Université de Montréal

IFT6561 – Simulation: aspects stochastiques

Devoir 2

par:

Eugénie Yockell
(20071932)
Cong Liu
(20161998)

Date de remise: 30 septembre 2020

Question 1

On cherche à mesurer le nombre de collisions dans un hypercube de trois dimensions divisé en 10^6 souscubes. Pour s'y faire, nous avons utilisé le générateur SWB avec les paramètres (b,r,k) = (231,8,48) qui est définie dans les exemples du cours. Pour définir les collisions, nous créons un tableau en 3 dimensions de longueur 100 où chaque indice correspond à un sous cube. Il suffit ensuite de multiplier par 100 le nombre aléatoire qu'on obtient et appliqué la fonction de plancher, ce nombre correspond maintenant à l'indice du sous cube. Nous faisons l'expérience pour $\mathbf{u}_i = (u_i, u_{i+1}, u_{i+2})$. Sur 10 réalisations indépendantes on obtient une moyenne de 53 collisions.

nombre de collisions
62
41
55
63
54
43
65
50
49
54

Table 1: Nombre de collisions sur 10 réalisations avec une moyenne de 53 collisions.

En faisant l'expérience pour un point aléatoire définit par $\mathbf{u}_i = (u_{25i}, u_{25i+20}, u_{25i+24})$, on obtient une moyenne de 2124 collisions.

nombre de collisions
2170
2138
2100
2104
2128
2086
2113
2117
2129
2158

Table 2: Nombre de collisions sur 10 réalisations avec une moyenne de 2124 collisions.

En faisant l'expérience pour un point aléatoire définit avec le générateur `MRG32k3a`, on obtient une moyenne de 49 collisions.

nombre de collisions
46
55
64
36
47
34
54
46
55
55

Table 3: Nombre de collisions sur 10 réalisations avec une moyenne de 49 collisions.

Le générateur SWB à une très longue période de 2^{1479}. Ce n'est toutefois pas suffisant pour qualifier un bon générateur. En effet, comme mentionné dans les notes de cours Exemple 1.16, les valeurs de u_n produites diffèrent de moins de $1/b = 2^{-31}$ que de ceux produits par le générateur LCG(linear congruential generator) avec modulo $m = b^k - b^r + 1$. En dimension 3, les points non nuls de la forme $(u_n, u_{n+k-r}, u_{n+k})$ générés par ce générateur LCG sont tous placés dans deux plans parallèles. Les deux plans sont définis par $u_n - u_{n+k+r} + u_{n+k} = q$, pour $q = 0$ et 1. On obtient un cas semblable lorsqu'on utilise un générateur SWB où les points sont générés par $\mathbf{u}_i = (u_{25i}, u_{25i+20}, u_{25i+24})$. Le manque de l'uniformité explique principalement pourquoi l'expérimentation générée par le générateur SWB ont un grand nombre de collisions.

D'après l'exemple 1.6, nous avons que si k est assez grand, alors l'espérance du nombre de collisions $C \approx \lambda_c = \frac{\lambda_m^2}{2k}$. La variable aléatoire C suit approximativement la loi de Poisson avec moyenne λ_c. Dans ce cas, nous avons $k = 10^6$, et $\lambda_m = 10^4$ identificateurs sont utilisés, alors théoriquement la variable aléatoire C aura une moyenne de $\lambda_c = \frac{\lambda_m^2}{2k} = \frac{10^8}{2*10^6} = 50$, ce qui correspond au résultat expérimental. Plus précisement, cette valeur est plus près du résultat obtenue avec le générateur `MRG32k3a` qui est le plus uniforme. On peut alors affirmer que `MRG32k3a` est un générateur plus efficace que le générateur SWB, car il génère des points plus uniforme et est plus près de la valeur théorique donné par la loi de Poisson.

Question 2

La fonction de densité de Gumbel est,

$$f(x) = \frac{1}{|\beta|} \exp\Big(- e^{(\delta-x)/\beta} + (\delta - x)/\beta\Big)$$

La CDF,

$$f(x) = \begin{cases} \exp\Big(- e^{(\delta-x)/\beta}\Big) & \text{if } \beta > 0 \\ 1 - \exp\Big(- e^{(\delta-x)/\beta}\Big) & \text{if } \beta < 0 \end{cases}$$

Définissons l'inverse de la CDF dans le cas $\beta > 0$,

$$\begin{aligned} &\exp\Big(- e^{(\delta-x)/\beta}\Big) = F(x) = U \\ &\Leftrightarrow e^{(\delta-x)/\beta} = -\ln(U) \\ &\Leftrightarrow (\delta - x)/\beta = \ln(-\ln(U)) \\ &\Leftrightarrow x = \delta - \beta \ln(-\ln(U)) \\ &\Leftrightarrow F^{-1}_{\beta>0}(x) = \delta - \beta \ln(-\ln(U)) \end{aligned}$$

Définissons l'inverse de la CDF dans le cas $\beta < 0$,

$$\begin{aligned} &1 - \exp\Big(- e^{(\delta-x)/\beta}\Big) = F(x) = U \\ &\Leftrightarrow \exp\Big(- e^{(\delta-x)/\beta}\Big) = 1 - U \\ &\Leftrightarrow e^{(\delta-x)/\beta} = -\ln(1 - U) \\ &\Leftrightarrow (\delta - x)/\beta = \ln(-\ln(1 - U)) \\ &\Leftrightarrow x = \delta - \beta \ln(-\ln(1 - U)) \\ &\Leftrightarrow F^{-1}_{\beta<0}(x) = \delta - \beta \ln(-\ln(1 - U)) \end{aligned}$$

On peut alors définir un algorithme pour générer une variable de Gumbel aléatoire.

procedure GENERATEGUMBEL(δ, β)
Genère variable aléatoire uniforme avec un random stream
$u = U(0, 1)$
  if $\beta > 0$ **then**
    $X = F^{-1}_{\beta>0}(u)$
  end if
  if $\beta < 0$ **then**
    $X = F^{-1}_{\beta<0}(u)$
  end if
end procedure X

Question 3

On veut estimer l'intégrale $\mu = \int_{[0,1)^s} f(\mathbf{u})d\mathbf{u}$ à l'aide de Monte Carlo (MC) et quasi-Monte Carlo randomiser (RQMC). Nous utilisons ici le programme `ProductExpCosRQMC` des exemples de cours pour mesurer les variances de Monte Carlo (MC) et quasi-Monte Carlo randomiser (RQMC). Un des avantages de RQMC est que c'est un estimateur sans biais avec une variance qui, sous certaines conditions, converge à taux plus rapide que les estimateurs Monte Carlo.

a) Pour une intégrale de dimension plus petite ou égal à 3, on s'attend qu'avec l'augmentation de b, la variance augmente pour RQMC. En effet, lorsque b augmente alors la fonction $f(\mathbf{u})$ oscille beaucoup plus. En effet, pour un nombre tel $c = bu$ où u est entre 0 et 1, un grand valeurs de b génère des nombres c très différents pour deux u différents. La fonction $cos(bu)$ va alors plus varier. Avoir une fonctions $f(\mathbf{u})$ qui varie beaucoup va augmenter la variance. Tandis que les plus grandes dimensions vont tendre à stabiliser la fonction $f(\mathbf{u})$ étant donné que c'est composé d'une somme de nombre entre -1 et 1. Ainsi, RQMC a plus de change de battre MC lorsque b est petit.

b) Les résultats sont présentés à la table 4 pour le cas $b = 0.5$. Dans le cas de $b = 0.5$ nous avons que la variance est 6.1×10^9 fois plus petite que pour la méthode MC. La méthode de RQMC diminue énormément la variance. La diminution de la variance permet d'augmenter la précision des estimations et ainsi rendre la méthode plus efficace.

	MC	**RQMC**
Variance	1.597	2.6×10^{-10}
Variance par exécution	1.597	1.68×10^{-5}

Table 4: Comparaison de variance sur l'estimation de l'intégral lorsque b=0.5

c) Les résultats sont présentés à table 5 pour le cas $b = 50$. Dans le cas de $b = 50$, la variance est 1.4×10^7 fois plus petite. On remarque alors que la variance à diminuer 100 fois plus avec $b = 0.5$. C'est exactement le résultat auquel on s'attendait selon l'hypothèse posé en **a)**. Plus b augmente, moins de chance à RQMC de battre MC. À noter que même dans le cas de $b = 50$, RQMC génère une variance nettement plus petite que MC.

	MC	**RQMC**
Variance	0.451	3.3×10^{-8}
Variance par exécution	0.451	2.15×10^{-3}

Table 5: Comparaison de variance sur l'estimation de l'intégral lorsque b=50

Question 4

On veut estimer $\mu = \mathbb{E}[X]$ à l'aide de la stratégie d'important sampling. On veut générer W_1 selon sa densité conditionnelle à $CW_1 \leq a \Rightarrow W_1 \leq a/C$. On veut aussi généré W_2 selon sa densité conditionnelle $CW_1W_2 \geq b \Rightarrow W_2 \geq b/W_1C$ tronqué à $[\frac{b}{CW_1}, \infty)$.
Définissons les nouvelles fonctions de densité,

$$g_1(w_1) = \frac{\pi_1(w_1)}{\mathbb{P}[W_1 \leq a/C]} = \frac{\pi_1(w_1)}{F_1(a/C)}$$

$$g_2(w_2|w_1) = \frac{\pi_2(w_2)}{\mathbb{P}[W_2 \geq \frac{b}{CW_1}]} = \frac{\pi_2(w_2)}{1 - F_2(\frac{b}{CW_1})}$$

Dérivons l'espérance,

$$\begin{aligned}
\mu &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X\pi_2(w2)\pi_1(w1)dw_1dw_2 \\
&= \int_{-\infty}^{a/C}\int_{b/Cw_1}^{\infty} X\frac{\pi_2(w2)\pi_1(w1)}{g_2(w_1|w_2)g_1(w_1)}g_2(w_1|w_2)g_1(w_1)dw_1dw_2 \\
&= \int_{-\infty}^{a/C}\int_{b/Cw_1}^{\infty} X\frac{\pi_2(w2)\pi_1(w1)}{\frac{\pi_1(w_1)}{F_1(a/C)}\frac{\pi_2(w_2)}{1-F_2(\frac{b}{CW_1})}}g_2(w_1|w_2)g_1(w_1)dw_1dw_2 \\
&= \int_{-\infty}^{a/C}\int_{b/Cw_1}^{\infty} XF_1(\frac{a}{C})\big(1 - F_2(\frac{b}{CW_1})\big)g_2(w_1|w_2)g_1(w_1)dw_1dw_2 \\
&= \mathbb{E}_g[X_is]
\end{aligned}$$

où $\mathbb{E}_g$ désigne l'espérance sous g_1 et g_2. On définit la variable de l'important sampling X_is comme,

$$X_is = XF_1(\frac{a}{C})\big(1 - F_2(\frac{b}{CW_1})\big)$$

Pour générer X, il faut calculer W_1 et W_2 selon les conditions définies. On définit alors la CDF tronquée inverse pour W_1 dans $[0, \frac{a}{C}]$,

$$\begin{aligned}
U &= F(w_1) = \frac{F(w_1) - F(0)}{F(a/C) - F(0)} \\
U &= \frac{F(w_1)}{F(\frac{a}{C})} \\
UF(\frac{a}{C}) &= F(w_1)
\end{aligned}$$

Pour la variable W_2 dans $[\frac{b}{CW_1}, \infty)$,

$$
\begin{aligned}
U = F(w_2) &= \frac{F(w_2) - F(\frac{b}{CW_1})}{F(\infty) - F(\frac{b}{CW_1})} \\
U &= \frac{F(w_2) - F(\frac{b}{CW_1})}{1 - F(\frac{b}{CW_1})} \\
U\big(1 - F(\frac{b}{CW_1})\big) &+ F(\frac{b}{CW_1}) = F(w_2)
\end{aligned}
$$

On peut alors simplement utiliser la fonction de CDF inverse lognormal sur les variables $UF(\frac{a}{C})$ et $U\big(1 - F(\frac{b}{CW_1})\big) + F(\frac{b}{CW_1})$.

Les résultats sont présentés à la table 6 pour $b = 102$ et à la table 7 pour $b = 112$. On remarque que dans le cas de $b = 102$ les espérances sont très semblables, mais la variance d'important sampling est 14 fois plus petite. Cette diminution dans la variance montre que la méthode d'important sampling est plus précise. Dans le cas de $b = 112$, les espérances sont encore une fois semblables, mais il y a un plus grand écart. La variance de l'important sampling est près de 2280 fois plus petite.

On peut alors conclure qu'avec une plus grande valeur de b, la méthode d'important sampling devient beaucoup plus précise. Ces résultats sont logiques puisque la méthode d'important sampling génère des distributions beaucoup moins dispersée dans le but de diminuer la variance. Dans notre cas, lorsque b est plus grand, nous avons un plus petit ensemble de points à générer ce qui diminue la variance, parce qu'on génère W_2 dans l'intervalle $[\frac{b}{CW_1}, \infty)$.

	Monte Carlo	**Important Sampling**
Espérance	0.20495	0.20860
Variance	0.86911	0.06215

Table 6: Comparaison de variance et de l'espérance lorsque b=102

	Monte Carlo	**Important Sampling**
Espérance	0.00793	0.00591
Variance	0.09122	0.00004

Table 7: Comparaison de variance et de l'espérance lorsque b=112