

Prof. Pierre L'Ecuyer

## DEVOIR 1

Devoir à remettre le *mercredi 16 septembre 2020, 13h30, avant le début du cours*, par courriel au professeur.

Les devoirs doivent être faits individuellement ou en équipe de deux. Il est très important de bien expliquer tout ce que vous faites et les preuves (quand il y en a) doivent être rigoureuses et convaincantes. Dans la correction, on accordera beaucoup d'importance à la clarté des explications. Attention au plagiat: il n'est pas permis de copier et/ou modifier les programmes ou les solutions venant de quelqu'un d'autre.

---

Les numéros des exercices réfèrent aux notes de cours. Les premiers sont des exercices simples de revision en probabilités. S.v.p. placez les exercices dans le bon ordre sur vos copies.

1. (5 points) Exercice 2.7
2. (5 points) Exercice 2.9
3. (6 points) Exercice 2.11
4. (7 points) Exercice 2.14
5. (7 points) (a) Prouvez que si  $X_1$  et  $X_2$  sont deux variables aléatoires prenant leurs valeurs dans  $[0, \infty)$ , alors  $\mathbb{E}[\max(X_1, X_2)] \geq \max(\mathbb{E}[X_1], \mathbb{E}[X_2])$  et  $\mathbb{E}[\min(X_1, X_2)] \leq \min(\mathbb{E}[X_1], \mathbb{E}[X_2])$ . Pensez à l'inégalité de Jensen.  
(b) Dans la Figure 1.4 des notes de cours, on voit qu'en remplaçant chaque  $Y_j$  par sa moyenne, on obtient une valeur plus petite que  $\mathbb{E}[T]$ , où  $T$  est la longueur du plus long chemin. Expliquer pourquoi cela est toujours vrai, peu importe le réseau. Et si  $T$  était la longueur du plus court chemin au lieu du plus long chemin, que se passerait-il? Pourquoi?
6. (8 points) Exercice 2.16: Let  $F$  be the cdf of some random variable  $Y$  and let  $-\infty < a < b < \infty$ .  
(a) Prove that if  $X = F^{-1}(U)$  where  $U \sim \text{Uniform}(F(a), F(b))$ , then  $X$  has the same distribution as  $Y$ , but truncated to the interval  $(a, b]$ , i.e.,  $X$  has the distribution of  $Y$  conditional on  $Y \in (a, b]$ . This provides an efficient way to generate  $X$  from this conditional distribution when  $F$  and  $F^{-1}$  are easy to evaluate. Hint: To show that, it suffice to show that the cdf of  $X$  is the same as the cdf of  $Y$  conditional on  $Y \in (a, b]$ .  
(b) If  $F$  is continuous, truncating to  $(a, b]$  is the same as truncating to the interval  $[a, b]$ , but not if  $F$  has a jump at  $a$ . Why? How would you generate  $X$  from distribution  $F$  truncated to  $[a, b]$  in the latter case?
7. (12 points) Voir l'exercice 1.19 des notes. Il n'y a pas de simulation à implanter pour cette question. L'idée est de comprendre ce qui se passe si on estime le volume d'une sphère de rayon 1 en  $s$  dimensions par la méthode Monte Carlo. Il s'agit bien sûr d'un exercice purement

académique, puisqu'on connaît déjà le volume de cette sphère, mais il permet de comprendre un type de difficulté qui survient dans de nombreuses applications pratiques. Pour estimer le volume, on tire  $n$  points au hasard dans le cube  $(0, 1)^s$ , on calcule la fraction  $\tilde{p}_n$  de ces points qui tombent dans la sphère (pour estimer la fraction  $p$  du cube occupé par la sphère), et l'estimateur du volume est  $\tilde{\mu}_n = 2^s \tilde{p}_n$ . L'*erreur relative*  $\text{RE}[\tilde{\mu}_n] = \sqrt{\text{Var}[\tilde{\mu}_n]}/\mathbb{E}[\tilde{\mu}_n]$ , représente l'erreur de cet estimateur (mesurée par l'écart-type) en proportion de la valeur que l'on veut estimer. Par exemple, une erreur relative de 0.10 veut dire que l'on a 10% d'erreur en moyenne quadratique.

(a) Prouvez que cet estimateur est sans biais. Donnez aussi (avec preuve) des formules exactes pour la variance et l'erreur relative de cet estimateur, en fonction de  $s$ .

(b) Pour avoir une erreur relative constante en fonction de  $s$ , disons inférieure à 0.01 pour tout  $s$ , à quelle vitesse (ou de quelle manière) doit-on augmenter  $n$  en fonction de  $s$ , lorsque  $s$  est grand? Donnez une formule pour  $n$  en fonction de  $s$  et expliquez ce que cela implique pour les grandes valeurs de  $s$ .

(c) Calculez les valeurs numériques de  $p$ ,  $V_s$ , et de  $n$  fois l'erreur relative au carré de  $\tilde{\mu}_n$ ,  $n\text{RE}^2[\tilde{\mu}_n] = n\text{Var}[\tilde{\mu}_n]/\mathbb{E}^2[\tilde{\mu}_n]$ , pour  $s = 2, 5, 10, 20$ .

Aide: Pour cette question, vous pouvez utiliser le fait que le volume d'une sphère de rayon 1 en  $s$  dimensions est  $V_s = \pi^{s/2}/\Gamma(1+s/2)$ , où  $\Gamma(\cdot)$  est la *fonction gamma*, définie par  $\Gamma(1/2) = \sqrt{\pi}$ ,  $\Gamma(1) = 1$ , et  $\Gamma(s+1) = s\Gamma(s)$ . Lorsque  $s$  est grand, l'approximation de Stirling donne

$$\Gamma(1 + s/2) \approx \lfloor s/2 \rfloor! \approx (\pi s)^{1/2} (s/(2e))^{s/2}.$$