10^{-4}) tandis que la variance dans la question 4 est d'environ 0.0036 (ordre de grandeur de 10^{-3}).

L'estimateur converge donc bien plus rapidement avec ce α^* optimal et l'intervalle de confiance, comme on l'observe, est bien plus resserré. Ainsi, la méthode d'échantillonnage préférentiel sans choix du α^* optimal, n'est pas meilleur que la méthode de Monte-Carlo classique en termes de variance (tous les deux d'ordre 10^{-3}).

On comprend donc mieux l'intérêt de trouver ce α^* optimal. Ceci a l'air d'être fortement confirmé avec le graphique suivant montrant l'évolution des estimateurs des Q1, Q4 et Q7.

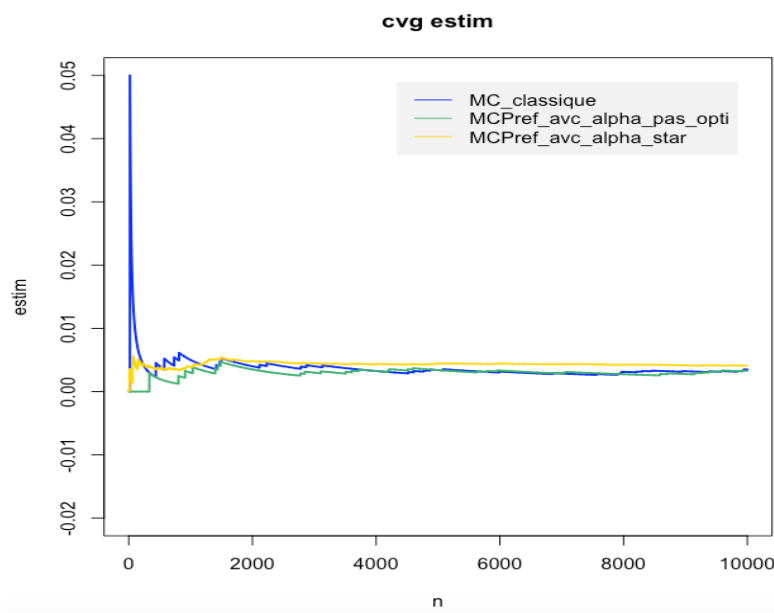


Figure 27 : Comparaison des estimateurs de MC classique, et les estimateurs d'échantillonnages préférentielles avec α quelconque, et α^* optimal.

Question 8.

8.a Par la loi forte des grands nombres (on peut utiliser la LFGN car l'intérieur de la somme sont des variables identiques, \perp par transformation mesurables de v.a \perp et dans L^1) :

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \prod_{k=1}^5 \frac{\lambda_k}{\alpha_k^*} U_{i,k}^{\lambda_k/\alpha_k^*-1} \left(1_{\{\rho(U_{i,1}, \dots, U_{i,5}) \geq t\}} \right) \\ & \xrightarrow{ps} E_U \left[\prod_{k=1}^5 \frac{\lambda_k}{\alpha_k^*} U_{i,k}^{\lambda_k/\alpha_k^*-1} \left(1_{\{\rho(U_{i,1}, \dots, U_{i,5}) \geq t\}} \right) \right] \\ & = \int_{[0,1]^5} \prod_{k=1}^5 \frac{\lambda_k}{\alpha_k^*} U_{i,k}^{\lambda_k/\alpha_k^*-1} \left(1_{\{\rho(U_{i,1}, \dots, U_{i,5}) \geq t\}} \right) du_1, \dots, du_5 \end{aligned}$$

On fait le changement de variable suivant :

$$x_i = -\frac{1}{\alpha_i^*} \ln u_i \rightarrow u_i = \exp(-\alpha_i^* x_i) \text{ et } dx_i = -\frac{1}{\alpha_i^* u_i} du_i \text{ et on obtient}$$

avec ce changement de variable :

$$\int_{(\mathbb{R}^+)^5} \left(\prod_{k=1}^5 \lambda_k \exp(-\lambda_k x_k) \right) 1_{\{\rho(\exp(-\alpha_1^* x_1), \dots, \exp(-\alpha_5^* x_5)) \geq t\}} dx_1 \cdots dx_5 \quad (1)$$

Puis, en remarquant que tous les termes se simplifient avec la forme de la fonction ρ (en composant les exponentielles par les logarithmes il ne nous reste que les X_i), on obtient :

$$\begin{aligned} & \rho(\exp(-\alpha_1^* x_1), \dots, \exp(-\alpha_5^* x_5)) \\ & = \min(X_1 + X_4, X_2 + X_5, X_1 + X_3 + X_5, X_2 + X_3 + X_4) \\ & = d(X) \end{aligned}$$

$$\begin{aligned} (1) \text{ devient alors : } & \int_{(\mathbb{R}^+)^5} \left(\prod_{k=1}^5 \lambda_k \exp(-\lambda_k x_k) \right) 1_{d(X) \geq t} dx_1 \cdots dx_5 \\ & = E_f [1_{\{d(X) \geq t\}}] \text{ où } f \text{ est la densité de } \otimes (\varepsilon(\lambda_i), 1 \leq i \leq 5) \\ & = \mathbb{P}[d(X) \geq t] \\ & = p \end{aligned}$$

Ainsi, on a bien la convergence souhaitée, à savoir :

$$\frac{1}{n} \sum_{i=1}^n \prod_{k=1}^5 \frac{\lambda_k}{\alpha_k^*} U_{i,k}^{\lambda_k/\alpha_k^*-1} \left(1_{\{\rho(U_{i,1}, \dots, U_{i,5}) \geq t\}} \right) \xrightarrow{ps} p$$

8.b D'après la question précédente :

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

$$= \frac{1}{n} \sum_{i=1}^n h(U_i) \stackrel{ps}{>} p.$$

où $(U_{i,k})_{1 \leq i \leq n} = (U_{i,1}, \dots, U_{i,5})_{1 \leq i \leq n}$ est une suite de V.A i.i.d de loi $U \sim U(0, 1)^5$.

Ici, on applique la méthode des variables antithétiques sur la loi uniforme $U(0, 1)^5$ qui présente des propriétés d'invariance. En effet, il existe bien une transformation mesurable A définie comme suit :

$$A : \mathbb{R}^5 \rightarrow \mathbb{R}^5 \quad \text{telle que } A(U) = 1 - U \sim U$$

$$X \rightarrow 1-X$$

La variable aléatoire $A(U)$ est appelée variable antithétique de U et l'estimateur de la variable antithétique est défini par :

$$\hat{p}_n^{(1)} = \frac{1}{n} \sum_{i=1}^n \frac{h(U_{i,k}) + h(A(U_{i,k}))}{2}$$

$$= \frac{1}{2n} \sum_{i=1}^n \prod_{k=1}^5 \frac{\lambda_k}{\alpha_k^*} U_{i,k}^{\lambda_k/\alpha_k^*-1} \left(1_{\{\rho(U_{i,1}, \dots, U_{i,5}) \geq t\}} \right) + \prod_{k=1}^5 \frac{\lambda_k}{\alpha_k^*} (1 - U_{i,k})^{\lambda_k/\alpha_k^*-1} \left(1_{\{\rho(1-U_{i,1}, \dots, 1-U_{i,5}) \geq t\}} \right)$$

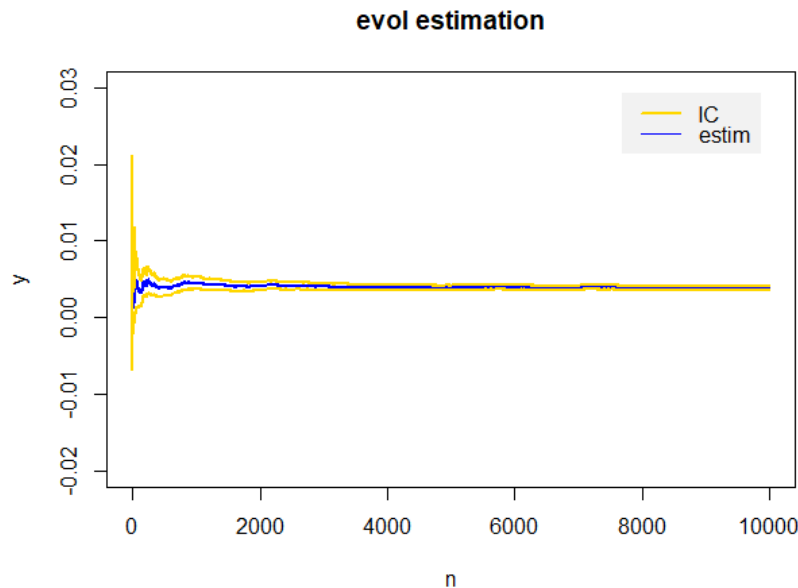


Figure 28 : Evolution de l'estimateur de MC par méthode de la variable antithétiques.

```

$spchap
[1] 0.003894425

$var
[1] 0.0001495996

$IC
$IC$IC
[1] 0.00365470 0.00413415

$IC$LongueurIC
[1] 0.00047945

```

Figure 29 : Valeurs de l'estimateurs de la variance et de l'IC avec méthode de la variable antithétique.

8.c

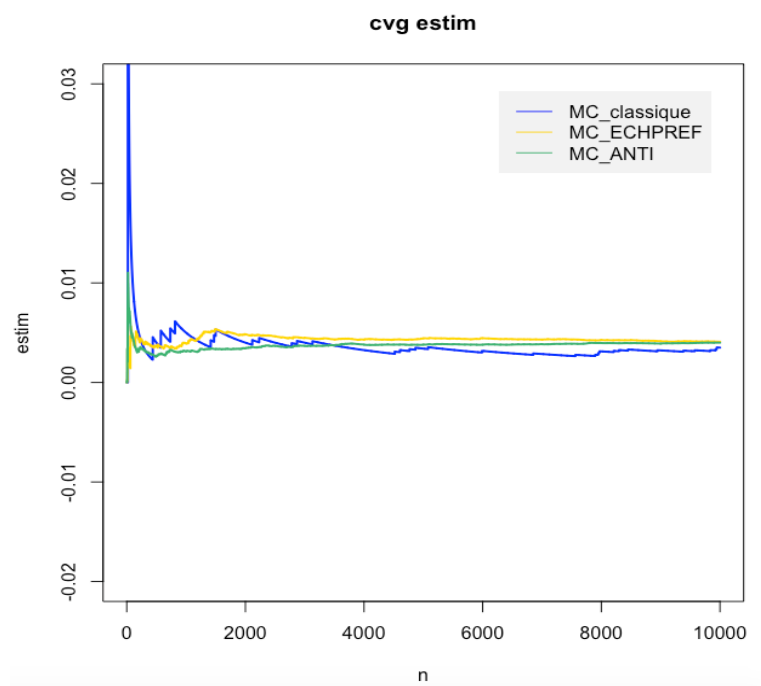


Figure 30 : Comparaison des estimateurs MC classique, MC avec variable antithétique et échantillonnage préférentiel avec α^* optimal.

Nous allons à présent comparer les méthodes des questions 1, 7 et 8.

Notons en premier lieu que les 3 estimateurs convergent bien vers la même valeur (≈ 0.0038).

Il n'est pas difficile de voir que la méthode de Monte-Carlo classique est la moins performante des 3 méthodes.

En effet sa variance n'est de l'ordre que de 10^{-3} (≈ 0.00398) tandis que les 2 autres méthodes présentes une variance de l'ordre de 10^{-4} .

Avec l'échantillonnage préférentiel on obtient une variance de 3×10^{-4} tandis qu'avec la méthode de variables antithétiques, on obtient une variance de $1,4 \times 10^{-4}$.

On peut ainsi être tenter de dire que la méthode avec variable antithétiques est un légèrement plus performantes à cause de ce léger écart.

Pour résumer, on peut ordonner par ordre de performance les 3 estimateurs suivant :

EstimMCclassique < EstimMCpréférentiel < ou = EstimMCva.anti