Egzamin: Symulacje Stochastyczne, II termin, ROZWIĄZANIA

MIM UW, Sierpień 2017.

1. Zmienna losowa S jest generowana przez następujący algorytm (zapisany w "pseudokodzie"):

   ```
   repeat
       Gen X ~ U(0,1);
       Gen Y ~ U(0,1);
       S := X + Y;
   until S < 1;
   return(S).
   ```

   (a) Podaj dystrybuantę zmiennej losowej S.

   (b) Podaj gęstość zmiennej losowej S.

   ROZWIĄZANIE: (X, Y) ma rozkład jednostajny na trójkącie $\{(x, y) : x > 0, y > 0, x + y < 1\}$, więc

   (a) $\mathbb{P}(S \leqslant s) = s^2$, dla $0 < s < 1$.

   (b) $f_S(s) = 2s$, dla $0 < s < 1$.

2. Dla podanych poniżej kodów w R, jaki jest rozkład zmiennych losowych Y?

   (a)
   ```
   X <- rexp(n)
   Y <- pexp(X)
   ```

   (b)
   ```
   X <- rnorm(n)
   Y <- pnorm(X)
   ```

   ROZWIĄZANIE: W obu przypadkach $Y = F_X(X)$, więc $Y \sim \mathrm{U}(0,1)$.

3. Dana jest funkcja `rboob(n)` zwracająca n niezależnych zmiennych losowych o skończonej wartości oczekiwanej μ i skończonej wariancji σ^2. Zarówno μ jak i σ są nieznane.

   Poniżej dany jest kod wywołany w interpreterze R, oraz rezultaty wywołania tego kodu:

```
> n <- 400
> X <- rboob(n)
> mean(X)
> [1] 50
> var(X)
> [1] 625
```

   (a) Podaj asymptotyczny przedział ufności dla μ o współczynniku ufności $1 - \alpha = 0.95$. (Kwantyl rzędu $1-\alpha/2 = 0.975$ rozkładu normalnego $\mathrm{N}(0,1)$ jest w przybliżeniu równy 2.)

   (b) Ile zmiennych należy wygenerować żeby obliczyć μ z dokładnością 0.001? (obliczenie z dokładnością ε oznacza, że jesteś w stanie podać przedział ufności o długości 2ε i współczynniku ufności $1 - \alpha = 0.95$.) *Uwaga:* Nie oczekujemy rygorystycznego wyprowadzenia najmniejszej możliwej liczby zmiennych które należy wygenerować, tylko rozsądnego oszacowania tej wartości wraz z uzasadnieniem.

<u>ROZWIĄZANIE:</u> Przedział ma postać $\mu = \bar{X} \pm \dfrac{z\hat{\sigma}}{\sqrt{n}}$, więc

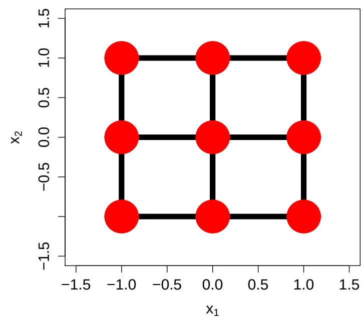
   (a) $50 \pm 2 * 25/20 = 50 \pm 2.5$.

   (b) $n \approx (10^3 * 2 * 25)^2 = 2.5 * 10^9$.

4. Przestrzenią stanów łańcucha Markowa jest $\mathcal{X} = \{-1, 0, 1\}^2$. Na przestrzeni $\mathcal{X}$ określona jest relacja „sąsiedztwa”:

$$(x_1, x_2) \sim (y_1, y_2) \text{ jeśli } |x_1 - y_1| + |x_2 - y_2| = 1.$$

Graf „sąsiedztwa” jest pokazany na rysunku.



Niech $N(x)$ oznacza liczbę sąsiadów punktu $x \in \mathcal{X}$. Prawdopodobieństwa przejścia łańcucha Markowa są dane wzorem:

$$Q(x, y) = \begin{cases} \dfrac{1}{N(x)} & \text{jeśli } x \sim y; \\ 0 & \text{w przeciwnym przypadku.} \end{cases}$$

(słowami: Q opisuje błądzenie po grafie, w każdym kroku przechodzimy do jednego z sąsiednich punktów z jednakowym prawdopodobieństwem).

(a) Znaleźć rozkład stacjonarny π dla łańcucha o macierzy Q.

Wskazówka: Zauważ, że $N(x)Q(x, y) = N(y)Q(y, x)$. Przypomnij sobie warunek odwracalności dla łańcuchów Markowa. Skorzystaj z symetrii grafu aby uniknąć wypisywania zbyt dużych wektorów i macierzy.

(b) Podaj prawdopodobieństwo przejścia w $n = 99$ krokach: $Q^{99}((0, 0), (0, 0))$.

ROZWIĄZANIE: $\pi(\cdot)$ jest unormowanym wektorem $N(\cdot)$, czyli $\pi(x) = N(x)/Z$, gdzie $Z = \sum_{x \in \mathcal{X}} N(x) = 1 * 4 + 4 * 3 + 4 * 2 = 24$.

(a) $\pi(x) = 4/24$ dla $x = (0, 0)$.

$\pi(x) = 3/24$ dla $x = (0, \pm 1)$ i $x = (\pm 1, 0)$.

$\pi(x) = 2/24$ dla $x = (\pm 1, \pm 1)$.

(b) Łańcuch ma okres 2, $Q^{99}((0, 0), (0, 0)) = 0$.

5. Zmienna losowa N jest generowana przez następujący algorytm (zapisany w "pseudokodzie"):

```
N := 0;
repeat
    Gen U ~ U(0,1);
    N := N + 1;
until U < p;
return(N).
```

(p jest danym parametrem, $0 < p < 1$).

(a) Podaj rozkład prawdopodobieństwa zmiennej losowej N (Oblicz $\mathbb{P}(N = n) = \cdots$).

(b) Zaproponuj lepszy algorytm generowania zmiennej losowej N o tym rozkładzie prawdopodobieństwa. Czas działania Twojego algorytmu nie powinien zależeć od wygenerowanej wartości N (zakładając, że koszt operacji arytmetycznych nie zależy od ich argumentów). *Wskazówka:* Łatwo generować zmienną X o rozkładzie wykładniczym. Może się przydać funkcja sufit '$\lceil\cdot\rceil$'.

ROZWIĄZANIE: $\mathbb{P}(N = n) = p(1-p)^{n-1}$ dla $n = 1, 2, \ldots$.

(a) $\mathbb{P}(N = n) = p(1-p)^{n-1}$ dla $n = 1, 2, \ldots$.

(b) Jeśli $X \sim \mathrm{Ex}(\lambda)$ to
$\mathbb{P}(\lceil X \rceil = n) = \mathbb{P}(n-1 < X \leqslant n) = e^{-\lambda(n-1)} - e^{-\lambda n} = (1 - e^{-\lambda})e^{-\lambda(n-1)}$, więc można przyjąć $N = \lceil X \rceil$ dla $X \sim \mathrm{Ex}(\lambda = -\log(1-p)$.