Corrigé du contrôle no 2, sujet B (durée : 3h)

*Documents, téléphones et calculatrices interdits. La plus grande importance a été accordée lors de la correction à la justification des réponses. Les exercices sont indépendants. On n'a pas tenu compte dans la correction des erreurs de syntaxe en **R**.*

Exercice 1.

(1) Nous remarquons que (changement de variable $t = \log(x)$)

$$\begin{aligned} \int_2^{+\infty} \frac{1}{x(\log(x))^2} dx &= \int_{\log(2)}^{+\infty} \frac{1}{t^2} dt \\ &= \left[-\frac{1}{t} \right]_{\log(2)}^{+\infty} \\ &= \frac{1}{\log(2)} . \end{aligned}$$

Nous pouvons donc écrire $I = \int_2^{+\infty} \frac{\sin(x)}{\log(2)} \times \frac{\log(2)}{x(\log(x))^2} dx$. Si $X_1, X_2, \dots$ sont i.i.d. de loi de densité $x \mapsto \mathbb{1}_{[0;+\infty[}(x) \frac{\log(2)}{x(\log(x))^2}$ alors, par la loi des grands nombres,

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\sin(X_i)}{\log(2)} \xrightarrow[n\to+\infty]{\text{p.s.}} I$$

(nous avons $\mathbb{E}(|\sin(X_1)/\log(2)|) < +\infty$ puisque la fonction sin est bornée). Nous pouvons aussi écrire

$$I = \int_2^{+\infty} \frac{e^{x-2}\sin(x)}{x(\log(x))^2} \times e^{2-x} dx .$$

Nous remarquons que $x \mapsto \mathbb{1}_{[2;+\infty[}(x) e^{2-x}$ est une densité de probabilité. Si $Y_1, Y_2, \dots$ sont i.i.d. de loi de densité $x \mapsto \mathbb{1}_{[2;+\infty[}(x) e^{2-x}$, nous avons

$$\begin{aligned} \mathbb{E}\left(\left| \frac{e^{Y_i-2}\sin(Y_i)}{Y_i(\log(Y_i))^2} \right| \right) &\leq \mathbb{E}\left(\frac{e^{Y_i-2}\sin(Y_i)}{Y_i(\log(Y_i))^2} \right) \\ &= \int_2^{+\infty} \frac{1}{x(\log x)^2} dx \\ &< \infty . \end{aligned}$$

Donc, par la loi des grands nombres,

$$\frac{1}{n} \sum_{i=1}^{n} \frac{e^{Y_i-2}\sin(Y_i)}{Y_i(\log(Y_i))^2} \xrightarrow[n\to+\infty]{\text{p.s.}} I .$$

(2) Nous voulons utiliser la première méthode. Calculons la fonction de répartition. Pour $y \geq 2$, nous utilisons le même changement de variable que ci-dessus pour calculer :

$$\begin{aligned} \int_2^{y} \frac{\log(2)}{x(\log x)^2} dx &= \int_{\log(2)}^{\log(y)} \frac{\log(2)}{t^2} dt \\ &= \left[-\frac{\log(2)}{t} \right]_{\log(2)}^{\log(y)} \\ &= 1 - \frac{\log(2)}{\log(y)} . \end{aligned}$$

La fonction de répartition est donc

$$F\,:\,y\mapsto\begin{cases}0 & \text{si } y<2\,,\\ 1-\frac{\log 2}{\log y} & \text{si } y\geq 2\,.\end{cases}$$

On inverse facilement cette fonction. Si $u\in[0;1]$ et $u=1-\log(2)/\log(y)$ $(y\geq 2)$, alors $y=\exp(\log 2/(1-u))$ (et inversement). Voir l'algorithme 1 pour le calcul de Monte-Carlo.

Algorithme 1 Boucle de Monte-Carlo

```
s=0
n=1000
for (i in 1:n)
{
  u=runif(1,0,1)
  y=exp(log(2)/(1-u))
  s=s+sin(y)/log(2)
}
print(s/n)
```

(3) On remarque que $\mathbb{E}((\sin(X_1)/\log(2))^2)<\infty$ (parce que la fonction sin est bornée). Le théorème central-limite nous dit que l'erreur

$$I-\frac{1}{n}\sum_{i=1}^{n}\frac{\sin(X_i)}{\log(2)}$$

est (approximativement) une gaussienne Z de variance σ^2/n. Nous voulons

$$\mathbb{P}(|Z|>0,02)\leq 0,01\,,$$

c'est à dire

$$\mathbb{P}\left(\left|\frac{\sqrt{n}Z}{\sigma}\right|>\frac{\sqrt{n}}{\sigma}0,02\right)\leq 0,01\,.$$

En utilisant les symétries de la gaussienne, nous voyons que

$$\mathbb{P}\left(\left|\frac{\sqrt{n}Z}{\sigma}\right|>\frac{\sqrt{n}}{\sigma}0,02\right)=2\mathbb{P}\left(\frac{\sqrt{n}Z}{\sigma}>\frac{\sqrt{n}}{\sigma}0,02\right)\,.$$

Comme $\sqrt{n}Z/\sigma$ est de loi $\mathcal{N}(0;1)$, nous voyons dans la table qu'il suffit de prendre n tel que

$$\frac{\sqrt{n}}{\sigma}0,02\geq 2,58\,,$$

c'est à dire

$$n\geq\left(\frac{2,58\sigma}{0,02}\right)^2\,.$$

Exercice 2.

(1) On calcule la fonction de répartition G associée à g. Si $x\leq 1$, $G(x)=0$. Si $x>1$ (changement de variable $u=\log(t)$)

$$\begin{aligned}\int_1^x\frac{1}{t(1+\log(t))^2}dt &= \int_0^{\log(x)}\frac{1}{(1+u)^2}du\\ &= \left[-\frac{1}{1+u}\right]_0^{\log(x)}\\ &= 1-\frac{1}{1+\log(x)}\,.\end{aligned}$$

On inverse cette fonction de répartition :

$$1-\frac{1}{1+\log x}=u \text{ avec } u\in[0;1], x\geq 1 \Leftrightarrow x=\exp\left(\frac{1}{1-u}-1\right).$$

Nous pouvons donc simuler une variable de densité g en utilisant l'inverse de la fonction de répartition (voir programme 2).

Algorithme 2 Simulation de variable aléatoire

```
u=runif(1,0,1)
x=exp(1/(1-u)-1)
print(x)
```

(2) Pour tout $x>1$, comme $\inf(\log(x),1)\leq\log(x)$, nous avons

$$f(x)\leq\frac{\log(x)}{Zx(1+(\log x)^2)}=\frac{1}{2Z}g(x).$$

La formule est encore vraie pour $x\leq 1$.

(3) Nous utilisons la technique d'échantillonage d'importance pour simuler suivant la loi de densité f (voir programme 3). Remarquons que pour $x>0$,

$$\frac{f(x)}{(1/(2Z))g(x)}=\frac{\inf(1,\log(x))}{\log(x)}.$$

Algorithme 3 Simulation de variable aléatoire

```
b=0
while (b==0)
{
  u=runif(1,0,1)
  x=sqrt(1/(1-u)-1)
  v=runif(1,0,1)
  if (v<inf(1,log(x))/log(x))
  {
    b=1
  }
print(x)
```

Exercice 3. Voir le cours.

Exercice 4.

(1) La constante Z est bien finie parce que $\mathcal{S}_{10}$ est fini (de cardinal 10!).

(2) Soient $(r_1,\dots,r_{10})$ et $(s_1,\dots,s_{10})$ tels que $Q((r_1,\dots,r_{10}),(s_1,\dots,s_{10}))=1/10$. Il existe $i\in\{1,2,\dots,10\}$ tel que $r_i=s_{i+1[10]}$, $r_{i+1[10]}=s_i$. Donc $Q((s_1,\dots,s_{10}),(r_1,\dots,r_{10}))=1/10$. Si $Q((r_1,\dots,r_{10}),(s_1,\dots,s_{10}))=0$, on a aussi $Q((s_1,\dots,s_{10}),(r_1,\dots,r_{10}))=0$. Donc Q est symétrique. Puisque Q est symétrique, pour montrer qu'il est irréductible, il suffit de montrer que pour tout $(r_1,\dots r_{10})\in\mathcal{S}_{10}$, il existe n tel que $Q^n((r_1,\dots,r_{10}),(1,2,\dots,10))>0$. Nous remarquons qu'il exsite k_1 tel que $Q^{k_1}((r_1,\dots,r_{10}),x_1)>0$ avec x_1 de la forme $(1,\dots)$ (il suffit de déplacer 1 de plus proche voisin en plus proche voisin jusqu'à la bonne place). Puis il existe k_2 tel que $Q^{k_2}(x_1,x_2)$ avec x_2 de la forme $(1,2,\dots)$. Et ainsi de suite.

(3) Voir l'algorithme 4 (dans lequel on affiche les valeurs de la chaîne au fur et à mesure). Si U est uniforme dans $[0;10]$ alors $\lfloor U\rfloor$ est uniforme dans $\{0,1,2,\dots,9\}$.

Algorithme 4 Metropolis

```
pot<-function(x)
{
  s=1
  for (i in 1:(N-1))
  {
    s=s*M[r[c[i]],r[c[i+1]]]
  }
  return(s)
}

n=1000
x=c(1,2,3,4,5,6,7,8,9,10)
print(x)
for (i in 1:n)
{
  u=runif(1,0,10)
  i=as.integer(u)+1
  j=(i+1) %% 10
  y=x
  y[i]=x[j]
  x[i]=y[j]
  v=runif(1,0,1)
  if (v<pot(y)/pot(x))
  {
    y=x
  }
  print(y)
}
```

(4) La loi π_0 est telle que $\pi_0(x) > 0$ pour tout x dans $\mathcal{S}_{10}$ donc, par théorème du cours, nous savons que la chaîne de Metropolis est irréductible, de loi invariante π_0. Un autre théorème du cours nous dit que

$$\frac{1}{n}\sum_{i=1}^{n} f(Y_i) \underset{n\to+\infty}{\overset{\text{p.s.}}{\longrightarrow}} \int_{\mathcal{S}_{10}} f(x)\pi_0(dx)$$

(la fonction f est bien π_0-intégrable puisque $\mathcal{S}_n$ est borné).