

Symulacje Komputerowe - Raport

Kwiecień 2024

1 Generator ACORN (Additive Congruential Random Number)

Generator ACORN k -tego rzędu oparty jest na ciągach liczbowych zdefiniowanych poprzez równania rekurencyjne

$$\begin{aligned} X_n^0 &= X_{n-1}^0, \quad n \geq 1 \\ X_n^m &= (X_n^{m-1} + X_{n-1}^m) \bmod M, \quad m = 1, \dots, k, n \geq 1 \\ Y_n^k &= \frac{X_n^k}{M}, \quad n \geq 1. \end{aligned} \quad (1)$$

Jeśli ziarno (pierwszy element wektora o elementach $X_0^m \in \{0 \dots M\}$), czyli $X_0^0 \in \{1 \dots M-1\}$, jest względnie pierwsze z M oraz M jest dostatecznie dużą liczbą naturalną, to generator posiada pożądane cechy a ciąg Y_n^k dla $n \geq 1$ przypomina ciąg IID zmiennych losowych z rozkładu jednostajnego $\mathcal{U}(0, 1)$. Zaimplementuj generator jako funkcję $ACORN(N, k, M, Lag)$ zwracającą N liczb pseudolosowych Y_n^k , dla $n = 1, \dots, N$, korzystając z algorytmu k -tego rzędu, przy ominięciu pierwszych Lag wyrazów. Przykładowymi parametrami mogą być $M = 2^{89} - 1$ (liczba pierwsza Mersenne'a), $k = 9$, $X_0^m = 0$ dla $m \neq 0$ (mogą być dowolne), $Lag = 10^3$ (czy to na pewno OK?¹). Sprawdź poprawność wyniku robiąc wykresy ciągu Y_n^k , funkcji $Y_{n-1}^k \mapsto Y_n^k$ oraz histogramy. Porównaj wydajność z wbudowanym generatorem *numpy*. Na rysunku 1 zobaczyć możemy pomocniczą reprezentację zależności pomiędzy ciągami ciągów X_n^k . Szczegóły oraz odniesienia do literatury znaleźć możemy na stronie autorów algorytmu <https://acorn.wikramaratna.org/index.html>.

18

R. S. WIKRAMARATNA

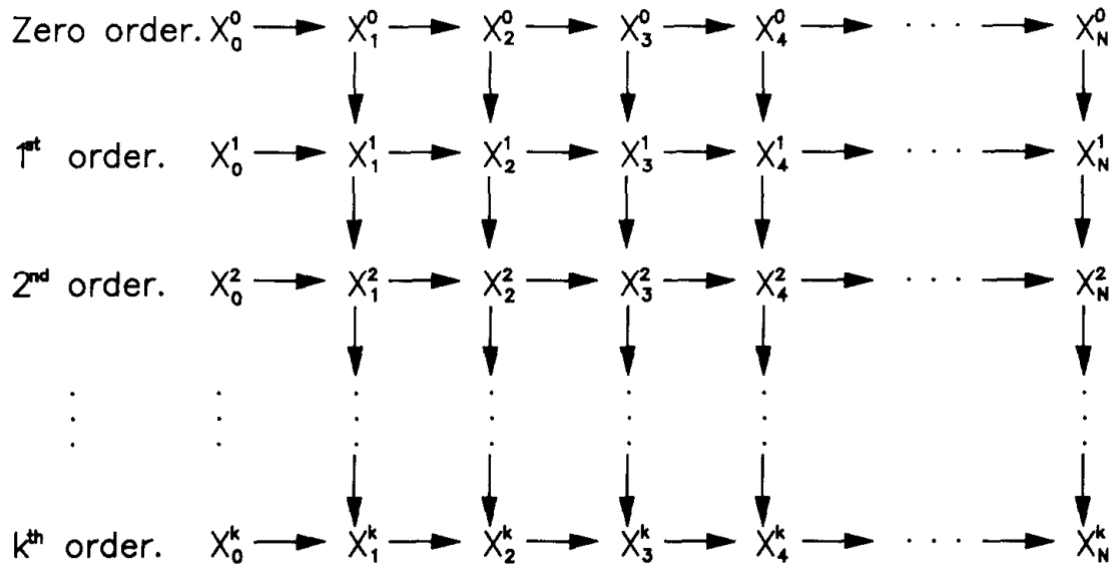


FIG. 1. Method of calculation of ACORN numbers.

Rysunek 1: Opis zależności pomiędzy wyrazami ciągów tworzącym generator ACORN. Obrazek z oryginalnego artykułu autora algorytmu ³. Ciąg k -ty to nieunormowany ciąg wyjściowy.

¹Julián M. Ortiz and Clayton V. Deutsch, *Random Number Generation with acorni: A Warning Note*

³R.S. Wikramaratna, *ACORN — A new method for generating sequences of uniformly distributed Pseudo-random Numbers*, Journal of Computational Physics, Volume 83, Issue 1, 1989, Pages 16-31, ISSN 0021-9991

2 Generowanie rozkładu normalnego

Na listach zadań poznaliśmy kilka metod generowania rozkładu normalnego - nie są to jedyne metody znane i powszechnie używane metody. W pierwszym kroku sprawdź dotychczas zaimplementowane algorytmy, czy działają poprawnie.

W drugim kroku, zaimplementuj dwie dodatkowe metody generowania rozkładu normalnego:

- metodę "tuzina przykładowa implementacja w linku poniżej
- metodę zigguratu
https://en.wikipedia.org/wiki/Ziggurat_algorithm
używając rozkładu wykładniczego. Przetestuj ile razy musimy obliczać funkcję wykładniczą dla różnych wyborów schodków.

Finalnie, porównaj efektywność generowania rozkładu normalnego za pomocą samodzielnie zaimplementowanych metod. Rozważ zarówno czas wykonania algorytmu, jak i dokładność symulacji.

Sprawdź też:

https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5982/7/56/1/ws_07_2019__02_porownanie_generatorow_liczb_pseudolosowych__piotr_sulewski.pdf

3 Metody redukcji wariancji

Zapoznaj się z metodami redukcji wariancji w metodach Monte Carlo: https://en.wikipedia.org/wiki/Variance_reduction W szczególności skupimy się na dwóch metodach:

- metoda odbić lustrzanych https://en.wikipedia.org/wiki/Antithetic_variates
- metoda zmiennej kontrolnej https://en.wikipedia.org/wiki/Control_variates

Krok 1.

Metodą Monte Carlo oblicz całkę

$$\int_0^1 \frac{4}{1+x^2} dx$$

(szacowanie liczby π)

Krok 2.

Zastosuj jedną z metod redukcji wariancji do ponownego obliczenia zadanej całki.

Krok 3.

Przedstaw analizę błędu względem ilości symulacji w formie tabeli lub wykresu.

4 Warunkowa wartość oczekiwana

Wartość oczekiwana $f(X)$ zmiennej Y warunkowanej zmienną X ma bardzo ciekawą interpretację, mianowicie spełnia własność

$$\mathbb{E}(Y|X = x) = f(x) = \arg \min_g \mathbb{E} \left((Y - g(X))^2 \right). \quad (2)$$

Oznacza to, że $\mathbb{E}(Y|X)$ to najlepsze przybliżenie w sensie L^2 zmiennej Y korzystające z danych pochodzących ze zmiennej X . Sprawdź symulacyjnie następujące stwierdzenia.

- Jeśli X, Y są zmiennymi niezależnymi a $\mathbb{E}(Y) = 0$, to dla $Z = XY + \sin X$ zachodzi $\mathbb{E}(Z|X) = \sin X$. Jako przykład możesz wziąć np. $X, Y \text{ iid } \sim \mathcal{N}(0, 1)$. **Zrób scatterplot $X \mapsto Z|X$ wysymulowanych wartości i zaznacz na nim wartości estymowane oraz teoretyczne $\mathbb{E}(Z|X)$.**
- Gdy N jest procesem Poissona o intensywności λ , to dla $T \geq t \geq 0$ zachodzi $\mathbb{E}(N_t|N_T) = tN_T/T$. Oznacza to, że np. zakładając przybywanie klientów do sklepu zgodnie z procesem Poissona i mając dane na temat dotychczasowej ilości klientów w sklepie w chwili T (czyli N_T), to najlepszym przybliżeniem dotychczasowej chwili klientów w chwili $t < T$ równej N_t jest tN_T/T . Jako przykład możesz wziąć np. $T = 1, \lambda = 10$ (zależnie od metody, niekoniecznie potrzebujesz parametru intensywności). **Zrób wykres $t \mapsto \mathbb{E}(N_t|N_T)$ dla kilku możliwych realizacji N_T i zaznacz na nim wartości estymowane.**
- Gdy N jest procesem Poissona o intensywności λ , to dla $t \geq s \geq 0$ zachodzi $\mathbb{E}(N_t|\mathcal{F}_s) = N_s + \lambda(t - s)$, gdzie \mathcal{F}_s to filtracja naturalna procesu N_s . Oznacza to, że np. zakładając przybywanie klientów do sklepu zgodnie z procesem Poissona i mając dane na temat dotychczasowej ilości klientów w sklepie w każdej chwili ω spełniającej $0 \leq \omega \leq s$ (czyli \mathcal{F}_s), to najlepszym przybliżeniem dotychczasowej chwili klientów w chwili $t \geq s$ jest $N_s + \lambda(t - s)$. Jako przykład możesz wziąć np. $s = 1, \lambda = 10$. **Zrób wykres $t \mapsto \mathbb{E}(N_t|\mathcal{F}_s)$ dla kilku możliwych realizacji $N_s, s \leq t$, i zaznacz na nim wartości estymowane. Podpowiedź: Wygeneruj kilka trajektorii $N_\omega, \omega \leq s$ i wybierz z nich te, które przybierają różne wartości N_s (w tym kroku nie potrzebowałeś użycia parametru intensywności). Następnie skorzystaj z własności Markowa i dla każdej z wybranych realizacji $N_\omega, \omega \leq s$ dosymuluj trajektorie N_ω dla $s < \omega \leq t$, za pomocą których wyestymuj $\mathbb{E}(N_t|\mathcal{F}_s)$.**

5 Czas ruiny

Niech X_t będzie procesem ruiny modelu Craméra-Lundberga

$$X_t = u + ct - \sum_{i=0}^{N_t} \xi_i, \quad (3)$$

gdzie $t, c, u \geq 0$, $\xi_i \sim \text{Exp}(\eta)$, $\xi_i \perp \xi_j$ dla $i \neq j$, $\mathbb{E}(\xi_i) = \eta$, a N_t jest procesem Poissona o intensywności λ . Czasem ruiny klasycznej nazywamy zmienną $\tau = \inf\{t > 0 | X_t < 0\}$. Prawdopodobieństwem ruiny w czasie nieskończonym nazywamy funkcję $\psi(u, c) = \mathbb{P}(\tau < \infty)$. Wzór Pollaczka-Chinczyna mówi, że

$$\psi(u, c) = \frac{\eta\lambda}{c} e^{-(\frac{1}{\eta} - \frac{\lambda}{c})u}. \quad (4)$$

Zweryfikuj symulacyjnie ten wynik. W tym celu skorzystaj z pomocniczego prawdopodobieństwa ruiny w czasie skończonym T , tj. $\Psi(u, c, T) = \mathbb{P}(\tau < T)$, $T > 0$, dla odpowiednio dużego T . [Sporządź wykresy funkcji \$u \mapsto \psi\(u, c_0\)\$ dla ustalonych \$c_0\$ oraz \$c \mapsto \psi\(u_0, c\)\$ dla ustalonych \$u_0\$ \(kilka trajektorii na jednym wykresie\) i porównaj je z wartościami wyestymowanymi.](#) Powtórz analizę dla funkcji odwrotnej

$$c(u, \psi) = \frac{\lambda u}{W_0\left(\frac{u\psi e^{\frac{u}{\eta}}}{\eta}\right)}, \quad (5)$$

gdzie W_0 to gałąź funkcji W Lamberta zdefiniowana poprzez równanie $W_0(xe^x) = x$ dla $x \geq 0$ (*scipy.special.lambertw*). Ta analiza odpowiada szukaniu wymaganej wartości wpłat przy danym kapitale początkowym w celu osiągnięcia wymaganego prawdopodobieństwa ruiny. Zweryfikuj symulacyjnie ten wynik. [Tym razem sporządź wykresy funkcji \$u \mapsto c\(u, \psi_0\)\$ dla ustalonych \$\psi_0\$ \(kilka trajektorii na jednym wykresie\) i porównaj je z wartościami wyestymowanymi.](#) Pamiętaj, że gdy $c \leq \lambda\eta$ to zachodzi $\psi = 1$!

6 Prawa arcusa sinusa

Ciągła zmienna losowa X ma rozkład arcusa sinusa ($X \sim \text{Arcsine}$), gdy jej funkcja gęstości prawdopodobieństwa p_X ma postać

$$p_X(x) = \frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}_{(0,1)}(x). \quad (6)$$

Dystrybuanta tej zmiennej jest wtedy równa

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \frac{2}{\pi} \arcsin(\sqrt{x}), & x \in [0, 1], \\ 1, & x > 1. \end{cases} \quad (7)$$

Prawa arcusa sinusa dla procesu Wienera W_t brzmią następująco.

Pierwsze prawo arcusa sinusa

$$T_+ = \lambda(\{t \in [0, 1] | W_t > 0\}) \sim \text{Arcsine}, \quad (8)$$

gdzie λ to miara Lebesgue'a. Oznacza to, że czas spędzony przez proces Wienera powyżej osi OX na odcinku $[0, 1]$ ma rozkład arcusa sinusa.

Drugie prawo arcusa sinusa

$$L = \sup\{t \in [0, 1] | W_t = 0\} \sim \text{Arcsine}. \quad (9)$$

Inaczej mówiąc, ostatni moment uderzenia procesu Wienera na odcinku $[0, 1]$ w oś OX ma rozkład arcusa sinusa.

Trzecie prawo arcusa sinusa

Niech M będzie liczbą spełniającą

$$W_M = \sup\{W_t | t \in [0, 1]\}. \quad (10)$$

Wtedy $M \sim \text{Arcsine}$. Oznacza to, że moment osiągnięcia maksymalnej wartości przez proces Wienera na odcinku $[0, 1]$ ma rozkład arcusa sinusa.

Zweryfikuj symulacyjnie prawa arcusa sinusa dla procesu Wienera. W tym celu użyj porównania histogramów oraz dystrybuant empirycznych wysymulowanych próbek z teoretycznymi wartościami.