

Travaux Pratiques

stoehr@ceremade.dauphine.fr

1 Simulations de variables aléatoires

Rappel R. `runif` permet de simuler des réalisations *i.i.d.* de loi $\mathcal{U}([0, 1])$. R propose des générateurs aléatoires pour la plupart des lois usuelles. Toutefois, ils ne seront pas utilisés dans cette partie sauf à titre de comparaison.

1.1 Fonction inverse et transformations

Exercice 1 (*Simulation d'une variable aléatoire discrète*). Soit X une variable aléatoire discrète sur l'ensemble $\{1, 2, 3, 4\}$. On définit la loi ν de X par

$$\mathbb{P}[X = 1] = 0.2, \quad \mathbb{P}[X = 2] = 0.5, \quad \mathbb{P}[X = 3] = 0.1, \quad \mathbb{P}[X = 4] = 0.2.$$

1. Simuler un échantillon \mathbf{u} de 10000 variables aléatoires *i.i.d.* suivant la loi $\mathcal{U}([0, 1])$.
2. À partir de \mathbf{u} , obtenir un échantillon \mathbf{x} de variables aléatoires *i.i.d.* suivant la loi ν .
3. Comparer le diagramme en bâtons de l'échantillon \mathbf{x} à celui de ν .

Rappel R. `barplot(height = ...)` trace le diagramme en bâtons de variables catégorielles ou discrètes. `height` est un vecteur contenant la hauteur des barres ou la table de contingence de l'échantillon. Pour un échantillon \mathbf{x} , la table de contingence s'obtient avec `table(x)`.

Exercice 2 (*Loi exponentielle et lois connexes*).

1. (a) Simuler 10000 réalisations de la loi exponentielle $\mathcal{E}(\lambda)$ de paramètre $\lambda = 2$ par la méthode de la fonction inverse.
(b) Tracer l'histogramme de cet échantillon et le comparer à la densité de la loi.
2. Soient X_1, \dots, X_n , n variables aléatoires *i.i.d.* de loi exponentielle $\mathcal{E}(\lambda)$.

(a) Montrer que $S_n = X_1 + \dots + X_n$ suit la loi gamma $\Gamma(n, \lambda)$, *i.e.*, a pour densité

$$f(x; n, \lambda) = \frac{\lambda^n}{\Gamma(n)} \exp(-\lambda x) x^{n-1} \mathbb{1}_{\{x \geq 0\}}, \quad \text{avec} \quad \Gamma(n) = \int_0^{+\infty} e^{-t} t^{n-1} dt.$$

- (b) Partant de ce résultat, simuler 10000 réalisations de la loi gamma $\Gamma(n, \lambda)$, avec $n = 10$.
 - (c) Tracer l'histogramme de cet échantillon et le comparer à la densité de la loi.
3. Soit $N = \sup\{n \geq 1 : S_n \leq 1\}$ (par convention $N = 0$ si $S_1 > 1$).

- (a) Montrer que N suit la loi de Poisson $\mathcal{P}(\lambda)$.
- (b) Partant de ce résultat, simuler 10000 réalisations de la loi de Poisson $\mathcal{P}(\lambda)$.
- (c) Tracer le diagramme en bâtons de cet échantillon et le comparer à la densité de la loi.

Rappel R.

- `hist(x, freq = F)` affiche l'histogramme d'un échantillon x . L'option `freq` spécifie si l'histogramme est représenté en densité d'effectifs (`freq = T` par défaut) ou en densité de probabilités (`freq = F`).
- `lines(x, y)` permet d'ajouter une courbe affine par morceaux reliant les points d'abscisse x et d'ordonnée y .
- Pour obtenir la densité aux points x , on a `dexp(x, rate = ...)` pour la loi exponentielle $\mathcal{E}(\text{rate})$, `dgamma(x, shape rate = ...)` pour la loi gamma $\Gamma(\text{shape}, \text{rate})$ et `dpois(x, lambda)` pour la loi de Poisson $\mathcal{P}(\text{lambda})$.

1.2 Loi normale, vecteurs gaussiens et mouvement brownien

Exercice 3 (Algorithme de Box-Muller).

1. Écrire une fonction `BM_car(n)` qui retourne n réalisations de la loi normale $\mathcal{N}(0, 1)$ par la méthode de Box-Muller cartésienne.
2. Valider l'algorithme à l'aide d'un outil graphique.

Rappel R.

- `dnorm`, `pnorm` et `qnorm` correspondent respectivement à la densité, à la fonction de répartition et au quantile d'une loi normale.
- `quantile(x, probs)` retourne les quantiles d'un échantillon x pour un vecteur de probabilités `probs`.
- `qqplot` permet de tracer un diagramme Quantile-Quantile.

Exercice 4 (Simulation de vecteurs gaussiens). Soit $\mathbf{X} = (X_1, X_2)$ de loi $\mathcal{N}(\mu, \Sigma)$, avec

$$\mu = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}$$

1. Simuler une suite de vecteurs $(X^{(n)})_{n \geq 1} = (X_{1,n}, X_{2,n})_{n \geq 1}$ qui suivent la loi de \mathbf{X} .
2. (a) Quelle est la loi de $X_1 + X_2$?
 (b) Comparer la densité de $X_1 + X_2$ avec l'histogramme de simulation.

Exercice 5 (*Simulation du brownien*). En utilisant que les accroissements d'un mouvement brownien $(W_t)_{0 \leq t \leq 1}$ sont indépendants et stationnaires simuler une réalisation du mouvement brownien.

1.3 Algorithme du rejet

Exercice 6 (*Rejet – un premier exemple*). Soit f une fonction de densité définie pour tout réel x par

$$f(x) = \frac{2}{\pi} \sqrt{1-x^2} \mathbb{1}_{\{x \in [-1,1]\}}.$$

1. Utiliser la méthode du rejet pour simuler 10000 réalisations suivant la loi de densité f .
2. Tracer l'histogramme de cet échantillon et le comparer à la densité f .

Exercice 7 (*Algorithme de Box-Muller – version polaire*).

1. Écrire une fonction `BM_pol(n)` qui retourne n réalisations de la loi normale $\mathcal{N}(0, 1)$ par la méthode de Box-Muller polaire.
2. Valider l'algorithme à l'aide d'un outil graphique.
3. Comparer les performances `BM_pol(n)` avec celles de `BM_car(n)`.

Exercice 8 (*Loi normale tronquée*). La loi normale tronquée de support $[b, +\infty[$ est définie par la densité f proportionnelle à

$$f_1(x) = \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \mathbb{1}_{\{x \geq b\}}, \quad \text{pour tout réel } x,$$

où μ et σ sont les paramètres de la loi normale $\mathcal{N}(\mu, \sigma^2)$.

Densité instrumentale n°1 : $\mathcal{N}(\mu, \sigma^2)$.

1. Écrire une fonction `tr_norm(n, b, mean, sd)` permettant de simuler n réalisations suivant la normale $\mathcal{N}(\text{mean}, \text{sd}^2)$ tronquée de support $[b, +\infty[$.
2. Simuler 10000 réalisation suivant la loi normale $\mathcal{N}(0, 1)$ tronquée de support $[2, +\infty[$ et en tracer l'histogramme.
3. Que se passe-t-il si l'on veut simuler suivant la loi normale tronquée de support $[50, +\infty[$?

Densité instrumentale n°2. La loi exponentielle translatée de b , $\tau\mathcal{E}(\lambda, b)$, a pour densité

$$g_\lambda(x) = \lambda e^{-\lambda(x-b)} \mathbb{1}_{\{x \geq b\}}, \quad x \in \mathbb{R}.$$

4. Écrire une fonction `tr_norm_2(n, b, mean, sd)` permettant de simuler n réalisations suivant la normale $\mathcal{N}(\text{mean}, \text{sd}^2)$ tronquée de support $[b, +\infty[$ avec la loi exponentielle translatée de b pour densité instrumentale (*c.f.*, TD n°2).
5. Comparer les performances de `tr_norm` et `tr_norm_2`.

2 Méthode de Monte Carlo classique

Exercice 9 (*Fonctions R*). On considère un estimateur de la forme

$$\hat{\delta}_n = \frac{1}{n} \sum_{k=1}^n y_k, \quad \text{avec } n \text{ la taille de l'échantillon et } (y_1, \dots, y_n) := y.$$

1. Écrire une fonction `mc_estim(y, level)` qui retourne $\hat{\delta}_n$ ainsi que l'intervalle de confiance au niveau `level` associé.
2. Écrire une fonction `mc_estim_evol(y, level)` qui retourne l'évolution de l'estimateur $\hat{\delta}_n$ de l'intervalle de confiance au niveau `level` associé en fonction de n .

Exercice 10 (*Estimation de π*).

1. Proposer une estimation de π par la méthode de Monte Carlo classique.
2. Tracer l'évolution de l'estimation de π et de l'intervalle de confiance au niveau de confiance 95% en fonction de la taille n de l'échantillon utilisé.
3. (a) En utilisant l'inégalité de Bienaymé-Tchebichev ou l'hypothèse du régime asymptotique, pour quelle valeur de n a-t-on une précision de 10^{-2} (au niveau de confiance 95%)?
(b) Comparer avec la valeur que vous obtenez.

Exercice 11 (*Mouvement brownien et finance*). Soit $(W_t^{(1)}), (W_t^{(2)})$ deux mouvements browniens avec corrélation $\rho = \frac{1}{2}$, i.e. $\mathbf{W} = (W_1^{(1)}, W_1^{(2)})$ suit la loi $\mathcal{N}(0, \Sigma)$, avec

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

On souhaite calculer

$$\mathcal{J} = \mathbb{E} \left[\left(\frac{1}{2} e^{-\sigma^2/2 + \sigma W_1^{(1)}} + \frac{1}{2} e^{-\sigma^2/2 + \sigma W_1^{(2)}} - 1 \right)^+ \right] \quad \text{avec } \sigma = 1.$$

1. Estimer \mathcal{J} par la méthode Monte-Carlo classique et donner l'intervalle de confiance au niveau 0.95.

3 Méthodes de réduction de variance

3.1 Échantillonnage préférentiel

Exercice 12 (*Loi de Cauchy*). Soit X une variable aléatoire de loi de Cauchy $\mathcal{C}(0, 1)$. On s'intéresse au calcul de

$$p := \mathbb{P}[X \geq 5] = \int_5^{+\infty} \frac{1}{\pi(1+x^2)} dx.$$

1. En utilisant R, donner la valeur de p .

Pour les différentes méthodes d'estimation ci-dessous, on tracera sur un même graphique l'évolution de l'estimateur et l'évolution de l'intervalle de confiance au niveau 0.95 associé en fonction du nombre de simulation. On prendra $n = 2000$.

2. En considérant x_1, \dots, x_n des réalisations de variables aléatoires *i.i.d.* suivant la loi de Cauchy $\mathcal{C}(0, 1)$, donner l'estimateur Monte Carlo classique de p .
3. Justifier que l'on peut écrire p sous la forme :

$$p = \int_0^{1/5} \frac{1}{\pi(1+x^2)} dx,$$

et en déduire une estimation de p par la méthode d'échantillonnage préférentiel.

4. Comparer la précision des différents estimateurs obtenus. Que pouvez-vous en conclure?

Exercice 13. Soit X une variable aléatoire de loi $\mathcal{N}(0, 1)$ et K une constante réelle. On s'intéresse au calcul de

$$p := \mathbb{P}[X \geq K].$$

Pour les applications numériques, on pourra prendre $K = 2$. Les intervalles de confiance seront donnés au niveau 0.95.

1. Donner la valeur de p en utilisant R.
2. Construire un estimateur de Monte Carlo classique de p .

Pour les questions suivantes, on tracera sur un même graphique l'évolution de la valeur de l'estimateur ainsi que de l'intervalle de confiance en fonction du nombre de simulations pour les valeurs du paramètre choisies.

3. Tester la méthode d'échantillonnage préférentiel pour la loi d'importance $\mathcal{N}(m, 1)$ avec différentes valeurs de m . Commenter les résultats obtenus.
4. Tester la méthode d'échantillonnage préférentiel avec pour loi d'importance la loi exponentielle $e(\lambda)$ translatée de K avec différentes valeurs de λ . Commenter les résultats obtenus.

3.2 Autres méthodes de réductions

Exercice 14 (*Mouvement brownien et finance – retour*). Soit $(W_t^{(1)}), (W_t^{(2)})$ deux mouvements browniens avec corrélation $\rho = \frac{1}{2}$. On souhaite calculer

$$\mathcal{J} = \mathbb{E} \left[\left(\frac{1}{2} e^{-\sigma^2/2 + \sigma W_1^{(1)}} + \frac{1}{2} e^{-\sigma^2/2 + \sigma W_1^{(2)}} - 1 \right)^+ \right].$$

Soient

$$\mathbf{Z} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \quad \text{et} \quad A = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

1. Montrer que le vecteur AZ suit la même loi que Z
2. Utiliser la méthode de la variable antithétique pour estimer \mathcal{J} .
3. Le nouvel estimateur est-il meilleur en terme de variance que l'estimateur classique calculé dans l'Exercice 11?

Exercice 15 (*Réduction de variance – un premier exemple*). On souhaite calculer

$$\mathcal{J} = \int_0^2 e^{-x^2} dx.$$

1. Utiliser R, pour calculer \mathcal{J} directement.
2. (a) Estimer \mathcal{J} par la méthode de Monte-Carlo classique de deux façons différentes.
(b) Donner les intervalles de confiance au niveau 0.95 correspondant aux estimateurs précédents.
3. (a) Pour chacun des estimateurs précédents, proposer une méthode alternative utilisant une variable antithétique.
(b) Donner les intervalles de confiance au niveau 0.95 correspondant aux estimateurs précédents.
4. (a) En écrivant un développement en série entière de l'intégrande, proposer une méthode d'estimation basée sur une variable de contrôle. On pourra tester la méthode pour $b = 1$ et pour la valeur optimale.
(b) Donner l'intervalle de confiance au niveau 0.95.
5. (a) Estimer \mathcal{J} par la méthode de stratification.
(b) Donner l'intervalle de confiance au niveau 0.95.
6. Comparer les différentes méthodes.

Exercice 16. On considère $\mathbf{X} = (X_1, X_2)$ une variable aléatoire de \mathbb{R}^2 admettant une densité f proportionnelle à

$$\tilde{f}(x_1, x_2) = \exp\left(-\frac{1}{2} \left\{ \frac{x_1^2}{4} + x_2^2 \right\}\right) \mathbb{1}_{\{|x_2| \leq 1\}},$$

On souhaite calculer

$$p = \mathbb{P}[e^{X_1} + e^{X_2} \geq K].$$

Pour les applications numériques, on pourra prendre $K = 5.3$. Pour chacun des estimateurs suivants, on donnera l'intervalle de confiance au niveau 95%.

1. Proposer une estimation de p par la méthode de Monte Carlo classique.

2. (a) Justifier que les variables aléatoires $(-X_1, X_2)$, $(X_1, -X_2)$ et $(-X_1, -X_2)$ sont toutes distribuées suivant la densité f_1 .
- (b) Parmi les transformations précédentes, lesquelles permettent de réduire la variance?
- (c) En déduire une estimation de p par la méthode des variables antithétiques.

3. On considère

$$h_{0,1}(X_1, X_2) = \mathbb{1}_{\{e^{X_1} \geq K_1\}} \quad \text{et} \quad h_{0,2}(X_1, X_2) = \mathbb{1}_{\{e^{X_2} \geq K_2\}}.$$

- (a) En utilisant une expérience Monte Carlo, déterminer laquelle de ces deux fonctions est la plus corrélée à $h(X_1, X_2) = \mathbb{1}_{\{e^{X_1} + e^{X_2} \geq K\}}$.
- (b) En déduire une estimation de p par la méthode de la variable de contrôle.
4. On partitionne la bande $\{(x, y) \in \mathbb{R}^2 : |y| \leq 1\}$ en strates D_ℓ , $\ell = 1, \dots, L$, définie par

$$D_\ell = [2\Phi^{-1}(\ell-1/L), 2\Phi^{-1}(\ell/L)] \times [-1, 1].$$

Proposer un estimateur stratifié de p pour l'allocation proportionnelle.