Dans cette série, on va s'intéresser à la simulation de variables aléatoires d'une loi donnée à partir de la distribution uniforme (la fonction *random* présente dans la plupart des logiciels).

Exercice 1. Méthode directe.

(a) Soient U une variable aléatoire uniformément distribuée sur l'intervalle $[0,1]$ et F une fonction de répartition (continue à droite). Soit $F^{-1}(u) = \inf\{x : F(x) \geq u\}$. Montrer que F est une fonction de répartition de la variable aléatoire $F^{-1}(U)$. En d'autres termes, si F est la fonction de répartition d'une variable aléatoire X, alors $F^{-1}(U) \stackrel{d}{=} X$.

Indication : montrer l'égalité $\{u \in [0,1] : \inf\{x : F(x) \geq u\} \leq v\} = [0, F(v)]$.

(b) Que peut-on dire de F^{-1} si F est continue et strictement croissante (i.e. F est inversible)?

(c) A quoi correspond F^{-1} pour une variable aléatoire à valeur dans $\mathbb{N}$? (Faire un dessin)

Exercice 2. Application de la méthode directe.
Générer, à l'aide de la méthode directe, 1000 réalisations $X_1, X_2, \ldots, X_{1000}$ d'une variable aléatoire exponentielle de paramètre $\lambda = 3$ (on utilisera `runif` pour générer des variables de loi uniforme) et comparer graphiquement l'histogramme de la simulation à la vraie densité. (Utiliser les fonctions `hist(...,breaks=30,freq=F)`,`curve(...,add=T)` et `dexp`).

Exercice 3. Méthode d'acceptation-rejet.
La méthode directe n'est pas applicable lorsqu'on ne connaît la densité de la variable à simuler qu'à une constante près, ou difficilement applicable lorsque la densité recherchée est très compliquée. Dans de tels cas, la méthode d'acceptation-rejet fournit une méthode de simulation exacte, mais plus coûteuse en temps de calcul.
Le principe est le suivant : on veut générer une variable aléatoire X de densité $f(x)$, et on sait facilement simuler une variable aléatoire Y de densité $g(y)$ avec la propriété suivante (c'est la seule restriction de la méthode) :

$$\text{il existe un } c \text{ tel que } \frac{f(y)}{g(y)} \leq c \text{ pour toutes les valeurs possibles de } y.$$

L'algorithme de simulation est le suivant :

PAS 1 : Simuler une v.a. Y de densité g et une v.a. U uniforme sur $[0,1]$.

PAS 2 : Si $U \leq \frac{f(Y)}{cg(Y)}$, donner à X la valeur Y, sinon retourner au PAS 1.

(a) Monter que $\mathrm{P}\left(U \leq \frac{f(Y)}{cg(Y)}\right) = \frac{1}{c}$.
Indication : Conditionner sur Y :

$$\mathrm{P}\left(U \leq \frac{f(Y)}{cg(Y)}\right) = \int \mathrm{P}\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y = y\right) g(y)dy.$$

(b) Montrer que le X défini dans l'algorithme satisfait bien $\mathrm{P}(X \leq t) = \int_{-\infty}^{t} f(x)dx$.

Indication : $\mathrm{P}(X \leq t) = \mathrm{P}\left(Y \leq t \mid U \leq \frac{f(Y)}{cg(Y)}\right) = \frac{\mathrm{P}\left(Y \leq t, U \leq \frac{f(Y)}{cg(Y)}\right)}{\mathrm{P}\left(U \leq \frac{f(Y)}{cg(Y)}\right)}$, continuer ce développement en conditionnant à nouveau sur Y.

Exercice 4. Application de la méthode d'acceptation-rejet.

(a) Utiliser la méthode d'acceptation-rejet pour générer 1000 réalisations d'une variable aléatoire $\mathcal{N}(0,1)$ en se basant sur la densité de Cauchy (`rcauchy`). Comparer graphiquement l'histogramme de la simulation à la vraie densité.

(b) À l'aide de la méthode de l'exercice 1, générer 1000 réalisations d'une variable aléatoire X de densité $f(x) = \frac{2}{5}(2 + \cos(x)) \exp(-x)$ en se basant sur la densité exponentielle (`rexp`). Comparer graphiquement l'histogramme de la simulation à la vraie densité.

$\boxed{\text{À rendre jusqu'au vendredi 15 mars 2013, à 12h}}$