

Prof. Pierre L'Ecuyer

## DEVOIR 6

Devoir à remettre le *mardi 25 février 2020, avant le début du cours.*

Les devoirs doivent être faits individuellement: un devoir par étudiant. Il est très important de bien expliquer tout ce que vous faites. 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 des solutions venant de quelqu'un d'autre, ou prises sur Internet ou dans des livres.

---

### 1. (13 points)

Soit  $F$  la cdf d'une variable aléatoire  $Y$  et soit  $-\infty \leq a < b \leq \infty$ .

(a) Prouvez que si  $X = F^{-1}(U)$  où  $U \sim \text{Uniforme}(F(a), F(b))$ , alors  $X$  a la même distribution que  $Y$ , mais tronquée à l'intervalle  $(a, b]$ , i.e.,  $X$  la loi de  $Y$  conditionnelle à  $Y \in (a, b]$ . Ce résultat donne une méthode efficace pour générer  $X$  selon cette loi conditionnelle lorsque  $F$  et  $F^{-1}$  sont faciles à évaluer.

Aide: Pour monter cela, il suffit de montrer que la cdf de  $X$  est la même que la cdf de  $Y$  conditionnelle à  $Y \in (a, b]$ .

(b) Si  $F$  est continue, tronquer à  $(a, b]$  est équivalent en pratique à tronquer à l'intervalle  $[a, b]$ , mais cela n'est pas vrai si  $F$  possède un saut au point  $a$ . Expliquez pourquoi.

### 2. (12 points)

Dans l'exemple de la page 41 des diapos sur la simulation, on considère une méthode de rejet pour une loi beta en utilisant une fonction chapeau  $h$  constante par morceaux (en bleu). Cette fonction chapeau est proportionnelle à une densité  $g$ , qui a une cdf correspondante  $G$ .

(a) Trouvez une formule pour  $G^{-1}$  et un algorithme pour générer  $X$  par inversion selon la densité  $g$ .

(b) Donnez un algorithme qui utilise cela pour générer  $X$  selon la densité  $f$  par la méthode de rejet.

### 3. (25 points)

Pour la question 3 du devoir 5, implantez la méthode de Gibbs avec balayage systématique pour le graphe donné en (c), pour  $k = 4$ . Faites  $n = 216,000$  itérations de la méthode et comptez le nombre de fois que chaque  $k$ -coloriage de  $\mathcal{C}$  est visité. Vous avez vu au devoir 5 qu'il y a 216 coloriations possibles avec  $k = 4$ , et si la chaîne est irréductible, sa loi stationnaire devrait être uniforme sur ces 216 coloriations, donc chacun devrait apparaître environ 1000 fois. Est-ce que vos résultats sont en accord avec cela?

Une façon simple de voir si les observations sont en accord avec cette hypothèse est de faire un test de chi-deux sur les résultats, pour tester l'hypothèse que la distribution est uniforme sur

l'ensemble des coloriage admissibles. Calculez le chi-deux, puis la  $p$ -valeur du test. Discutez vos résultats.

Rappel: S'il y a  $k$  possibilités, et si  $N_j$  est le nombre observé d'observations et  $e_j$  le nombre espéré d'observations pour la possibilité  $j$ , pour  $j = 1, \dots, 216$ , alors le chi-deux est défini par

$$\chi^2 = \sum_{j=1}^k \frac{(N_j - e_j)^2}{e_j}$$

et cette variable aléatoire suit approximativement la loi de chi-deux avec  $k - 1$  degrés de liberté. Si on observe  $\chi^2 = x$ , alors la  $p$ -valeur du test est  $p = \mathbb{P}[X > x]$  où  $X$  est une variable aléatoire qui suit la loi du chi-deux à  $k - 1$  degrés de liberté. Lorsque  $p$  est trop petit, on peut mettre en doute l'hypothèse. Dans ce cas-ci, on a  $k = 216$  et  $e_j = n/k = 1000$  pour tout  $j$ .

Formellement, le test du chi-deux est valide sous l'hypothèse que les  $n$  observations sont indépendantes. Ici, elles ne le sont pas, car ce sont les états visités par une chaîne de Markov, mais la loi du chi-deux est quand même une bonne approximation quand  $n$  est grand, et le test peut détecter si la loi dévie beaucoup de l'uniformité, par exemple.

Si vous voulez faire vos simulations en Java, il y a la librairie SSJ disponible ici: [ssj-github](#). Avant de l'utiliser, regardez bien la section "Documentation and tutorial". Ceci dit, vous pouvez utiliser le logiciel que vous voulez.