Wprowadzenie do symulacji i metod Monte Carlo 14.11.2018

Lista nr 3 ver 3

Paweł Lorek

$X \sim$ Gumbel z parametrami $\mu \in \mathbf{R}$ oraz $\sigma > 0$: dystrybuanta $F(x;\mu,\sigma)$ =exp$(-e^{-(x-\mu)/\sigma})$.

1. Ciągła wersja metody eliminacji. Cel: chcemy wygenerować zmienną losową X o gęstości $f(x)$ umiejąc generować Y o gęstości $g(y)$, gdzie $f(x) \leq cg(x)$ dla $c \geq 1$. Algorytm:

   1 Wygeneruj Y z gęstością $g(y)$

   2 Jeśli $cUg(Y) \leq f(Y)$ (gdzie $U \sim U(0,1)$ niezależna od Y) to podstaw $X := Y$

   3 w przeciwnym przypadku GOTO 1

   Podaj dowód poprawności algortymu (tzn., że wynikiem jest zmienna losowa X o gęstości $f(x)$).

2. Niech $U \sim U(0,1)$. Dla $p \in (0,1)$ zdefiniujmy
$$X = \left\lfloor \frac{\log U}{\log(1-p)} \right\rfloor.$$
   Jaki rozkład ma zmienna X?

3. Podaj (i wyjaśnij) całą procedurę generowania zmiennej losowej o rozkładzie Gamma(α,1) dla $\alpha > 0$ (skrypt: Przykład 6.4) metodą eliminacji (rozpatrz osobno przypadek $\alpha < 1$ i $\alpha \geq 1$).

4. Niech $U \sim U(0,1)$. Pokaż, że $X = \tan(\pi(U - 1/2))$ ma rozkład Cauchyego, tj. gęstość $f(x) = \dfrac{1}{\pi(1+x^2)}$.

5. Pokaż, iż dla $U \sim U(0,1)$ zmienna $X = -\log(-\log(U))$ ma rozkład Gumbela z parametrami $\mu = 0, \sigma = 1$. Jak z X uzyskać zmienną o rozkładzie Gumbela z dowolnymi parametrami $\mu, \sigma > 0$?

6. W algorytmie do generowania zmiennej o rozkladzie $N(0,1)$ metodą eliminacji używaliśmy zmiennej losowej Exp(1), tj. o gęstości $g(y) = e^{-y}\mathbf{1}(y \geq 0)$. Przedstaw wersję algorytmu, w którym używalibyśmy rozkladu Exp(λ), $\lambda > 0$, tj. o gęstości $g_2(y) = \lambda e^{-\lambda y}\mathbf{1}(y \geq 0)$. Dla jakiego λ prawdopodobieństwo akceptacji jest największe?

7. Podaj (i wyjaśnij) metodę eliminacji do symulowania zmiennej losowej z rozkładu $X \sim N(0,1)$ przy pomocy zmiennej losowej Y z rozkładu Cauchyego. Wskaż optymalne c.

   Czy można symulować metodą eliminacji zmienną losową z rozkładu Cauchyego za pomocą zmiennej losowej z rozkładu normalnego?

8. Rozpatrzmy następujący algorytm:

   1. Wysymuluj $U_1 \sim U(0,1)$ i podstaw $Y = \left\lfloor -\frac{\ln(U_1)}{\ln(2)} \right\rfloor$
   2. Wysymuluj $U_2 \sim U(0,1)$. Jeśli $U \leq \frac{2^{Y+1}}{4Y!}$ to zwróć $X := Y$
   3. w przeciwnym przypadku GOTO 1

   Jaki rozkład ma wynik działania algorytmu, tj. zmienna X?

9. Rozważmy szacowanie całki $I = \int_0^1 e^{-x^2} dx$.

   UWAGA: W poniższych wyliczeniach wariancji możesz przyjąć:

   $\int_0^1 e^{-x^2} dx \approx 0.7468, \qquad \int_0^1 \left(e^{-x^2}\right)^2 dx \approx 0.59814$

   Rozważmy dwa estymatory

   1. $\hat{Y}_n^{HoM} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, gdzie $Y_i = \mathbf{1}(e^{-U_{2i}} > U_{2i-1})$, a $U_1, \ldots U_{2n}$ są i.i.d. $U(0,1)$.
   2. $\hat{Y}_n^{CMC} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, gdzie $Y_i = e^{-U_i^2}$, a $U_1, \ldots U_n$ są i.i.d. $U(0,1)$.

   Jaki są wariancje powyższych estymatorów? Ile wynosi $\frac{Var\hat{Y}^{HoM}}{Var\hat{Y}^{CMC}}$? Ile musimy wykonwać symulacji (tzn. jakie musi być n) by mieć móc stwierdzić iż $Pr(I \in [\hat{Y}_n - b, \hat{Y}_n + b]) = 0.95$ dla $b = 0.005$ (odpowiedz dla $\hat{Y}_n$ będącego $\hat{Y}_n^{HoM}$ lub $\hat{Y}_n^{CMC}$).

10. Poprzednie zadanie było trochę "sztuczne" w tym sensie, że do wyliczenia wariancji użyliśmy wiedzy na temat tego co trzeba policzyć... Mamy następujące oszacowania:

    $$1 - x^2 + \frac{1}{2}x^4 - \frac{1}{6}x^6 \leq e^{-x^2} \leq 1 - x^2 + \frac{1}{2}x^4.$$

    Wykorzystaj je, by oszacować $Var\hat{Y}^{HoM}$ oraz $Var\hat{Y}^{HoM}$.

    Korzystając z tych oszacowań powiedz też ile musimy wykonwać symulacji (tzn. jakie musi być n) by mieć móc stwierdzić iż $Pr(I \in [\hat{Y}_n - b, \hat{Y}_n + b]) = 0.95$ dla $b = 0.005$ (odpowiedz dla $\hat{Y}_n$ będącego $\hat{Y}_n^{HoM}$ lub $\hat{Y}_n^{CMC}$).

    UWAGA: Nie musisz ręcznie liczyć całek (co do zasady prostych - to są całki z wielomianów), możesz wspomóc się tutaj komputerem.