Projet de Monte Carlo

Irène DUPREZ & Guillaume LEVESQUE

Exercice 1

Simulation suivant la densité f

1. Supposons qu'il existe une constante m$\in \mathbb{R}_+^*$, telle que $m \times a \geq 1$, et une densité g, pour laquelle on dispose d'un générateur aléatoire telles que

$$\forall (x,y) \in \mathbb{R}^2, \psi(x,y) \leq mg(x,y). \tag{1}$$

Alors on a, en posant $M = m \times a \in \mathbb{R}_+^*$:

$$\forall (x,y) \in \mathbb{R}^2, f(x,y) = a\psi(x,y) \leq Mg(x,y)$$

Ainsi, nous pouvons appliquer la méthode d'acceptation-rejet de la façon suivante :

Soient $(U_n)_{n\geq 1}$ une suite de variable $i.i.d$ de loi $\mathcal{U}([0,1])$ et $(Z_{1,n}, Z_{2,n})_{n\geq 1}$ une suite de variable $i.i.d$ de loi de densité g, avec (U_m) indépendante de $Z_{1,m}, Z_{2,m}$ $\forall m \in \{1,...,n\}$. Z_T suit la loi de densité de f où

$$T := inf\{n \geq 1 : U_n \leq \frac{f(Z_n)}{Mg(Z_n)} = \frac{a\psi(Z_n)}{mag(Z_n)} = \frac{\psi(Z_n)}{mg(Z_n)}\}$$

qui ne depend pas de a. Donc, grâce à la méthode d'acceptation-rejet, de m et de g, il n'est pas nécessaire, ici, de connaître a pour simuler suivant f.

Trouvons maintenant m et g satisfaisant (1) ; Soit (x,y) $\in \mathbb{R}^2$,

$$\psi(x,y) = \left[\left|\sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right)\right| + 4\cos(x)^2 + y^4\right] \exp^{-2(x+|y|)} 2\pi \overbrace{\frac{1}{2}\mathbf{1}_{\{y\in[-1;1]\}}\frac{1}{\pi}\mathbf{1}_{\{x\in[-\frac{\pi}{2};\frac{\pi}{2}]\}}}^{g(x,y)}$$

Afin d'optimiser la simulation suivant le principe de la méthode d'acceptation-rejet nous allons chercher à approcher au mieux le supremum de la fonction $\left[\left|\sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right)\right| + 4\cos(x)^2 + y^4\right] \exp^{-2(x+|y|)} 2\pi$ sur $[\frac{-\pi}{2}; \frac{\pi}{2}] \otimes [-1;1]$.Tout d'abord remarquons que pour $x \in [-\frac{\pi}{2}; \frac{\pi}{2}]$:

$$0 \leq x^2 \leq \frac{\pi^2}{4} \Leftrightarrow -\frac{\pi}{4} \leq \frac{2}{\pi}x^2 - \frac{\pi}{4} \leq \frac{\pi}{4}$$

Ainsi $\left|\sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right)\right| \leq \frac{\sqrt{2}}{2}$. Et on a sur $[\frac{-\pi}{2}; \frac{\pi}{2}] \otimes [-1;1]$:

$$\upsilon(x,y) = \overbrace{\left[\frac{\sqrt{2}}{2} + 4\cos(x)^2 + y^4\right] \exp^{-2x}}^{\tau(x,y)} \exp^{-2|y|} 2\pi = \overbrace{\tau(x,y)}^{\geq 0}\overbrace{\exp^{-2|y|} 2\pi}^{\geq 0}$$

On remarque alors que maximiser $\upsilon(x,y)$ peut se faire en maximisant $\tau(x,y)$ en fonction de x puis en maximisant le tout en fonction de y, les deux termes étant positifs.

$$\frac{\partial}{\partial x}\tau(x,y) = -2\exp^{-2x}\left[\frac{\sqrt{2}}{2} + 4\cos(x)^2 + y^4\right] + \exp^{-2x}(-8\cos(x)\sin(x))$$

En factorisant par $\exp^{-2x}$ on obtient les signes suivants sur $[-\frac{\pi}{2};\frac{\pi}{2}]$:

$$\frac{\partial}{\partial x}\tau(x,y)=\overbrace{\exp^{-2x}}^{\geq 0}\left[-\sqrt{2}-8\overbrace{\cos(x)}^{\geq 0}\overbrace{(\cos(x)+\sin(x))}^{\geq 0\ sur\ [-\frac{\pi}{4};\frac{\pi}{2}]}-2\overbrace{y^4}^{\geq 0}\right]$$

Ce qui implique que la fonction $\tau(x,y)$ sera décroissante en x sur $[-\frac{\pi}{4};\frac{\pi}{2}]$. Mais qu'en est-il du signe de $\left[-\sqrt{2}-8\cos(x)\left(\cos(x)+\sin(x)\right)-2y^4\right]$ sur $[-\frac{\pi}{2};-\frac{\pi}{4}]$?

En resolvant l'équation $\left[-\sqrt{2}-8\cos(x)\left(\cos(x)+\sin(x)\right)\right]=0$ sur cet intervalle on trouve deux solutions $x=-1.325$ et $x=-1.031$. Ainsi la fonction $\tau(x,y)$ sera croissante en x seulement pour $x\in[-1.325;-1.031]$, en supposant que y=0 (dès qu'on prend des y differents sur $[-1;1]$ cela ne fait que plus lisser $\tau(x,y)$ (toujours avec y variable muette), jusqu'à la rendre décroissante sur $[-\frac{\pi}{2};\frac{\pi}{2}]$)).

Ainsi nous pouvons retenir 2 candidats pour la maximisation de $\tau(x,y)$: $x=-1.031$ ou $x=-\frac{\pi}{2}$. $\tau(-\frac{\pi}{2},y)\approx 16.4+23y^4$ étant clairement supérieur à $\tau(-1.03,y)\approx 13.8+7.9y^4$ $\forall y\in[-1;1]$ on en déduit que la fonction $\tau(x,y)$ atteint son maximum en x=$-\frac{\pi}{2}$ sur $[-\frac{\pi}{2};\frac{\pi}{2}]$ lorsque $y\in[-1;1]$. Il nous suffit maintenant de maximiser $\varphi(y)=\left(\frac{\sqrt{2}}{2}+y^4\right)\exp^{-2|y|}\exp^{\pi}2\pi$ sur $[-1;1]$. Commençons d'abord avec l'intervalle $[-1;0]$ (où $|y|=-y$)

$$\frac{\partial\varphi(y)}{\partial y}=\left[\left(\frac{\sqrt{2}}{2}+y^4\right)2\exp^{2y}+\exp^{2y}4y^3\right]\exp^{\pi}2\pi=\left(\sqrt{2}+2y^4+4y^3\right)\exp^{2y}\exp^{\pi}2\pi$$

Etudions sur $[-1;0]$ le signe de :

$$\begin{aligned}\frac{\partial^2\varphi(y)}{\partial^2}&=\left[(\sqrt{2}+2y^4+4y^3)2\exp^{2y}+(8y^3+12y^2)\exp^{2y}\right]\exp^{\pi}2\pi\\&=(2\sqrt{2}+4y^4+4y^3+\overbrace{12y^3+12y^2}^{\geq 0\ car\ y^2-y^3\geq 0\ sur\ [-1;0]})\exp^{2y}\exp^{\pi}2\pi\\&\geq(2\sqrt{2}+4y^4+4y^3)\overbrace{\exp^{2y}\exp^{\pi}2\pi}^{\geq 0}\end{aligned}\tag{2}$$

Etudions maintenant le signe de$(2\sqrt{2}+4y^4+4y^3)=4(\frac{\sqrt{2}}{2}+y^4+y^3)$

$$\frac{\partial}{\partial}4(\frac{\sqrt{2}}{2}+y^4+y^3)=4(4y^3+3y^2)=\overbrace{4y^2}^{\geq 0}(4y+3)\geq 0\ pour\ y\geq\frac{-3}{4}$$

Donc $4(\frac{\sqrt{2}}{2}+y^4+y^3)$ est décroissante sur $[-1;\frac{-3}{4}]$ et croissante sur $[\frac{-3}{4};0]$ Donc admet un minimum en $y=\frac{-3}{4}$ qui est $4(\frac{\sqrt{2}}{2}+(\frac{-3}{4})^4+(\frac{-3}{4})^3)=4(0.705+0.316-0.42)\geq 0$. Donc $4(\frac{\sqrt{2}}{2}+y^4+y^3)\geq 0$ $\forall y\in[-1;0]$ et :

$$\frac{\partial^2\varphi(y)}{\partial^2}\geq(2\sqrt{2}+4y^4+4y^3)\overbrace{\exp^{2y}\exp^{\pi}2\pi}^{\geq 0}\geq 0$$

Ainsi $\varphi(y)$ est convexe sur $[-1;0]$, et par parité de $\varphi(y)$, sera convexe sur $[0;1]$ (car les fonctions "valeur absolue" et "puissance 4" sont des fonctions paires).

De ce fait le maximum de $\varphi(y)$ pour $y\in[-1;1]$ est en -1, 0 ou 1

$$\varphi(-1)=\overbrace{\left(\frac{\sqrt{2}}{2}+1\right)\exp^{-2}}^{\approx 0.23}\exp^{\pi}2\pi$$

$$\varphi(0) = \overbrace{\left(\frac{\sqrt{2}}{2}\right)}^{\approx 0.7} \exp^{\pi} 2\pi$$

$$\varphi(1) = \overbrace{\left(\frac{\sqrt{2}}{2} + 1\right) \exp^{-2}}^{\approx 0.23} \exp^{\pi} 2\pi$$

Donc $\varphi(y) \leq \varphi(0) = \frac{\sqrt{2}}{2} 2\pi \exp^{\pi} = \sqrt{2}\pi \exp^{\pi}$ sur $y \in [-1; 1]$. \ Nous avons donc m=$\sqrt{2}\pi \exp^{\pi}$ et g le produit de deux uniformes sur $[-\frac{\pi}{2}; \frac{\pi}{2}]$ et $[-1; 1]$

2. Application de la méthode d'acceptation-rejet avec R

```r
# Question 2

# Fonction qui simule selon g
rgen_g <- function(n) {
  g1 <- runif(n, min = (-pi / 2), max = (pi / 2))
  g2 <- runif(n, min = -1, max = 1)
  return(cbind(g1, g2))
}

# Fonction qui calcule psi(x,y)
psi <- function(x, y) {
  result1 <- abs(sin((2 / pi) * (x**2) - (pi / 4))) + (4 * cos(x)**2) + y**4
  result2 <- exp(-2 * (x + abs(y)))
  return(result1 * result2)
}

# Variables utilisees dans la suite de l'exercice
n <- 10000
m <- (sqrt(2) * pi) * exp(pi)
c_g <- Sys.time() # Temps de calcul pour question 6
g <- rgen_g(n)
c_g <- Sys.time() - c_g
c_u <- Sys.time() # Temps de calcul pour question 6
uni <- runif(n, 0, 1)
ratio <- psi(g[, 1], g[, 2]) / (m / (2 * pi))
c_u <- Sys.time() - c_u
val_ratio <- ratio

# Fonction qui simule selon f par la methode du rejet
rgen_f <- function(g_1, g_2, uni, ratio) {
  n_l <- 0
  while (uni > ratio) {
    uni <- runif(1, 0, 1)
    g <- rgen_g(1)
    g_1 <- g[1]
    g_2 <- g[2]
    ratio <- psi(g_1, g_2) / (m / (2 * pi))
    val_ratio <- c(val_ratio, ratio)
    n_l <- n_l + 1
```

```
  }
  assign("val_ratio", val_ratio, envir = .GlobalEnv)
  return(cbind(g_1, g_2, n_l))
}


# Question 3

c_f <- Sys.time() # Temps de calcul pour question 6
z <- Vectorize(rgen_f)(g[, 1], g[, 2], uni, ratio)
c_f <- Sys.time() - c_f
z <- t(z) # Pour une visualisation plus agreable
n_l <- sum(z[, 3]) / n
n_t <- length(val_ratio)
```

Estimation de a

Pour les 4 prochaines questions remarquons que supp(g)=supp(f) et que g(x,y)>0 sur supp(g) pour éviter les divisions par 0

4. a) D'après la question 1, on a

$$\forall (x,y) \in supp(g), \quad a = \frac{f(x,y)}{m\rho(x,y)g(x,y)}$$

Ainsi, on en déduit

$$\int_{supp(f)} a \, dxdy = a \overbrace{\int_{[\frac{-\pi}{2};\frac{\pi}{2}]\otimes[-1;1]} dxdy}^{Aire\ d'un\ rectangle\ o\grave{u}\ L=\pi\ et\ l=2} = a2\pi = \int_{supp(f)} f(x,y)\frac{1}{m\rho(x,y)g(x,y)} \, dxdy$$

D'où :

$$a = \mathbb{E}_f\left(\frac{1}{2\pi m\rho(X,Y)g(X,Y)}\right)$$

Or, on sait que supp(f)⊆supp(g) puisque supp(g)=supp(f). Donc par échantillonage préférentiel on obtient :

$$a = \mathbb{E}_g\left(\frac{h(X,Y)f(X,Y)}{g(X,Y)}\right) \text{ avec } h : (x,y) \mapsto \frac{1}{2\pi m\rho(x,y)g(x,y)} \tag{3}$$

Comme on connait la fonction f a une constante près, on utilise l'estimateur auto-normalisé :

$$\hat{b}_n(\rho) = \frac{\sum_{k=1}^n w_k h(Z_k)}{\sum_{k=1}^n w_k} \tag{4}$$

avec $Z_k = (Z_{1,k}, Z_{2,k})_{k\geq 1}$ une suite de variable $i.i.d$ de loi de densité g et $w_k = \frac{\psi(Z_k)}{g(Z_k)} = m\rho(Z_k)$

Biais :

$$\mathbb{E}_g\left(\hat{b}_n(\rho)\right) = \mathbb{E}_g\left(\frac{\sum_{k=1}^n w_k h(Z_k)}{\sum_{k=1}^n w_k}\right) \neq \frac{\mathbb{E}_g\left(\sum_{k=1}^n w_k h(Z_k)\right)}{\mathbb{E}_g\left(\sum_{k=1}^n w_k\right)}$$

Or, on a pour $i \in \{1, ..., n\}$ (en rappelant que supp(g)=supp(f)):

$$\mathbb{E}_g\left(w_i\right) = \mathbb{E}_g\left(m\rho(Z_i)\right) = \int_{supp(g)} m\frac{\psi(x,y)}{mg(x,y)}g(x,y)\ dxdy = \int_{supp(f)} \frac{1}{a}f(x,y)\ dxdy = \frac{1}{a}$$

$$\mathbb{E}_g\left(w_i h(Z_i)\right) = \mathbb{E}_g\left(\frac{m\rho(Z_i)}{2\pi m\rho(Z_i)g(Z_i)}\right) = \int_{supp(g)} \frac{1}{2\pi g(x,y)}g(x,y)\ dxdy = \frac{1}{2\pi}\overbrace{\int_{[\frac{-\pi}{2};\frac{\pi}{2}]\otimes[-1;1]} dxdy}^{2\pi} = 1$$

Ainsi, par linéarité de l'espérance on obtient :

$$\mathbb{E}_g\left(\hat{b}_n(\rho)\right) \neq \frac{\sum_{k=1}^n \mathbb{E}_g\left(w_k h(Z_k)\right)}{\sum_{k=1}^n \mathbb{E}_g\left(w_k\right)} \overset{i.i.d}{\widehat{=}} \frac{n\mathbb{E}_g\left(w_1 h(Z_1)\right)}{n\mathbb{E}_g\left(w_1\right)} = \frac{1}{\frac{1}{a}} = a$$

Les $(w_i)_{i\geq 1}$ et les $(w_i h(Z_i))_{i\geq 1}$ étant $i.i.d$ comme transformation mesurables des $(Z_i)_{i\geq 1}$ $i.i.d.$ Donc l'estimateur $\hat{b}_n(\rho)$ est biaisé

Convergence : Les $(w_i)_{i\geq 1}$ sont $i.i.d$ et intégrables car d'espérance $\frac{1}{a} < +\infty$ (car $a > 0$) suivant g. De ce fait, la loi forte des grands nombres nous permet d'écrire :

$$\frac{1}{n}\sum_{k=1}^n w_k \xrightarrow[\text{p.s.}]{n\to+\infty} \mathbb{E}_g(w_1) = \frac{1}{a} \tag{5}$$

De plus, on posant d : $\mathbb{R}_+^* \to \mathbb{R}_+^*$ la fonction fonction inverse sur les réels positifs. d est continue et d'après le "mapping theorem" on a :

$$\frac{1}{n}\sum_{k=1}^n w_k \xrightarrow[\text{p.s.}]{n\to+\infty} \frac{1}{a} \Rightarrow d(\frac{1}{n}\sum_{k=1}^n w_k) = \frac{n}{\sum_{k=1}^n w_k} \xrightarrow[\text{p.s.}]{n\to+\infty} d(\frac{1}{a}) = a$$

De la même façon, les $(w_i h(Z_i))_{i\geq 1}$ sont $i.i.d$ et intégrables, car d'espérance 1, suivant g. Toujours d'après la loi forte des grands nombres on a :

$$\frac{1}{n}\sum_{k=1}^n w_k h(Z_k) \xrightarrow[\text{p.s.}]{n\to+\infty} \mathbb{E}_g(w_1 h(Z_1)) = 1$$

Ainsi, nous pouvons conclure :

$$\frac{1}{n}\sum_{k=1}^n w_k h(Z_k)\frac{n}{\sum_{k=1}^n w_k} = \hat{b}_n(\rho) \xrightarrow[\text{p.s.}]{n\to+\infty} 1 \times a = a$$

Et l'estimateur $\hat{b}_n(\rho)$ de a converge presque sûrement vers a.

Intervalle de confiance asymptotique: On a toujours que les $(w_i)_{i\geq 1}$ sont $i.i.d$ et intégrables, car d'espérance 1/a (où $a > 0$), suivant g. Montrons maintenant qu'elles admettent un moment d'ordre 2 afin de pouvoir appliquer le TCL.

$$\begin{aligned} var_g(w_1) &= \mathbb{E}_g(w_1^2) - \mathbb{E}_g(w_1)^2 = m^2\mathbb{E}_g(\rho(Z_k)^2) - \frac{1}{a^2} = m^2\int_{supp(g)} \frac{\psi^2(x,y)}{m^2g^2(x,y)}g(x,y)\ dxdy - \frac{1}{a^2} \\ &= \int_{supp(g)} \frac{a\psi^2(x,y)}{ag(x,y)}\ dxdy - \frac{1}{a^2} = \int_{supp(f)} \frac{f(x,y)\psi(x,y)}{ag(x,y)}\ dxdy - \frac{1}{a^2} \\ &= \mathbb{E}_f\left(\frac{\psi(Z_k)}{ag(Z_k)}\right) - \frac{1}{a^2} \leq \mathbb{E}_f\left(\frac{mg(Z_k)}{ag(Z_k)}\right) - \frac{1}{a^2} = \frac{ma-1}{a^2} \text{ (d'après (1))} \end{aligned} \tag{6}$$

qui est finie car $a > 0$. Ainsi, d'après le théorème central limite, on a :

$$\frac{\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} w_k - \frac{1}{a}\right)}{\sqrt{var_g(w_1)}} \xrightarrow[\mathcal{L}]{n\to+\infty} \mathcal{N}(0,1)$$

En posant d : $\mathbb{R}_+^* \to \mathbb{R}_+^*$ la fonction fonction inverse sur les réels positifs. On remarque que d est dérivable et de dérivé $d'(\frac{1}{a}) = -a^2 \neq 0$. Ainsi d'après la Delta Méthode on obtient :

$$\begin{aligned}
\frac{\sqrt{n}\left(d(\frac{1}{n}\sum_{k=1}^{n} w_k) - d(\frac{1}{a})\right)}{\sqrt{var_g(w_1)}} \xrightarrow[\mathcal{L}]{n\to+\infty} \mathcal{N}(0, d'(\frac{1}{a})^2) &\Leftrightarrow \frac{\sqrt{n}\left(\frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k)} - a\right)}{\sqrt{var_g(w_1)}} \xrightarrow[\mathcal{L}]{n\to+\infty} \mathcal{N}(0, a^4) \\
&\Leftrightarrow \frac{\sqrt{n}\left(\frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k)} - a\right)}{\sqrt{var_g(w_1)}a^2} \xrightarrow[\mathcal{L}]{n\to+\infty} \mathcal{N}(0,1)
\end{aligned} \tag{7}$$

Or $\hat{b}_n(\rho)$ est un estimateur fortement consistant, donc consistant de a. Donc $\frac{a^2}{\hat{b}_n(\rho)^2}$ convergera en probabilité vers 1. Ainsi d'après le théorème de Slutsky :

$$\frac{\sqrt{n}\left(\frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k)} - a\right)}{\sqrt{var_g(w_1)}a^2} \frac{a^2}{\hat{b}_n(\rho)^2} \xrightarrow[\mathcal{L}]{n\to+\infty} 1 \times \mathcal{N}(0,1)$$

Par ailleurs, on sait que $\hat{\sigma_n^2} = \frac{1}{n-1}\sum_{k=1}^{n}(w_k - \frac{\sum_{k=1}^{n} w_k}{n})^2 \in \mathbb{R}_+^*$ converge $p.s$ vers $var_g(w_1) \in \mathbb{R}_+^*$ donc $\sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\left(w_k - \frac{\sum_{k=1}^{n} w_k}{n}\right)^2}$ converge $p.s$ vers $\sqrt{var_g(w_1)}$ par continuité de la fonction racine carré sur $\mathbb{R}_+^*$. Or la convergence $p.s$ du quotient vers 1 entrainant la convergence en probabilité. On obtient par Slutsky:

$$\frac{\sqrt{n}\left(\frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k)} - a\right)}{\sqrt{var_g(w_1)}\hat{b}_n(\rho)^2} \frac{\sqrt{var_g(w_1)}}{\sqrt{\hat{\sigma_n^2}}} \xrightarrow[\mathcal{L}]{n\to+\infty} 1 \times \mathcal{N}(0,1)$$

Ainsi,

$$\mathbb{P}_g\left(-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)} \leq \frac{\sqrt{n}\left(\frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k)} - a\right)}{\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2} \leq q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\right) = 1 - \alpha$$

Et on a :

$$\mathbb{P}_g\left(-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2 \leq \sqrt{n}\left(\frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k)} - a\right) \leq q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2\right) = 1 - \alpha$$

Où l'ordre des inégalités reste inchangé car $\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2 \geq 0$. On obtient alors :

$$\mathbb{P}_g\left(\frac{-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2}{\sqrt{n}} - \frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k} \leq (-a) \leq \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2}{\sqrt{n}} - \frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k}\right) = 1 - \alpha$$

Ou encore,

$$\mathbb{P}_g\left(\frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2}{\sqrt{n}} + \frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k} \geq a \geq \frac{-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}\hat{b}_n(\rho)^2}{\sqrt{n}} + \frac{1}{\frac{1}{n}\sum_{k=1}^{n} w_k}\right) = 1 - \alpha$$

Qui nous permet de conclure :

$$\text{IC}_a^{1-\alpha} = \left[\frac{1}{\frac{1}{n}\sum_{k=1}^n w_k} - \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)} \sqrt{\hat{\sigma_n^2}} \hat{b}_n(\rho)^2}{\sqrt{n}} ; \frac{1}{\frac{1}{n}\sum_{k=1}^n w_k} + \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)} \sqrt{\hat{\sigma_n^2}} \hat{b}_n(\rho)^2}{\sqrt{n}} \right]$$

Pour alléger le code R de la question suivante remarquons que dans notre cas $w_k h(Z_k)$ (codé sous le nom wk_hzk) est égale à :

$$w_k h(Z_k) = \frac{m\rho(Z_k)}{2\pi m\rho(Z_k) g(Z_k)} = \frac{1}{2\pi \times \frac{1}{2\pi}} = 1$$

```
# Question 4b)

c_b_hat <- Sys.time() # Temps de calcul pour question 6
# Evaluation de bn chapeau
wk <- 2 * pi * psi(g[, 1], g[, 2])
wk_hzk <- rep(1, n)
bn_hat <- sum(wk_hzk) / sum(wk)
c_b_hat <- Sys.time() - c_b_hat
# Evaluation de l'intervalle de confiance asymptotique
sigma_hat_wk <- (sum((wk - sum(wk) / n)**2)) / (n - 1)
ecart <- ((qnorm(0.975) * sqrt(sigma_hat_wk) * (bn_hat**2)) / sqrt(n))
ic_95_a_inf <- (n / sum(wk)) - ecart
ic_95_a_sup <- (n / sum(wk)) + ecart


# Question 4c)

# Estimation du biais de bn chapeau par une methode de bootstrap
k <- 1000
bn_bootstrap <- matrix(sample(wk, n * k, replace = TRUE), ncol = n)
bn_bootstrap <- n / rowSums(bn_bootstrap)
estim_biais <- mean(bn_bootstrap) - bn_hat
```

5. a) Si nous partons maintenant du principe que nous pouvons simuler suivant f, grâce à la première partie de cet exercice, alors nous allons pouvoir utiliser l'estimateur de Monte Carlo classique pour estimer a. Sachant que $f(x, y) = a\psi(x, y)$, on a que :

$$\int_{supp(f)} a\ dxdy = \int_{supp(f)} \frac{f(x, y)}{\psi(x, y)}\ dxdy = \mathbb{E}_f \left(\frac{1}{\psi(X)} \right) \tag{8}$$

Or,

$$\int_{supp(f)} a\ dxdy = a \overbrace{\int_{[\frac{-\pi}{2};\frac{\pi}{2}]\otimes[-1;1]} dxdy}^{Aire\ d'un\ rectangle\ où\ L=\pi\ et\ l=2} = a2\pi \tag{9}$$

Avec X est une variable aléatoire suivant la densité f. Posons $h : \mathbb{R}^2 \mapsto \mathbb{R}$ definit par $h(X) = \frac{1}{2\pi\psi(X)}$. Soit $(X_i)_{i\in\{1,\ldots,n\}}$ une suite de variables aléatoires $i.i.d$ suivant la densité f alors les $(h(X_i))_{i\in\{1,\ldots,n\}}$ sont $i.i.d$, car transformations mesurables des $(X_i)_{i\in\{1,\ldots,n\}}$, et intégrables par rapport à f d'après (8) et (9) (a étant fini). Donc d'après la loi forte des grands nombres :

$$\hat{a}_n = \frac{1}{n}\sum_{i=1}^n h(X_i) \xrightarrow[\text{p.s.}]{n\to+\infty} \mathbb{E}_f(h(X_1)) = \frac{1}{2\pi}\mathbb{E}_f\left(\frac{1}{\psi(X)}\right) = \frac{1}{2\pi}a2\pi = a$$

Donc $\hat{a}_n$ est un estimateur fortement consistant de a.

Biais : Par linéarité de l'espérance on a :

$$\mathbb{E}_f(\hat{a}_n) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_f(h(X_i)) \overbrace{=}^{i.i.d} \mathbb{E}_f(h(X_1)) = a$$

Ainsi $\hat{a}_n$ est également un estimateur non biaisé de a.

Intervalle de confiance asymptotique : On sait d'après la question précédente que les $h(X_i)_{i\in\{1,...,n\}}$ sont *i.i.d* et intégrables par rapport à f, mais admettent-elles un moment d'ordre 2 ? Soit $i \in \{1, ..., n\}$

$$\begin{aligned}
Var_f(h(X_i)) &= \mathbb{E}_f(h(X_i)^2) - \mathbb{E}_f(h(X_i))^2 = \mathbb{E}_f\left(\left[\frac{1}{2\pi\psi(X_i)}\right]^2\right) - a^2 \\
&= \int_{supp(f)} \frac{1}{4\pi^2\psi(x,y)^2} f(x,y)\frac{a}{a}\, dxdy - a^2 = \frac{a}{4\pi^2}\int_{supp(f)} \frac{1}{\psi(x,y)}\, dxdy - a^2 \\
&= \frac{a}{4\pi^2}\int_{supp(f)} \frac{\exp\left(2(x+|y|)\right)}{\left[\left|\sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right)\right| + 4\cos(x)^2 + y^4\right]}\, dxdy - a^2 \\
&\leq \frac{a}{4\pi^2}\int_{supp(f)} \underbrace{\frac{\exp\left(\pi+2\right)}{\left[\left|\sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right)\right| + 4\cos(x)^2 + y^4\right]}}_{\text{est continu et défini sur supp(f) car (*)}}\, dxdy - a^2
\end{aligned}$$

(*) : le dénorminateur ne s'annule pas sur supp(f) car il est positif (par positivité de la valeur absolue, de la fonction carré et de la fonction "puissance 4") et différent de 0 puisque $\cos(x)^2 = 0$ sur $x \in [-\frac{\pi}{2}; \frac{\pi}{2}]$ quand $x = -\frac{\pi}{2}$ ou $x = \frac{\pi}{2}$ mais dans ce cas $\left|\sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right)\right| = \sin(\frac{\pi}{4}) = \frac{\sqrt{2}}{2} \neq 0$. Donc l'intégrale sera finie et les $h(X_i)_{i\in\{1,...,n\}}$ admettent un moment d'ordre 2. De plus leur variance sera différente de 0 car ce ne sont pas des diracs.

Nous pouvons donc appliquer le théorème central limite de la façon suivante :

$$\sqrt{\frac{n}{var_f(h(X_1))}}(\hat{a}_n - a) \xrightarrow[\mathcal{L}]{n\to+\infty} \mathcal{N}(0,1)$$

Par ailleurs, nous savons que $\hat{\sigma_n^2} = \frac{1}{n-1}\sum_{k=1}^{n}(h(X_i) - \frac{\sum_{k=1}^{n} h(X_i)}{n})^2 \in \mathbb{R}_+^*$ converge *p.s* vers $var_f(h(X_1)) \in \mathbb{R}_+^*$ donc $\sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\left(h(X_i) - \frac{\sum_{k=1}^{n} h(X_i)}{n}\right)^2}$ converge *p.s* vers $\sqrt{var_f(h(X_1))}$ par continuité de la fonction racine carré sur $\mathbb{R}_+^*$. La convergence *p.s* du quotient vers 1 entrainant la convergence en probabilité, on obtient par Slutsky:

$$\sqrt{\frac{n}{var_f(h(X_1))}}\frac{\sqrt{var_f(h(X_1))}}{\sqrt{\hat{\sigma_n^2}}}(\hat{a}_n - a) \xrightarrow[\mathcal{L}]{n\to+\infty} 1 \times \mathcal{N}(0,1)$$

Qui nous donne :

$$\mathbb{P}_f\left(\frac{-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}}{\sqrt{n}} - \hat{a}_n \leq -a \leq \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}}{\sqrt{n}} - \hat{a}_n\right) = 1 - \alpha$$

ou encore,

$$\mathbb{P}_f\left(\frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}}{\sqrt{n}} + \hat{a}_n \geq a \geq \frac{-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma_n^2}}}{\sqrt{n}} + \hat{a}_n\right) = 1 - \alpha$$

Et on obtient l'intervalle de confiance asymptotique de a au niveau $1-\alpha$ suivant :

$$\mathrm{IC}_a^{1-\alpha} = \left[\hat{a}_n - \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)} \sqrt{\hat{\sigma_n^2}}}{\sqrt{n}} ; \hat{a}_n + \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)} \sqrt{\hat{\sigma_n^2}}}{\sqrt{n}} \right]$$

```
# Question 5b)

c_a_hat <- Sys.time() # Temps de calcul pour question 6
# Estimation de an chapeau
h_xi <- 1 / (2 * pi * psi(z[, 1], z[, 2]))
an_hat <- (sum(h_xi)) / n
c_a_hat <- Sys.time() - c_a_hat
# Evaluation de l'intervalle de confiance asymptotique
sigma_hat_hx <- sqrt(sum((h_xi + sum(h_xi) / n)**2) / (n - 1))
ic2_95_a_inf <- an_hat - ((qnorm(0.975) * sigma_hat_hx) / sqrt(n))
ic2_95_a_sup <- an_hat + ((qnorm(0.975) * sigma_hat_hx) / sqrt(n))
```

6. Les coûts d'estimation de $\hat{b}_n$ et $\hat{a}_n$, pour une certaine précision ϵ, sont les suivants :

   - coût $\hat{b}_n$: $c_b \frac{var(\hat{b}_n)}{\epsilon^2}$, $\quad c_b = n(c_\rho + c_g)$
   - coût $\hat{a}_n$: $c_a \frac{var(\hat{a}_n)}{\epsilon^2}$, $\quad c_a = n(c_h + c_f)$

   On obtient donc le rapport des coûts suivant :

$$\mathcal{R}\left(\hat{b}_n, \hat{a}_n\right) = \frac{c_b var\left(\hat{b}_n\right)}{c_a var\left(\hat{a}_n\right)} \tag{10}$$

   On peut estimer $var\left(\hat{b}_n\right)$ grace a une methode de bootstrap par $\frac{1}{n-1}\sum_{i=1}^n \left(\hat{b}_n^{(i)} - \sum_{j=1}^n \hat{b}_n^{(i)}\right)^2$ où $\hat{b}_n^{(i)} = \frac{\sum_{k=1}^n w_k^{(i)} h(Z_k^{(i)})}{\sum_{k=1}^n w_k^{(i)}}$ avec $Z_k^{(i)} = (Z_{1,k}^{(i)}, Z_{2,k}^{(i)})_{k\geq 1} \overset{iid}{\sim} g$ et $w_k^{(i)} = \frac{\psi(Z_k^{(i)})}{g(Z_k^{(i)})} = m\rho(Z_k^{(i)})$, $k \in \{1, ..., K\}$.

   On peut estimer $var\left(\hat{a}_n\right)$ par $\frac{1}{n}\hat{\sigma}_n^2$ (de la question 5b)), $\hat{a}_n$ étant l'estimateur de Monte Carlo classique.

```
# Question 6

# Estimation du cout de calcul des estimateurs
c_b <- c_g + c_b_hat
c_a <- c_f + c_a_hat + c_u
# Estimation des variances des estimateurs
sigma_hat_b <- var(bn_bootstrap)
sigma_hat_a <- (sigma_hat_hx^2) / n
# Efficacite relative de bn chapeau par rapport a an chapeau
r_b_a <- (as.numeric(c_b) * sigma_hat_b) / (as.numeric(c_a) * sigma_hat_a)
```

Comme $\mathcal{R}\left(\hat{b}_n, \hat{a}_n\right) \ll 1$, on conclut que $\hat{b}_n$ est bien plus efficace que $\hat{a}_n$.

7. a) On note :

   $c_x = \left[\left|\sin\left(\frac{2}{\pi}x^2 - \frac{\pi}{4}\right)\right| + 4\cos(x)^2\right]$

Pour $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, on a :

$$f_X(x) = ae^{-2x} \int_0^1 (c(x) + y^4)e^{-2y}\ dy + ae^{-2x} \int_{-1}^0 (c(x) + y^4)e^{2y}\ dy$$

En faisant un changement de variable $z = -y$ dans la deuxième intégrale on obtient :

$$\begin{aligned}
&= ae^{-2x} \int_0^1 (c(x) + y^4)e^{-2y}\ dy - ae^{-2x} \int_1^0 (c(x) + z^4)e^{-2z}\ dz \\
&= ae^{-2x} \int_0^1 (c(x) + y^4)e^{-2y}\ dy + ae^{-2x} \int_0^1 (c(x) + z^4)e^{-2z}\ dz \\
&= 2ae^{-2x} \int_0^1 (c(x) + y^4)e^{-2y}\ dy = \mathbb{E}_{\mathcal{U}[0,1]}(2ae^{-2x} h_x(U))
\end{aligned} \tag{11}$$

avec $h_x : y \mapsto (c_x + y^4)e^{-2y}$. Ainsi on obtient l'estimateur de Monte Carlo suivant :

$$\hat{\delta}_{X,n}(x) = \frac{2ae^{-2x}}{n} \sum_{k=1}^{n} h_x(X_k)$$

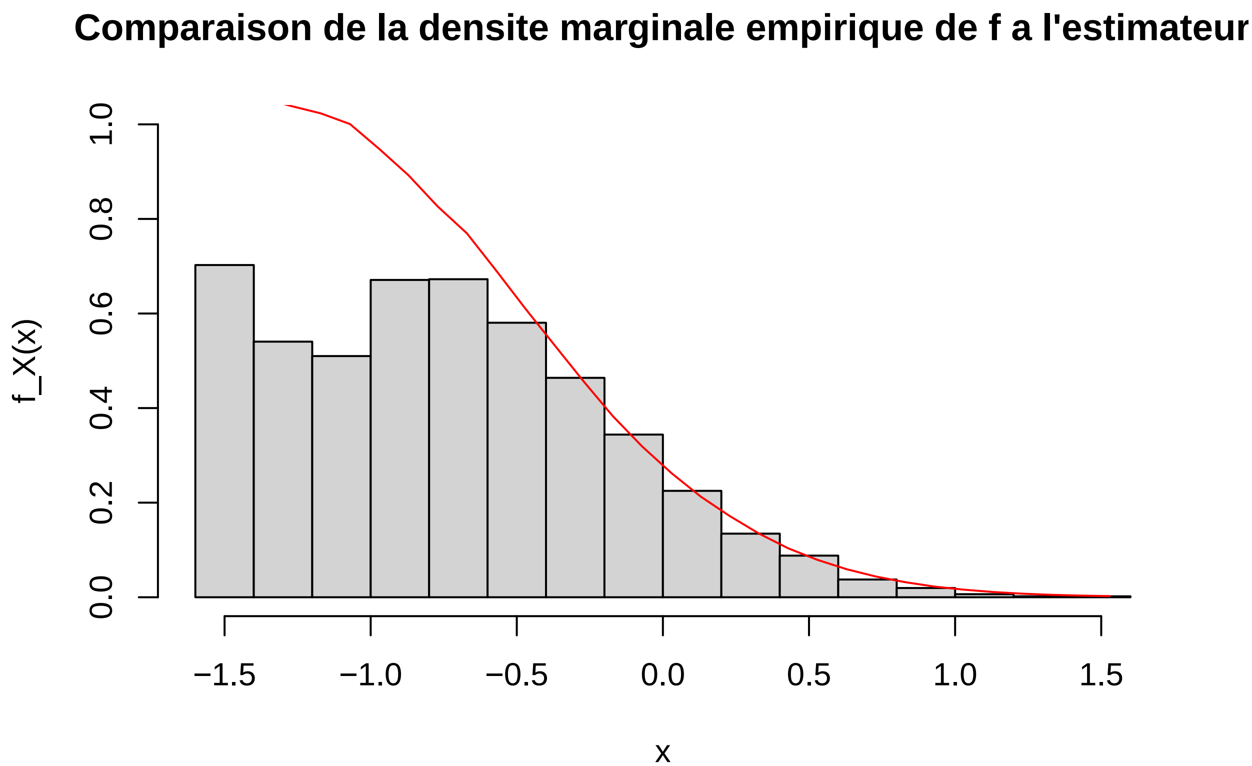
avec $(X_k)_{k\geq 1}$ une suite de variable $i.i.d$ suivant une loi uniforme sur $[0, 1]$

Comme $\hat{\delta}_{X,n}(x)$ et $\hat{a}_n$ sont des estimateurs consistans respectivement de $f_X(x)$ et de a, on a par le théorème de Slutsky :

$$\hat{f}_{X,n}(x) = \frac{2\hat{a}_n e^{-2x}}{n} \sum_{k=1}^{n} h_x(X_k) = \frac{2ae^{-2x}}{n} \sum_{k=1}^{n} h_x(X_k) \times \frac{\hat{a}_n}{a} \xrightarrow[\mathcal{L}]{n\to+\infty} 1 \times f_X(x) \tag{12}$$

Qui nous donne un estimateur de $f_X(x)$ en fonction de $\hat{a}_n$ sur $[\frac{-\pi}{2}; \frac{\pi}{2}]$.

```r
# Question 7b)

# Fonction qui a x et n associe fXn(x) chapeau
fn_hat_x <- function(x, n) {
  x_i <- matrix(runif(n * length(x)), ncol = n)
  # Et non simplement runif(n) afin que la fonction supporte un vecteur
  # en entree et qu'elle puisse ainsi nous permettre le tracer
  # de sa courbe dans les lignes ci-dessous.
  cx <- abs(sin((2 * (x^2) / pi) + (pi / 4))) + 4 * (cos(x)^2)
  hx_x_i <- (cx + x_i^4) * exp(-2 * x_i)
  resultat <- 2 * an_hat * exp(-2 * x) * rowSums(hx_x_i) / n
  return(resultat)
}

titre <- "Comparaison de la densite marginale empirique de f a l'estimateur"
axex <- "x"
axey <- "f_X(x)"
t <- seq(-pi / 2, pi / 2, 0.1)
hist(z[, 1], freq = FALSE, ylim = c(0, 1), xlab = axex, ylab = axey, main = titre)
lines(t, fn_hat_x(t, n), col = "red")
```



Estimateur ponctuel

8. On note $U_k^{(x)} = \frac{\psi(x,Y_k)w(X_k)}{\psi(X_k,Y_k)}$

   On a :

$$\begin{aligned}
\mathbb{E}_f(U_1^{(x)}) &= \int_{supp(f)} \frac{\psi(x,v)w(u)}{\psi(u,v)} f(u,v)dudv \\
&= \int_{supp(f)} \frac{1}{a} f(x,v)w(u) \frac{a}{f(u,v)} f(u,v)dudv \\
&= \int_{supp(f)} f(x,v)w(u)dudv \\
&\overbrace{=}^{\text{Fubini positif}} \underbrace{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} w(u)du}_{=1 \text{ car supp(w)} \subseteq \text{supp}(f_X) = [-\frac{\pi}{2};\frac{\pi}{2}]} \underbrace{\int_{-1}^{1} f(x,v)dv}_{=f_X(x)} \\
&= f_X(x)
\end{aligned}$$

Convergence : Ainsi les $\left(U_k^{(x)}\right)_{k\in\{1,\ldots,n\}}$ étant *i.i.d* comme transformation mesurable des $(X_i,Y_i)_{i\in\{1,\ldots,n\}}$ et intégrables. On a par la loi forte des grands nombres :

$$\frac{1}{n}\sum_{k=1}^{n} U_k^{(x)} = \hat{w}_n(x) \xrightarrow[\text{p.s.}]{n\to+\infty} \mathbb{E}_f(U_1^{(x)}) = f_X(x)$$

Intervalle de confiance asymptotique : On sait que les $\left(U_k^{(x)}\right)_{k\in\{1,\dots,n\}}$ sont $i.i.d$ et intégrables par rapport à f. Montrons qu'elles admettent moment d'ordre 2.

$$\begin{aligned}
\mathbb{E}_f\left[\left(U_k^{(x)}\right)^2\right] &= \int_{supp(f)} \frac{f(x,v)^2 w(u)^2}{f(u,v)^2} f(u,v)dudv \\
&= \int_{supp(f)} \frac{f(x,v)^2 w(u)^2}{f(u,v)} dudv \\
&< \infty \quad \text{car } f \text{ ne s'annule pas sur son support}
\end{aligned}$$

De plus leur variance est différente de 0 car ce ne sont pas des diracs.

Nous pouvons donc appliquer le théorème central limite avec $\left(U_k^{(x)}\right)_{k\in\{1,\dots,n\}}$:

$$\sqrt{\frac{n}{var_f\left(U_1^{(x)}\right)}}(\hat{w}_n(x) - f_X(x)) \xrightarrow[\mathcal{L}]{n\to+\infty} \mathcal{N}(0,1)$$

Par ailleurs, nous savons que $\hat{\sigma}^2_{U,n}(x) = \frac{1}{n-1}\sum_{k=1}^n \left(U_k^{(x)} - \hat{w}_n(x)\right)^2 \in \mathbb{R}_+^*$ converge $p.s$ vers $var_f\left(U_1^{(x)}\right) \in \mathbb{R}_+^*$ donc $\sqrt{\hat{\sigma}^2_{U,n}(x) = \frac{1}{n-1}\sum_{k=1}^n \left(U_k^{(x)} - \hat{w}_n(x)\right)^2}$ converge $p.s$ vers $\sqrt{var_f\left(U_1^{(x)}\right)}$ par continuité de la fonction racine carré sur $\mathbb{R}_+^*$. La convergence $p.s$ du quotient vers 1 entrainant la convergence en probabilité, on obtient par Slutsky:

$$\sqrt{\frac{var_f\left(U_1^{(x)}\right)}{\hat{\sigma}^2_{U,n}(x)}}\sqrt{\frac{n}{var_f\left(U_1^{(x)}\right)}}(\hat{w}_n(x) - f_X(x)) = \sqrt{\frac{n}{\hat{\sigma}^2_{U,n}(x)}}(\hat{w}_n(x) - f_X(x)) \xrightarrow[\mathcal{L}]{n\to+\infty} 1\times\mathcal{N}(0,1)$$

Ainsi nous obtenons :

$$\mathbb{P}_f\left(\frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma}^2_{U,n}(x)}}{\sqrt{n}} + \hat{w}_n(x) \geq f_X(x) \geq \frac{-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma}^2_{U,n}(x)}}{\sqrt{n}} + \hat{w}_n(x)\right) = 1-\alpha$$

Et on obtient l'intervalle de confiance asymptotique de $f_X(x)$, en fonction de x, au niveau $1-\alpha$ suivant :

$$\mathrm{IC}_{f_X}^{1-\alpha}(x) = \left[\hat{w}_n(x) - \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma}^2_{U,n}(x)}}{\sqrt{n}} ; \hat{w}_n(x) + \frac{q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\sqrt{\hat{\sigma}^2_{U,n}(x)}}{\sqrt{n}}\right]$$

9. La densité w^* telle que l'estimateur est de variance minimale est définie par:

$$\begin{aligned}
w^* &= \arg\min_w \{var_f(\hat{w}_n(x))\} \\
&= \arg\min_w \left\{var_f\left(\frac{1}{n}\sum_{k=1}^n U_k^{(x)}\right)\right\} \\
&= \arg\min_w \left\{\frac{1}{n} var\left(U_1^{(x)}\right)\right\} \\
&\text{Les } \left(U_k^{(x)}\right)_{k\in\{1,\dots,n\}} \text{ étant } i.i.d
\end{aligned}$$

Or, par indépendance de X et Y, on a :

$$\begin{aligned} U_1^{(x)} &= \frac{\psi(x, Y_1) w(X_1)}{\psi(X_1, Y_1)} \times \frac{a}{a} \\ &= \frac{f_{X,Y}(x, Y_1) w(X_1)}{f_{X,Y}(X_1, Y_1)} \\ &= \frac{f_X(x) f_Y(Y_1) w(X_1)}{f_X(X_1) f_Y(Y_1)} \\ &= \frac{f_X(x) w(X_1)}{f_X(X_1)} \end{aligned}$$

Ainsi, on a $var_f(\hat{w}_n(x)) = \frac{f_X(x)^2}{n} var_f\left(\frac{w(X)}{f_X(X)}\right)$, et la variance atteint son minimum pour tout w tq $\exists \lambda \in \mathbb{R}, w = \lambda f_X$. Comme w est une densité, on a forcement $w^* = f_X$.

Ce résultat parait difficilement exploitable, puisqu'il implique de connaitre le paramètre que l'on cherche a estimer. Cependant, en pratique, on peut utiliser la densité empirique de notre echantillon z.

```r
# Question 10

# Liste des coordonnees (x,y) qui forment la densite empirique associee a z[,1]
fx_emp <- density(z[, 1], from = -1.6, to = 1.6, n = 321)
fx_emp <- cbind(fx_emp[["x"]], fx_emp[["y"]])

# Fonction qui a x associe fX(x)
# Calcule empiriquement avec l'echantillon z
densite_emp <- function(x) {
  resultat <- fx_emp[which(round(fx_emp[, 1], 2) == round(x, 2)), 2]
  return(cbind(resultat))
}

c_w_hat <- Sys.time() # Temps de calcul pour question 11
# Estimation de wn(-1) chapeau
x <- -1
wk_emp <- (sapply(z[, 1], densite_emp))
uk_x <- (psi(x, z[, 2]) * wk_emp) / psi(z[, 1], z[, 2])
w_hat_x <- sum(uk_x) / n
c_w_hat <- Sys.time() - c_w_hat
# Evaluation de l'intervalle de confiance asymptotique de fX(-1)
sigma_hat_uk <- var(uk_x)
ic_95_fx_inf <- w_hat_x - qnorm(0.975) * sqrt(sigma_hat_uk / n)
ic_95_fx_sup <- w_hat_x + qnorm(0.975) * sqrt(sigma_hat_uk / n)
```

11. Les coûts d'estimation de $\hat{f}_{X,n}(-1)$ et $\hat{w}_n(-1)$, pour une certaine précision ϵ, sont les suivants :

- coût $\hat{f}_{X,n}(-1)$: $c_{f_X(-1)} \frac{var\left(\hat{f}_{X,n}(-1)\right)}{\epsilon^2}, \quad c_{f_X(-1)} = n(c_a + c_{\mathcal{U}} + c_{h_{-1}})$
- coût $\hat{w}_n(-1)$: $c_{w(-1)} \frac{var(\hat{w}_n(-1))}{\epsilon^2}, \quad c_{w(-1)} = n(c_w + c_f + c_\psi)$

On obtient donc le rapport des coûts suivant :

$$\mathcal{R}\left(\hat{w}_n(-1), \hat{f}_{X,n}(-1)\right) = \frac{c_{w(-1)} var\left(\hat{w}_n(-1)\right)}{c_{f_X(-1)} var\left(\hat{f}_{X,n}(-1)\right)} \tag{13}$$

On peut estimer $var\left(\hat{w}_n(-1)\right)$ grace a une methode de bootstrap, et $var\left(\hat{f}_{X,n}(-1)\right) = \frac{4\hat{a}_n^2 e^4}{n} var\left(h_{-1}(X)\right)$ a l'aide de la variance empirique de $h_{-1}(X)$.

```r
# Question 11
c_fx_hat <- Sys.time() # Temps de calcul
# Estimation de fXn(-1) chapeau
x_i <- runif(n)
cmoins1 <- abs(sin((2 * (x^2) / pi) + (pi / 4))) + 4 * (cos(x)^2)
hmoins1_x <- (cmoins1 + x_i^4) * exp(-2 * x_i)
fn_hat_moins1 <- 2 * an_hat * exp(-2 * x) * sum(hmoins1_x) / n
c_fx_hat <- Sys.time() - c_fx_hat
# Estimation des variance de fXn(-1) chapeau et wn(-1) chapeau
uk_x_bootstrap <- matrix(sample(uk_x, n * k, replace = TRUE), ncol = n)
w_x_bootstrap <- rowSums(uk_x_bootstrap) / n
sigma_hat_fx <- var(hmoins1_x) * ((4 * (an_hat^2) * exp(4))) / n
sigma_hat_w <- var(w_x_bootstrap)
# Estimation du cout de calcul de estimateurs
c_w_moins1 <- as.numeric(c_w_hat + c_f + c_u)
c_fx_moins1 <- as.numeric(c_a + c_fx_hat)
# Efficacite relative de wn(-1) chapeau par rapport a fXn(-1) chapeau
r_fx_w_moins1 <- (c_w_moins1 * sigma_hat_w) / (c_fx_moins1 * sigma_hat_fx)
```

Comme $\mathcal{R}\left(\hat{w}_n(-1), \hat{f}_{X,n}(-1)\right) \ll 1$, on conclut que $\hat{w}_n(-1)$ est plus efficace que $\hat{f}_{X,n}(-1)$.

Exercice 2

1.

```r
# Question 1

# Fonction qui simule selon une loi normale multivariee
rmvnorm <- function(n, mu, sigma) {
  normale <- matrix(rnorm(length(mu) * n), nrow = length(mu))
  return(t(chol(sigma)) %*% normale + mu)
  # chol() donne la matrice triangulaire superieur de Cholesky
  # --> On transpose
}

# Variables utilisees dans la suite de l'exercice
mu <- c(0.1, 0, 0.1)
sigma <- matrix(c(0.047, 0, 0.0117, 0, 0.047, 0, 0.0117, 0, 0.047), nrow = 3)
n <- 10000
x <- rmvnorm(n, mu, sigma)
x <- t(x) # Pour une visualisation plus agreable
```

2. a) On a $\delta = \mathbb{E}(h(\mathbf{X}))$ avec $h : (x_1, x_2, x_3) \mapsto \min\left(3, \frac{1}{3}\sum_{k=1}^{3} e^{-x_k}\right)$. On définit donc l'estimateur de Monte Carlo comme suit :

$$\bar{\delta}_n = \frac{1}{n}\sum_{k=1}^{n} h(\mathbf{X}_k) \tag{14}$$

   b) Comme $\bar{\delta}_n$ est un estimateur sans biais de δ, on a $\text{MSE}_{\bar{\delta}_n} = var(\bar{\delta}_n) = \frac{1}{n} var(h(\mathbf{X}))$

```r
# Question 2b)

# Application de h a notre echantillon x
hx <- rowSums(exp(-x)) / 3
hx[hx > 3] <- 3
delta_barre <- mean(hx)
mse_barre <- var(hx) / n
```

3. a) On pose $A : x \mapsto 2\mu - x$. $A(\mathbf{X})$ est bien une transformation mesurable de $\mathbf{X}$ laissant la loi $\mathcal{N}(\mu, \Sigma)$ invariante.

   Ainsi, $A(\mathbf{X}) \sim \mathcal{N}(\mu, \Sigma)$, c'est à dire $A(\mathbf{X})$ et $\mathbf{X}$ suivent la même loi, et donc $h(\mathbf{X})$ et $h \circ A(\mathbf{X})$ aussi.

   Comme h et A sont décroissantes, $h \circ A$ est croissante. Ainsi, $cov(h(\mathbf{X}), h \circ A(\mathbf{X})) < 0$ et on a bien $var(\hat{\delta}_n) \leq \frac{var(\bar{\delta}_n)}{2}$ avec

$$\hat{\delta}_n = \frac{1}{2n} \sum_{k=1}^{n} h(\mathbf{X}_k) + h \circ A(\mathbf{X}_k) \tag{15}$$

   l'estimateur de δ par la methode de la variable antithétique.

   On a,

$$\begin{aligned} var(\hat{\delta}_n) &= var\left(\frac{1}{2n} \sum_{k=1}^{n} h(\mathbf{X}_k) + h \circ A(\mathbf{X}_k)\right) \\ &\overbrace{=}^{i.i.d} \frac{1}{4n} var(h(\mathbf{X}) + h \circ A(\mathbf{X})) \\ &= \frac{1}{4n}\left[var(h(\mathbf{X})) + var(h \circ A(\mathbf{X})) + 2cov(h(\mathbf{X}), h \circ A(\mathbf{X}))\right] \\ &= \frac{1}{2n} var(h(\mathbf{X})) + \frac{1}{2n} cov(h(\mathbf{X}), h \circ A(\mathbf{X})) \quad \text{car } h(\mathbf{X}) \text{ et } h \circ A(\mathbf{X}) \text{ ont même loi.} \end{aligned} \tag{16}$$

   Ainsi,

$$\begin{aligned} R_1 &= \frac{\frac{1}{2n} var(h(\mathbf{X})) + \frac{1}{2n} cov(h(\mathbf{X}), h \circ A(\mathbf{X}))}{\frac{1}{n} var(h(\mathbf{X}))} \\ &= \frac{1}{2}\left(1 + \frac{cov(h(\mathbf{X}), h \circ A(\mathbf{X}))}{var(h(\mathbf{X}))}\right) \end{aligned} \tag{17}$$

   b) On a que :

   $\hat{\delta}_n$ est un estimateur sans biais de δ, donc $\text{MSE}_{\hat{\delta}_n} = var(\hat{\delta}_n) \overbrace{=}^{i.i.d} \frac{1}{n} var\left(\frac{h(\mathbf{X}) + h \circ A(\mathbf{X})}{2}\right)$

```r
# Question 3b)

# Application de h(A) a notre echantillon x
ax <- 2 * mu - t(x)
ax <- t(ax) # Pour une visualisation plus agreable
hax <- rowSums(exp(-ax)) / 3
hax[hax > 3] <- 3
```

```
delta_hat <- sum(hx + hax) / (2 * n)
mse_hat <- var((hx + hax) / 2) / n
r1 <- mse_hat / mse_barre
print(r1)
```

```
## [1] 0.02028074
```

Dans notre cas la méthode de la variable antithétique a reduit la variance de 98% par rapport à l'estimateur de Monte Carlo classique.

4. a) On teste les 3 fonctions suivantes :

   - $h_{0,1} : (x_1, x_2, x_3) \mapsto \frac{1}{3}\sum_{i=1}^{3} x_i$
   - $h_{0,2} : (x_1, x_2, x_3) \mapsto \frac{1}{3}\sum_{i=1}^{3} \left(x_i - \frac{1}{3}\sum_{j=1}^{3} x_j\right)^2$
   - $h_{0,3} : (x_1, x_2, x_3) \mapsto \frac{1}{3}\sum_{i=1}^{3} x_i^2$

   et on compare $\rho(h(\mathbf{X}), h_{0,i}(\mathbf{X})), \quad i = 1, 2, 3.$

```
# Question 4 a)

# Test 1
h0x1 <- rowSums(x) / 3
rho1 <- cor(hx, h0x1)
# Test 2
h0x2 <- rowSums((x - colSums(x) / 3)^2) / 2
rho2 <- cor(hx, h0x2)
# Test 3
h0x3 <- rowSums(x^2) / 3
rho3 <- cor(hx, h0x3)
```

Ainsi, on choisit $h_0 = h_{0,1}$ et on obtient l'estimateur par la methode de variable de contrôle simple pour $b \in \mathbb{R}$:

$$\hat{\delta}(b) = \frac{1}{n}\sum_{i=1}^{n} \left(h(\mathbf{X}_i) - b(h_0(\mathbf{X}_i) - m)\right) \tag{18}$$

avec $\mathbf{X}_i = (X_{1,i}, X_{2,i}, X_{2,i}) \overset{i.i.d}{\sim} \mathcal{N}(\mu, \Sigma) \quad i = 1, ..., n$ et $m = \mathbb{E}(h_0(\mathbf{X})) = \frac{1}{3}\sum_{i=1}^{3} \mu_i$

b) D'après le cours, la valeur de b qui minimise la variance de l'estimateur est $b^* = \frac{cov(h_0(\mathbf{X}), h(\mathbf{X}))}{var(h_0(\mathbf{X}))}$. On peut estimer b^* à l'aide de la variance et de la covariance empirique :

$$\hat{b}^* = \frac{\sum_{i=1}^{l} \left(h_0(\tilde{\mathbf{X}}_i) - m\right)\left(h(\tilde{\mathbf{X}}_i) - \frac{1}{l}\sum_{j=1}^{l} h(\tilde{\mathbf{X}}_j)\right)}{\sum_{i=1}^{l} \left(h_0(\tilde{\mathbf{X}}_i) - m\right)^2} \tag{19}$$

Avec $\tilde{\mathbf{X}}_i = (\tilde{X}_{1,i}, \tilde{X}_{2,i}, \tilde{X}_{2,i}) \overset{i.i.d}{\sim} \mathcal{N}(\mu, \Sigma) \quad i = 1, ..., l$

```r
# Question 4 b)

m <- mean(mu)
# Estimation de b* avec la methode de burn-in
l <- 1000
x_tilde <- t(rmvnorm(l, mu, sigma))
h0x_tilde <- rowSums(x_tilde) / 3
hx_tilde <- rowSums(exp(-x_tilde)) / 3
hx_tilde[hx_tilde > 3] <- 3
b_star_hat <- cov(h0x_tilde, hx_tilde) / var(h0x_tilde)
delta_hat_b_star <- mean(hx - b_star_hat * (h0x1 - m))
mse_hat_b <- var(hx - b_star_hat * (h0x1 - m)) / n
r2 <- mse_hat_b / mse_barre
print(r2)
```

```
## [1] 0.02225058
```

Comme $\hat{\delta}_n(b)$ est un estimateur sans biais de δ, on a $\text{MSE}_{\hat{\delta}_n(b)} = var(\hat{\delta}_n(b)) = \frac{1}{n}var(h(\mathbf{X}) - b(h_0(\mathbf{X}) - m))$

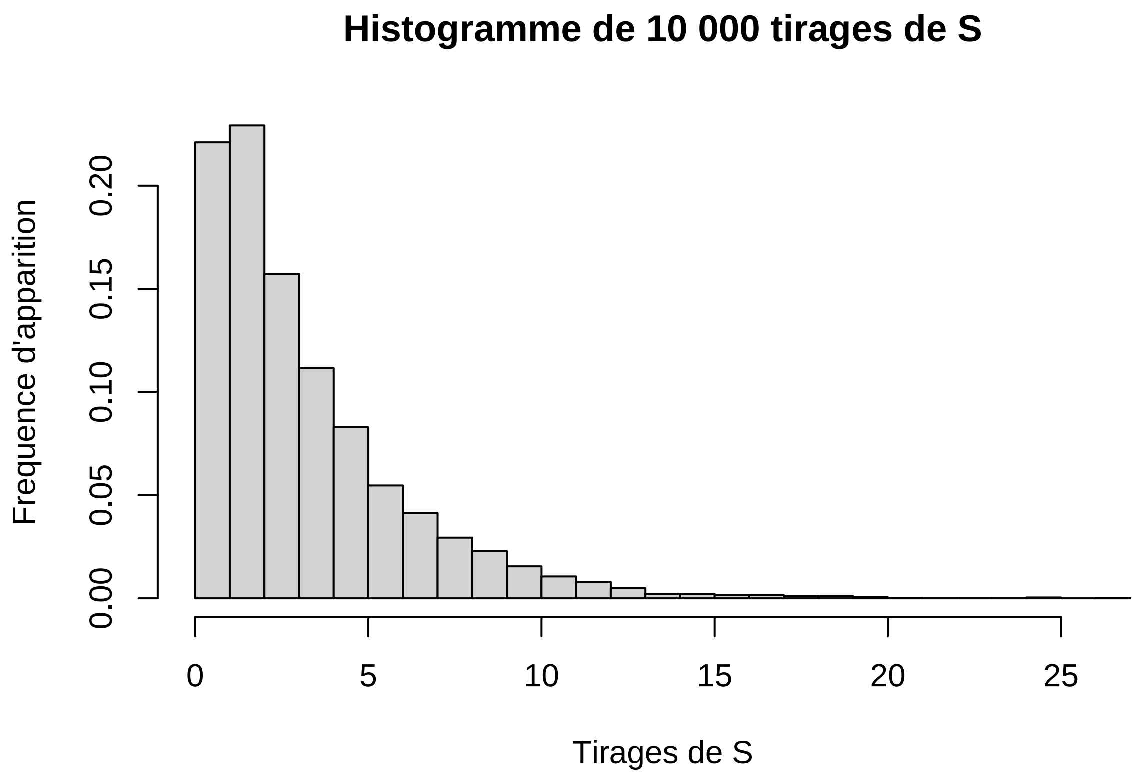
Ainsi, dans notre cas, la méthode de la variable de contrôle simple permet de réduire la variance de 98% par rapport à l'estimateur de Monte Carlo classique, au prix de simuler l variables en plus.

Exercice 3

1.

```r
# Question 1

# Fonction qui simule S pour un Y donne
simul_s <- function(y) {
  return(sum(log(rgamma(y, m, theta) + 1)))
}

# Variables utilisees dans la suite de l'exercice
n <- 10000
p <- 0.2
m <- 2
theta <- 2

c_mc <- Sys.time() # Temps de calcul pour question 2b)
geo <- rgeom(n, p) + 1 # Sur R, rgeom est a valeur dans N et non N*
tirage_s <- sapply(geo, simul_s) # echantillon de n tirages de S
c_mc <- Sys.time() - c_mc
delta_mc <- (sum(tirage_s)) / n

# estimateur sans biais --> mse(delta_mc)=var(delta_mc)
mse_mc <- var(tirage_s) / n

# Histogramme pour se donner une idee du tirage obtenu
bords <- seq(0, round(max(tirage_s)) + 1)
titre <- "Histogramme de 10 000 tirages de S"
axex <- "Tirages de S"
```

```r
axey <- "Frequence d'apparition"
hist(tirage_s, breaks = bords, freq = F, xlab = axex, ylab = axey, main = titre)
```



2. Nous allons appliquer la méthode de stratification à Y. Ainsi nous proposons l'ensemble de 15 strates suivant : $D_1 = \{1\}, D_2 = \{2\}, D_3 = \{3\}, ..., D_{13} = \{13\}, D_{14} = \{14\}, D_{15} = \mathbb{N} \setminus \{0, 1, ..., 14\}$.

D'après le cours, nous avons que :

$$Y^{(k)} = F^{\leftarrow}\left[F(d_{k-1}) + U\{F(d_k) - F(d_{k-1})\}\right]$$

suit la loi de Y|Y$\in D_k = \{d_k\}$ où Y suit une loi géométrique de paramètre p avec F sa fonction de répartition et U$\sim \mathcal{U}([0,1])$.

Calculons maintenant l'allocation proportionnelle $(n_1, ..., n_{15})$:

$$n_k = \mathbb{P}(Y \in D_k) \times n = \mathbb{P}(Y = k) \times n = p(1-p)^{k-1} \times n \text{ pour k } \in \{1, ..., 14\}$$

et $n_{15} = (n - \sum_{i=1}^{14} n_i)$

Ainsi, on obtient l'estimateur stratifié avec allocation proportionnelle suivant :

$$\hat{\delta}_n(n_1, ..., n_{15}) = \frac{1}{n} \sum_{k=1}^{15} \sum_{i=1}^{n_k} h(Y_i^{(k)})$$

où $h : y \mapsto \sum_{i=1}^{y} \log(X_i + 1)$

```r
# Question 2b)

c_strat <- Sys.time()
vect_proba <- dgeom(0:13, p) # Car la geometrique commence en 0 sur R
vect_proba[15] <- 1 - sum(vect_proba)
allocation <- round(n * vect_proba)
allocation[10] <- allocation[10] + 1 # Afin d'avoir "somme des allocations = n"
# Simulation de Y sachant qu'elle appartient a la derniere strate
unif <- runif(allocation[15])
y_15 <- qgeom((pgeom(13, p) + unif * (1 - pgeom(13, p))), p) + 1
# On evalue F en 13 = 14-1 car la geometrique commence a 0 sur R.
# On ajoute 1 pour la meme raison.
y_2 <- c(rep(1:14, allocation[1:14]), y_15)
tirage_s_2 <- sapply(y_2, simul_s)
c_strat <- Sys.time() - c_strat
delta_strat <- (sum(tirage_s_2)) / n
```

$$Var[\hat{\delta}_n(n_1, ..., n_{15})] = \sum_{k=1}^{15} \frac{p_k^2}{n_k}\sigma_k^2 = \frac{1}{n}\sigma^2(q_1, ..., q_{15}) = \frac{1}{n}\sigma^2(\frac{n_1}{n}, ..., \frac{n_{15}}{n})$$

Qui est finalement la variance de notre vecteur de probabilité puisque notre allocation est proportionnelle à la probabilité d'appartenance à chaque strate.

```r
# estimateur sans biais --> mse(delta_strat) = var(delta_strat)
mse_strat <- var(vect_proba) / n
# Efficacite relative de la methode
eff_relative <- (as.numeric(c_mc) * mse_mc) / (as.numeric(c_strat) * mse_strat)
```

Donc l'efficacité relative de la méthode de stratification avec allocation proportionnelle est 1300 fois meilleure que celle de la méthode de Monte-Carlo classique. Cela s'explique par le fait que l'estimateur stratifié avec allocation proportionnelle est de variance plus faible que l'estimateur de Monte Carlo classique et aussi car, grâce à la méthode de stratification, nous avons eu à générer 440 géométrique au lieu de 10 000.