

## IFT-3245

## Démonstration 4, 29 octobre 2012.

1. Introduction à TestU01 : fichier `standardminimal_testu01.c`.
2. Génération de variables non-uniforme par inversion : cas de la Gumbel et de la Géométrique. Expliquez comment générer
  - une variable aléatoire exponentielle
  - une variable aléatoire géométrique
 en utilisant le principe d'inversion. En utilisant la description d'une variable géométrique en termes d'expériences de Bernoulli, est-il possible de construire un autre générateur pour cette variable discrète ? Si oui, expliquez comment procéder, sinon justifiez votre réponse. Pouvez-vous donner d'avantage de commentaires sur cette dernière approche ?

Commençons par l'exponentielle. Dans ce cas,

$$F(x) = 1 - e^{-\lambda x} \quad \text{for } x > 0.$$

Il est facile d'inverser  $F$  :

$$X = -\ln(1 - U)/\lambda.$$

De manière équivalente, on peut générer  $X$  comme suit :

$$X = -\ln(U)/b.$$

Le cas de la géométrique est un peu plus difficile vu que le variable est discrete. On a

$$F(x) = P[X \leq k] = 1 - (1 - p)^{1+x}, \quad \text{for } x \in \mathcal{N}.$$

Par conséquent, on suivant le schéma vu au cours,

$$F^{-1}(u) = \inf \left\{ x \in \mathcal{N} \mid x \geq \frac{\ln(1 - u)}{\ln(1 - p)} - 1 \right\}$$

Ceci conduit à la méthode de génération suivante :

$$X = \left\lfloor \frac{\ln U}{\ln(1 - p)} \right\rfloor.$$

Une géométrique peut se formuler comme une répétition d'essais Bernoulli( $p$ ) jusqu'à l'obtention d'un premier succès. Dès lors, on pose  $X$  à 0, et on répète la procédure suivante. On génère  $U$  suivant une uniforme  $(0, 1]$ . Si  $U > 1$ , on ajoute 1 à  $X$ , sinon, on stoppe. Dans ce cas, le nombre de générations d'uniformes a comme espérance  $E[X] = (1 - p)/p$ . Si  $p$  est proche de 1, le coût peut être plus faible que la méthode d'inversion vu qu'on ne doit pas utiliser l'opérateur logarithmique. Pour  $p$  petit, la méthode d'inversion, en  $O(1)$ , sera plus rapide.

3. Introduction choix discrets.
4. L'hypothèse première en modèles de flot de trafic est que la vitesse est une fonction décroissante de la densité. Comme la densité croît, l'espace entre les véhicules diminue, et les conducteurs réagissent en ralentissant. Un des modèles les plus populaires est la formule de Greenshields :

$$u(K) = u_f - \frac{u_f}{K_j} K,$$

où  $u_f$  est la vitesse à flot libre (qui est observé pour une densité nulle) et  $K_j$  est la densité de congestion (aucun véhicule ne bouge plus). On s'en servira pour représenter des modèles simples de trafic.

5. Notion de fonction de répartition empirique. Retour sur la loi exponentielle, et principe de maximum de vraisemblance. On souhaite estimer le paramètre  $\lambda$  d'une loi exponentielle à partir d'un  $n$ -échantillon. La fonction de densité est

$$f(x, \lambda) = e^{-\lambda} \frac{\lambda^x}{x!}$$

La vraisemblance, dont nous cherchons le maximum, s'écrit :

$$L(x_1, \dots, x_i, \dots, x_n; \lambda) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$$

La vraisemblance étant positive, nous pouvons considérer son logarithme, qui est une fonction monotone croissante :

$$\begin{aligned} \ln L(x_1, \dots, x_i, \dots, x_n; \lambda) &= \ln e^{-\lambda n} + \ln \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} \\ &= -\lambda n + \left( \sum_{i=1}^n x_i \right) \ln \lambda - \sum_{i=1}^n \ln(x_i!). \end{aligned}$$

La dérivée première s'annule quand :

$$-n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0,$$

ou encore

$$\hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n}.$$

La dérivée seconde s'écrit :

$$\frac{\partial^2 \ln L(x_1, \dots, x_i, \dots, x_n; \lambda)}{\partial \lambda^2} = -\frac{\sum_{i=1}^n x_i}{\lambda^2} \leq 0.$$

Ce ratio étant toujours négatif alors, l'estimation est donnée par :

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}.$$

L'estimateur de maximum de vraisemblance correspond donc simplement au nombre d'occurrences par unité de temps.