Metody Probabilistyczne i Statystyka

LISTA 9

Lista zadań do wykładu 10

Informatyka algorytmiczna (I st.) WIT – 2021/2022

UWAGA – Zadania oznaczone symbolem ★ są nieobowiązkowe (zadania „jak starczy czasu", niekoniecznie trudne).

Zadanie 1. Jak wygenerować najłatwiej zmienną losową o gęstości prawdopodobieństwa $f(x) = 1,5 \cdot \sqrt{x}$ z wartościami na odcinku $[0,1]$?

Zadanie 2. Dlaczego metoda oparta na obliczeniu $F_X^{-1}(u)$ dla u z rozkładu jednostajnego generuje poprawnie zmienną losową o dystrybuancie F_X? Nie dowodziłem tego na wykładzie pozostawiając to Waszym intuicjom.

Zadanie 3. Pytanie dodatkowe do tej samej metody: obliczenie $F_X^{-1}(u)$ może być obarczone pewnym błędem. Same wartości u też są pobierane z pewną dokładnością (np. mają typ integer). Czy to może istotnie wpłynąć na zniekształcenie generatora? Jeśli tak, to w jakich sytuacjach?

Zadanie 4. Oszacuj złożoność obliczeniową przedstawionej na wykładzie metody generowania zmiennej losowej z rozkładem Poissona. Oczywiście, metoda ta nie ma stałego czasu obliczeń, więc oszacuj wartość oczekiwaną przyjmując jako parametr koszt każdej operacji arytmetycznej.

Zadanie 5. Zimplementuj "własny" generujący punkty w kwadracie $[0,1] \times [0,1]$ z rozkładem jednostajnym:

- znajdź w bibliotekach (np. Javy) funkcję haszującą, np. SHA-1,
- użyj outputu do wygenerowania losowych punktów kwadratu (czy wartość powiedzmy 160-bitowa może być wykorzystana do wygenerowania kilku punktów?)
- wygeneruj plot otrzymanych punktów i sprawdź jak to wygląda.

Najwygodniej byłoby to zrobić w MATLABIE, ale niestety chyba nie ma hasza wśród standardowo zaimplementowanych funkcji.

Zadanie 6. Transformacja Box-Muller dla zmiennych losowych U_1, U_2 ma postać

$$Z_1 = \sqrt{-2\ln(U_1)} \cdot \cos(2\pi U_2) \tag{1}$$

$$Z_2 = \sqrt{-2\ln(U_1)} \cdot \sin(2\pi U_2) \tag{2}$$

Jeśli U_1, U_2 są niezależne i mają rozkład jednostajny na $[0,1]$ to Z_1, Z_2 są niezależne mają rozkład normalny.

- zakładając, że powyższe stwierdzenie jest prawdziwe, porównaj złożoność generowania Z_1, Z_2 opisaną metodą do generowania np. za pomocą *rejection method*,
- ★ pokaż, że jest to prawda (wystarczy wyliczyć, że $\Pr(Z_1 < a \wedge Z_2 < b) = \Phi(a) \cdot \Phi(b)$)

- transformacja pozwala na jeszcze dalej idące sprytne uproszczenie: zamiast generować U_1, U_2 jak wyżej, można wziąć losowe u, v z odcinka $[0, 1]$ i zastosować *rejection sampling*: losować ponownie, gdy $(u, v) = (0, 0)$ lub $u^2 + v^2 > 1$. Wtedy można przejść do współrzędnych $U_1 = u^2 + v^2$ oraz U_2 jako kąt między osią x-ów a odcinkiem $((0, 0), (u, v))$.
  1. Rozrysuj tę sytuację.
  2. Sprawdź, że U_1, U_2 mają rozkład jednostajny.
  3. Zastąp wzory trygonometryczne $\cos(2\pi U_2)$ oraz $\sin(2\pi U_2)$ za pomocą wartości obliczonych bezpośrednio z u oraz v.
- Jaka jest ostatecznie złożoność generowania zmiennych losowych wg rozkładu normalnego tą metodą?