

Simulation de la loi Gamma et du χ^2 décentré.

Responsable du projet: Aurélien Alfonsi, alfonsi@cermics.enpc.fr

La méthode du rejet.

On considère une suite $(X_i)_{i \geq 1}$ de variables aléatoires i.i.d. à valeurs dans \mathbb{R}^d . Pour $A \subset \mathbb{R}^d$ mesurable tel que $\mathbb{P}(X_1 \in A) = p > 0$, on pose

$$T = \inf\{i \in \mathbb{N}^*, X_i \in A\},$$

avec pour convention $\inf \emptyset = +\infty$.

1. Montrer que T suit une loi géométrique de paramètre p . En déduire que $\mathbb{P}(T < +\infty) = 1$.
2. Pour $B \subset \mathbb{R}^d$ mesurable et $n \in \mathbb{N}^*$, calculer $\mathbb{P}(X_T \in B, T = n)$.
3. En déduire que la loi de X_T est égale à la loi de X_1 sachant $X_1 \in A$, i.e.

$$\mathbb{P}(X_T \in B) = \mathbb{P}(X_1 \in A \cap B)/p.$$

Simulation de la loi Gamma

On rappelle que la loi $\Gamma(\alpha, \beta)$ pour $\alpha, \beta > 0$ est la loi de densité

$$\frac{1}{\Gamma(\alpha)} \beta^\alpha e^{-\beta x} x^{\alpha-1} \mathbf{1}_{\{x>0\}},$$

où $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} dz$, $\alpha > 0$ est la fonction Gamma d'Euler.

4. Montrer que $Z \sim \Gamma(\alpha, \beta)$ si, et seulement si $\beta Z \sim \Gamma(\alpha, 1)$. Expliquer alors pourquoi il est suffisant, pour simuler n'importe quelle variable aléatoire Gamma, d'être capable de simuler la loi $\Gamma(\alpha, \beta(\alpha))$ pour tout $\alpha > 0$ où $\beta : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ est une fonction quelconque.

Méthode de Cheng and Feast pour simuler $\Gamma(\alpha, \alpha - 1)$ lorsque $\alpha > 1$. On considère la méthode du rejet avec X_i ayant même loi que (U, V) où U, V sont indépendantes et de loi uniforme sur $[0, 1]$, et

$$A = \left\{ (x_1, x_2) \in (0, 1)^2, \frac{2}{\alpha - 1} \log(x_1) - \log\left(\frac{\lambda x_2}{x_1}\right) + \frac{\lambda x_2}{x_1} - 1 \leq 0 \right\},$$

avec $\lambda > 0$. On considère $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction test mesurable bornée.

5. Montrer que

$$\mathbb{E} \left[f \left(\frac{\lambda(X_T)_2}{(X_T)_1} \right) \right] = \frac{1}{p} \int_0^\infty f(v) \left(\int_0^1 \mathbf{1}_{\{x_1 \leq \exp(\frac{\alpha-1}{2}(1-v)) v^{\frac{\alpha-1}{2}}\}} \mathbf{1}_{\{x_1 < \frac{\lambda}{v}\}} \frac{x_1}{\lambda} dx_1 \right) dv$$

6. Calculer $\max_{v \geq 0} \exp(\frac{\alpha-1}{2}(1-v)) v^{\frac{\alpha-1}{2}}$ et $\max_{v \geq 0} \exp(\frac{\alpha-1}{2}(1-v)) v^{\frac{\alpha+1}{2}}$.

Montrer que pour $\lambda \geq \underline{\lambda} = e^{-1} \left(\frac{\alpha+1}{\alpha-1} \right)^{\frac{\alpha+1}{2}}$, $\frac{\lambda(X_T)_2}{(X_T)_1} \sim \Gamma(\alpha, \alpha-1)$.

7. Montrer que $\lambda^{CF} = \frac{\alpha-1/(6\alpha)}{\alpha-1} \geq \underline{\lambda}$. Implémenter la méthode de simulation en utilisant $\underline{\lambda}$ et λ^{CF} pour simuler la loi $\Gamma(5/2, 1)$. Calculer les probabilités de rejet par méthode de Monte-Carlo dans les deux cas. Tracer un histogramme réalisé avec 10000 simulations, et tracer sur le même graphique la densité de la loi $\Gamma(5/2, 1)$.

8. Soit $B = \left\{ (x_1, x_2) \in (0, 1)^2, \frac{2}{\alpha-1}(x_1-1) + \frac{x_1}{\lambda x_2} + \frac{\lambda x_2}{x_1} - 2 \leq 0 \right\}$. Montrer que $B \subset A$. Implémenter la méthode du rejet en testant d'abord si $X_i \in B$, et ensuite si nécessaire en testant $X_i \in A$. Donner en utilisant la méthode de Monte-Carlo un intervalle de confiance pour $\mathbb{P}(X_i \in B)$. Comparer le temps de calcul nécessaire à simuler 10^6 simulations avec cette méthode et avec la méthode qui teste directement si $X_i \in A$, pour $\lambda = \lambda^{CF}$.

Méthode d'Ahrens and Dieter pour simuler $\Gamma(\alpha, 1)$ lorsque $\alpha \in]0, 1]$. On considère à nouveau la méthode du rejet avec X_i ayant même loi que (U, V) où U, V sont indépendantes et de loi uniforme sur $[0, 1]$. On pose $\lambda = \frac{\alpha+e}{e}$ et considère

$$A = \left\{ (x_1, x_2) \in (0, 1)^2, x_2 \leq \mathbf{1}_{\{\lambda x_1 \leq 1\}} \exp(-(\lambda x_1)^{\frac{1}{\alpha}}) + \mathbf{1}_{\{\lambda x_1 > 1\}} (-\log(\lambda(1-x_1)/\alpha))^{\alpha-1} \right\}.$$

9. Montrer que $\mathbf{1}_{\{\lambda(X_T)_1 \leq 1\}} (\lambda(X_T)_1)^{\frac{1}{\alpha}} - \mathbf{1}_{\{\lambda(X_T)_1 > 1\}} \log(\lambda(1-(X_T)_1)/\alpha)$ suit la loi $\Gamma(\alpha, 1)$.

10. Implémenter cette méthode. En utilisant la méthode de Monte-Carlo, donner un intervalle de confiance pour la probabilité de rejet lorsque $\alpha = 1/2$. Tracer un histogramme réalisé avec 10000 simulations, et tracer sur le même graphique la densité de la loi $\Gamma(1/2, 1)$. Donner le temps de calcul nécessaire à simuler 10^6 simulations avec cette méthode et comparer aux temps obtenus à la question 8.

Simulation de la loi du χ^2 décentré

Pour $\nu > 0$ et $d \geq 0$, on note $\chi^2(\nu, d)$ la loi de densité

$$p_{\nu, d}(z) = \sum_{i=0}^{\infty} \frac{e^{-d/2} (d/2)^i}{i!} \frac{1/2}{\Gamma(i+\nu)} \left(\frac{z}{2} \right)^{i+\nu-1} e^{-z/2}, \quad z > 0.$$

Le paramètre ν est appelé le degré de liberté et d est appelé le paramètre de décentralité.

11. Vérifier que $\chi^2(\nu, 0)$ est la loi $\Gamma(\nu/2, 1/2)$, et que lorsque $\nu \in \mathbb{N}^*$, il s'agit de la loi de la somme de ν carrés de variables aléatoires gaussiennes centrées réduites indépendantes.

On suppose par la suite $d > 0$.

12. On considère N une loi de Poisson de paramètre $d/2$ et, conditionnellement à N , Y de loi $\Gamma(\nu + N, 1/2)$. Montrer que $Y \sim \chi^2(\nu, d)$. Calculer $\mathbb{E}[e^{-uY}]$ pour $u \geq 0$.
13. A l'aide de la question précédente, implémenter une méthode de simulation pour $\chi^2(\nu, d)$. Tracer, sur le même graphique et sur une grille régulière de valeurs de u , la valeur théorique de $\mathbb{E}[e^{-uY}]$ et la valeur obtenue par méthode de Monte-Carlo avec un nombre suffisant de simulations (on indiquera également l'intervalle de confiance à 95% sur ce graphique), pour $\nu = 1/2, 3/2$ et $d = 1$.

Amélioration du temps de calcul. On commence par supposer $\nu \geq 1/2$.

14. Soit $G \sim \mathcal{N}(0, 1)$ une gaussienne centrée réduite. Montrer que $(G + \sqrt{d})^2 \sim \chi^2(1/2, d)$. Pour $\nu > 1/2$, on considère $X \sim \Gamma(\nu - 1/2, 1/2)$ indépendante de G . Montrer que $X + (G + \sqrt{d})^2 \sim \chi^2(\nu, d)$.
15. Implémenter cette méthode de simulation, et comparer le temps de calcul avec la méthode de la question 13 pour simuler 100000 simulation avec $\nu = 7/3$ et $d = 5$.

On suppose $\nu \in]0, 1/2[$.

16. Montrer que $\partial_d[e^{d/2}p_{\nu,d}(z)] = \frac{1}{2}e^{d/2}p_{\nu+1,d}(z)$, puis que

$$p_{\nu,d}(z) = e^{-d/2}p_{\nu,0}(z) + \int_{e^{-d/2}}^1 p_{\nu+1,2\log(u)+d}(z)du.$$

17. A l'aide des questions 11, 14 et d'une variable aléatoire U de loi uniforme sur $[0, 1]$, proposer une méthode de simulation pour $\chi^2(\nu, d)$ alternative à celle de la question 13. Comparer les temps de calcul pour 100000 simulation avec $\nu = 1/6$ et $d = 5$.