

Zadania do wykładu: Symulacje stochastyczne i teoria Monte Carlo

Tomasz Rolski

Instytut Matematyczny, Uniwersytet Wrocławski

Wrocław, 2015

1 Zadania

1

Zadania teoretyczne

- 1.1 Niech $X_1, X_2, \dots$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie $\mathcal{U}(\bar{M})$, gdzie M jest liczbą naturalną oraz $\bar{M} = \{0, 1, \dots, M-1\}$. Niech $M = 2^k$ oraz $b(x)$ dla $x \in \bar{M}$ będzie rozwinięciem dwójkowym przedstawionym w reprezentacji k cyfr zero lub jeden. Pokazać, że ciąg zmiennych losowych zero-jedynkowych zdefiniowany przez

$$\xi_1, \xi_2, \dots = b(X_1), b(X_2) \dots$$

jest ciągiem prób Bernoulliego. Pokazać na przykładzie, że tak nie musi być jeśli M nie jest potęgą dwójki.

- 1.2 Uzasadnić, że procedura generowania losowej permutacji z wykładu daje losową permutację po rozwiązaniu następującego zadania z rachunku prawdopodobieństwa, w którym indukcyjnie definiujemy ciąg n permutacji zbioru $[n]$. Zerową permutacją π_0 jest permutacja identycznościowa. Dla $k = 1, 2, \dots, n-1$ definiujemy indukcyjnie π_k -tą z π_{k-1} -szej następująco: zamieniamy element k -ty z J_k -tym gdzie J_k jest losowe (tj. o rozkładzie jednostajnym) na $\{k, \dots, n\}$, niezależnie od $J_1, \dots, J_{k-1}$. Pokazać, że π_{n-1} jest permutacją losową.

- 1.3 Dla przykładu **[??przykład:zad.o.kolizjach??]**, przy $n = 2^{14}$ i $k = 2^{20}$ uzupełnić tabelkę:

liczba kolizji $\leq$	100	110	120	130	140	150	160	170
z prawdopodobieństwem	?	?	?	?	?	?	?	?

- 1.4 Dany jest następujący ciąg zero-jedynkowy : 101011011111100001010100
0101100010100010101110110
100101010011010101011111
0111000101011000100000001
0011100111101001111001111
0001110110001011100111000

¹Poprawiane 13.10.2017

```

1011000001111001110001011
0100110110100101011010100
1111000010011011001000001
0001010001100100001100111
1111011110011001001001110

```

Przeprowadzić testy sprawdzające zgodność z hipotezą, że pochodzi z ciągu symetrycznych prób Bernoulliego.

1.5 Przez *funkcję błędu* $\operatorname{erf}(x)$ rozumie się

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

oraz przez *uzupełniającą funkcję błędu* $\operatorname{erfc}(x)$

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt.$$

Pokazać, że jeśli $\Phi(x)$ jest dystrybuantą standardowego rozkładu normalnego, to

$$\Phi(x) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}(x/\sqrt{2}) = \frac{1}{2}\operatorname{erfc}(-x/\sqrt{2}).$$

1.6 Udowodnić nierówność Czernowa. Niech $S_n = \xi_1 + \dots + \xi_n$, gdzie (ξ_i) są próbami Bernoulliego $\operatorname{Ber}(p)$. Pokazać, że dla $\delta < 1$

$$\begin{aligned} \mathbb{P}(S_n > (1 + \delta)np) &\leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^{np}, \\ \mathbb{P}(S_n < (1 - \delta)np) &\leq \left(\frac{e^{-\delta}}{(1 - \delta)^{1-\delta}} \right)^{np}. \end{aligned}$$

1.7 Obliczamy całkę $\bar{f} = \int_{[0,1]^d} f(\mathbf{u}) d\mathbf{u}$ przy założeniu, że $\sigma_f^2 = \operatorname{Var} f(\mathbf{U}) < \infty$. Korzystając z nierówności Czebyszewa pokazać, że z prawdopodobieństwem co najmniej δ mamy

$$\left| \frac{1}{n} \sum_{i=1}^n f(\mathbf{U}_i) - \bar{f} \right| < \frac{\sigma_f}{\sqrt{\delta n}}.$$

- 1.8 Rzucono parę kostek $n = 144$ razy i otrzymano następujące wyniki. W tabelce przez E_s oznaczamy liczbę oczekiwanych obserwacji a przez O_s liczbę zaobserwowanych.

s	2	3	4	5	6	7	8	9	10	11	12
E_s	4	8	12	16	20	24	20	16	12	8	4
O_s	2	4	10	12	22	29	21	15	14	9	6

Zweryfikować testem chi-kwadrat hipotezę, że rzuty są niezależne i na każdej kostce wyniki są jednakowo prawdopodobne. Obliczyć p -funkcje.

- 1.9 Pokazać, że dla układów pokerowych mamy następujące prawdopodobieństwa:

Losujemy 5 kart z 52 i liczymy średnią liczbę wystąpień określonych sekwencji. Liczba wszystkich możliwych sekwencji wynosi $W = \binom{52}{5} = 2\,598\,960$. Prawdopodobieństwa wystąpień sekwencji kart w pokerze: - Jedna para: $1098240/W = 0.422569 = 1/2,4$ - Dwie Pary: $123552/W = 0.047539 = 1/21$ - Trójka: $54912/W = 0.0211 = 1/44$ - Full: $3744/W = 0.00144 = 1/694$ - Karetą: $624/W = 0.00024 = 1/4165$ - Kolor: $5148/W = 0.00198 = 1/50$ - Strit: $9 * 4^5/W = 9216/W = 0.003546 = 1/282$ - Poker: $9 * 4/W = 36/W = 0.0000138 = 1/72193$

- 1.10 [Feller, t. 1, p.94] Klasyczne zadanie o rozmieszczeniu n kul w m komórkach zakłada ich "losowość", tzn. każde rozmieszczenie ma prawdopodobieństwo m^{-n} . Pokazać, że prawdopodobieństwo $p_c(n, m)$ znalezienia dokładnie c pustych wynosi

$$\binom{m}{c} \sum_{j=0}^{m-c} (-1)^j \binom{m-c}{j} \left(1 - \frac{n+j}{m}\right).$$

Pokazać, że gdy m, n rosną tak, że $\lambda = me^{-\frac{n}{m}}$ pozostaje ograniczone, to dla każdego c

$$p_c(n, m) \rightarrow \frac{\lambda^c}{c!} e^{-\lambda}.$$

Zadania laboratoryjne

Zadania na wykorzystanie instrukcji MATLAB-owskich: `rand`, `rand(m,n)`, `rand('state', 0)`, `randperm`, `randn`.

- 1.11 Przeprowadzić test szybkości generatora: `rand`, `rand('state')`, `rand('seed')`, `rand('twister')` przy użyciu następującego skryptu:

```
clc; clear;
tic
for j=1:1000000
    rand;
end
toc
```

- 1.12 Napisać program na sprawdzenie generatorów MATLABowskich przy użyciu
- testu Kołmogorowa–Smirnowa,
 - testu serii z $t = 3$. Obliczyć p -wartość dla $d = 4$, $n = 2^m$; ($m = 10, 11, \dots, 20$).

- 1.13 [Savicky] Uruchomić procedurę

```
Z=rand(28,100000);
condition = Z(1,:)<1/16;
scatter(Z(16,condition),Z(28,condition),'.'');
```

- 1.14 Uzasadnić, że w przypadku teoretycznym gdy Z jest macierzą składającą się z zmiennych losowych niezależnych o rozkładzie jednostajnym to na rysunku powinniśmy otrzymać chmurę par (U_{1i}, U_{2i}) ($i = 1, \dots, 100000$), gdzie U_{ij} są niezależnymi zmiennymi losowymi o jednakowym rozkładzie jednostajnym $U(0,1)$. Rysunek powinien więc wyglądać jak

```
Z=rand(2,100000);
condition = Z(1,:)<1/16;
scatter(Z(1,:),Z(2,:),'.');
```

Różnicę łatwo zauważyć. A co się dzieje gdyby użyć metody 'twister'. Wyjaśnić jak powinna wyglądać chmura w wypadku 'idealnym'.

- 1.15 Napisać procedurę na losowanie
- z n obiektów wybieramy losowo k ze zwracaniem,
 - z n obiektów wybieramy k bez zwracania.
 - wybierającą losowy podzbiór k elementowy $\{1, \dots, n\}$.

- 1.16 Napisać procedurę generowania losowej permutacji `randpermmy` na podstawie algorytmu z podrozdziału [??ssec1:alg.los.per?]. Porównać szybkość swojej procedury z matlabowska `randperm`.
- 1.17 Obliczyć przez symulacje prawdopodobieństwo p_n tego, że w permutacji losowej liczb $1, \dots, n$, żadna liczba nie jest na swoim miejscu. Zrobić obliczenia dla $n = 1, \dots, 10$. Do czego może zdążyć p_n gdy $n \rightarrow \infty$.
- 1.18 Napisać procedurę n rzutów monetą $(\xi_1, \dots, \xi_n)$.
Jeśli orzeł w i -tym rzucie to niech $\xi_i = 1$ jeśli reszka to $\xi_i = -1$. Niech $S_0 = 0$ oraz $S_k = \xi_1 + \dots + \xi_k$. Obliczyć pierwszy moment i wariancję ξ_1 . Zrobić wykres $S_0, S_1, \dots, S_{1000}$. Na wykresie zaznaczyć linie: $\pm \text{Var}(\xi)\sqrt{n}$, $\pm 2\text{Var}(\xi)\sqrt{n}$, $\pm 3\text{Var}(\xi)\sqrt{n}$.
Zrobić zbiorczy wykres 10 replikacji tego eksperymentu.
- 1.19 Rzucamy N razy symetryczną monetą i generujemy błędzenie przypadkowe $(S_n)_{0 \leq n \leq N}$. Napisać procedury do obliczenia:
1. Jak wygląda absolutna różnica pomiędzy maximum i minimum błędzenia przypadkowego gdy N rośnie?
- 1.20 Rozważamy ciąg $(S_n)_{0 \leq n \leq 2N}$. Przez prowadzenie w odcinku $(i, i+1)$ rozumiemy, że $S_i \geq 0$ oraz $S_{i+1} \geq 0$. Niech X będzie łączną długością prowadzeń dla N rzutów.
1. Jakie jest prawdopodobieństwo, że jeden gracz będzie przeważał pomiędzy 50% a 55% czasu? Lub więcej niż 95% czasu? Zrobić obliczenia symulacją. Eksperymentalnie sprawdzić ile powtórzeń należy zrobić aby osiągnąć stabilny wynik.
3. Niech $N = 200$. Zrobić histogram wyników dla 10000 powtórzeń tego eksperymentu.
- 1.21 Napisać procedurę do zadania dni urodzin.
1. Jakie jest prawdopodobieństwo, że wśród n losowo wybranych osób przynajmniej dwie mają ten sam dzień urodzin?
2. Jakie jest prawdopodobieństwo, że wśród n losowo wybranych osób przynajmniej dwie mają urodziny w przeciągu r dni jeden od drugiego?
3. Przypuśćmy, że osoby pojawiają się kolejno. Jak długo trzeba czekać aby pojawiły się dwie mające wspólne urodziny? Zrobić histogram dla 1000 powtórzeń.

- 1.22 [Test odstępów dni urodzin] Używając symulacji znaleźć funkcję prawdopodobieństwa statystyki R . Sprawdzić czy wyniki się pokrywają jeśli do obliczeń używa się różnych generatorów.
- 1.23 [Zagadnienie komiwojażera] Komiwojażer opuszcza miasto 0 i musi odwiedzić wszystkie miasta $1, \dots, n$ przed powrotem do punktu wyjściowego. Odległości pomiędzy miastem i a miastem j jest c_{ij} . Jeśli nie ma drogi do kładziemy $c_{ij} = \infty$. Jest $n = 100$ miast do odwiedzenia, których odległości są zadane w następujący sposób. Zacząć od `rand('state', 0)`. Przydzielić odległości kolejno wierszami c_{ij} gdzie $i = 0, \dots, 100$ oraz $j = 0, 1, \dots, 100$ - odległość między adresem i oraz j . (W ten sposób wszyscy będą mieli taką samą macierz odległości). Znaleźć jak najlepszą drogę $0, \pi_1, \dots, \pi_{100}, 0$ minimizującą przejechaną trasę. Czy uda się odnaleźć najkrótszą trasę. Znaleźć najkrótszą po 100, 200, itd. losowaniach.

Projekt

Projekt 1 Bob szacuje, że w życiu spotka 3 partnerki, i z jedną się ożeni. Zakłada, że jest w stanie je porównywać. Jednakże, będąc osobą honorową, (i) jeśli zdecyduje się umawiać z nową partnerką, nie może wrócić do już odrzuconej, (ii) w momencie decyzji o ślubie randki z pozostałymi są niemożliwe, (iii) musi się zdecydować na którąś z spotkanych partnerek. Bob przyjmuje, że można zrobić między partnerkami ranking: 1=dobra, 2=lepsz, 3=najlepsza, ale ten ranking nie jest mu wiadomy. Bob spotyka kolejne partnerki w losowym porządku.

Napisz raport jaką strategię powinien przyjąć Bob. W szczególności w raporcie powinny się znaleźć odpowiedzi na następujące pytania. Przy sporządzaniu raportu przeprowadzić analizę błędów, korzystając z fundamentalnego wzoru

$$b = \frac{1.96\sigma}{\sqrt{n}}$$

(patrz podsekcja [??fundamentalnyZwiązek??], gdzie b - błąd, n - liczba replikacji, oraz σ wariancja estymatora). A więc jeśli chcemy wyestymować r z błędem mniejszym niż $b = 0.02$ możemy skorzystać z fundamentalnego wzoru z wariancją uzyskaną wcześniej z pilotażowej symulacji 1000 repliacji.

- Bob decyduje się na ślub z pierwszą spotkaną partnerką. Oblicz prawdopodobieństwo P (teoretyczne), że ożeni się z najlepszą. Oblicz też oczekiwany rangę r
- Użyj symulacji aby określić P i r dla następującej strategii. Bob nigdy nie żeni się z pierwszą, żeni się z drugą jeśli jest lepsza od pierwszej, w przeciwnym razie żeni się z trzecią.
- Rozważyć zadanie z 10-cio ma partnerkami, które mają rangi 1, 2, ..., 10. Obliczyć teoretycznie prawdopodobieństwo P i oczekiwaną rangę r dla strategii jak w zadaniu (a).
- Zdefiniować zbiór strategii, które są adaptacją strategii z zadania (b) na 10 partnerek. Podać jeszcze inną strategię dla Boba. Dla każdej strategii obliczyć P i r . Wsk. Należy pamiętać, że jak się nie zdecydujemy na jakąś kandydatkę, to nie ma potem do niej powrotu.

Projekt 2 Rozpatrzyć następujące generatory liczb losowych: "state", "twister" oraz excellowski $u_i = (0.9821u_{i-1} + 0.211327) \bmod 1$. Ponadto rozpatrzyć pseudo-losowe ciągi bitów date.e, date.sqrt2 (z strony www.math.uni.wroc.pl/~tolsb)

- (i) Dla tych ciągów przeprowadzić analizę ich “losowości”, korzystając z przynajmniej 3 testów (odpowiednio dla każdego ciągu). Na przykład Kołmogorowa-Smirnow, test serii i test odstępów dni urodzin. Zamiast tego ostentego można skorzystać z testu kolizji.
- (ii) Dla generatora “twister” zbadać dwa przypadki. Pierwszy gdy generator startuje z ziarna $u_0 = 0$, a drugi gdy z ziarna $u_0 = 1812433253$.

2 Zadania

2

Zadania teoretyczne

- 2.1 Niech X ma dystrybuantę F i chcemy wygenerować zmienną losową o warunkowym rozkładzie $X|X \in (a, b)$ gdzie $\mathbb{P}(X \in (a, b)) > 0$. Niech

$$V = F(a) + (F(b) - F(a))U.$$

Jaki rozkład ma V . Pokazać, że $Y = F^{-1}(V)$ ma żadaną warunkową dystrybuantę G , gdzie

$$G(x) = \begin{cases} 0 & x < a, \\ \frac{F(x) - F(a)}{F(b) - F(a)}, & a \leq x < b, \\ 1 & x \geq b \end{cases}$$

- 2.2 Podać procedurę generowania liczb losowych o rozkładzie *trójkątnym* $\text{Tri}(a, b)$, z gęstością

$$f(x) = \begin{cases} c(x - a), & a < x < (a + b)/2, \\ c(b - x), & (a + b)/2 < x < b, \end{cases}$$

gdzie stała normująca $c = 4/(b - a)^2$. Wsk. Rozważyć najpierw gęstość $U_1 + U_2$.

- 2.3 Pokazać, że

$$X = \log(U/(1 - U))$$

ma rozkład *logistyczny* o dystrybuancie $F(x) = 1/(1 + e^{-x})$.

- 2.4 Podać dwie procedury na generowanie liczby losowej z gęstością postaci

$$f(x) = \sum_{j=1}^N a_j x^j, \quad 0 \leq x \leq 1,$$

gdzie $a_j > 0$. Wsk. Obliczyć gęstość $\max(U_1, \dots, U_n)$ i następnie użyć metody superpozycji. Jaki warunek muszą spełniać (a_i) ?

²Poprawiane 1.2.2016

2.5 Podać procedurę generowania zmiennej losowej X z gęstością $f(x) = n(1-x)^{n-1}$.

2.6 Podać dwie procedury generowania liczby losowej o rozkładzie Cauchego.

2.7 Niech F jest dystrybuantą rozkładu Weibulla $W(\alpha, c)$,

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \exp(-cx^\alpha), & x \geq 0. \end{cases}$$

Zakładamy $c > 0, \alpha > 0$. Pokazać, że

$$X = \left(\frac{-\log U}{c} \right)^{1/\alpha}$$

ma rozkład $W(\alpha, c)$.

2.8 Podać przykład pokazujący, że c w metodzie eliminacji (zarówno dla przypadku dyskretnego jak i ciągłego) nie musi istnieć skończone.

2.9 a) Udowodnić, że $\mathbb{P}(\lfloor U^{-1} \rfloor = i) = \frac{1}{i(i+1)}$, dla $i = 1, 2, \dots$

b) Podać algorytm i obliczyć prawdopodobieństwo akceptacji na wygenerowanie metodą eliminacji dyskretną liczbę losową X z funkcją prawdopodobieństwa $\{p_k, k = 1, 2, \dots\}$, gdzie

$$p_k = \frac{6}{\pi^2} \frac{1}{k^2}, \quad k = 1, 2, \dots$$

c) Podać algorytm i obliczyć prawdopodobieństwo akceptacji na wygenerowanie metodą eliminacji dyskretną liczbę losową X z funkcją prawdopodobieństwa $\{p_k, k = 1, 2, \dots\}$, gdzie $p_k = \frac{90}{\pi^4} \frac{1}{n^4}$.

2.10 Podaj metodę generowania liczby losowej z gęstością:

$$f(x) = \begin{cases} e^{2x}, & -\infty < x < 0 \\ e^{-2x}, & 0 \leq x < \infty \end{cases}$$

Taki rozkład nazywa się *podwójny wykładniczy*.

2.11 *Rozkład beta* $\text{Beta}(\alpha, \beta)$ ma gęstość

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < x < 1.$$

gdzie beta funkcja jest zdefiniowana przez

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx.$$

Przy założeniu $\alpha, \beta > 1$ podać algorytm wraz z usadnieniem generowania liczb losowych $\text{Beta}(\alpha, \beta)$ metodą eliminacji. Wsk. Przyjąć $g(x) = \mathbf{1}(0 < x < 1)$.

2.12 Gęstość rozkładu zadana jest wzorem

$$f(x) = Cxe^{-x^4}, \quad x \geq 0,$$

gdzie $C = 4/\sqrt{\pi}$. Podać procedurę generowania zmiennych losowych o rozkładzie f .

2.13 Gęstość dwuwymiarowego wektora losowego $\mathbf{Y} = (Y_1, Y_2)$ jest dana

$$f(x_1, x_2) = ce^{-|\mathbf{x}|^4} = ce^{-(x_1^2+x_2^2)^2}, \quad \mathbf{x} \in \mathbb{R}^2,$$

gdzie c jest znaną stałą normującą. Zaproponować algorytm na sumulację wektora losowego $\mathbf{Y}$. (Wsk. Przedyskutować dwie metody: eliminacji oraz przejście do współrzędnych biegunowych oraz zastanowić się którą lepiej używać).

2.14 a) Niech η będzie zmienną losową niezależną od X , $\mathbb{P}(\eta = -1) = \mathbb{P}(\eta = 1) = 1/2$, oraz X ma rozkład z gęstością daną w ([?eq:polnormalny?]). Pokazać, że ηX ma rozkład standardowy normalny $N(0,1)$.
b) Pokazać, że jeśli X ma rozkład $\text{Gamma}(\alpha, 1)$ to X/β ma rozkład $\text{Gamma}(\alpha, \beta)$.

2.15 Podać dwa algorytmy na generowanie punktu losowego o rozkładzie $U(B)$, gdzie $B = \{x \in \mathbb{R}^2 : |x| \leq 1\}$. Wsk. Do jednego z algorytmów wykorzystać lemat [?!:7.5?].

- 2.16 Niech (X, Y) będzie wektorem losowym z gęstością $f(x, y)$ i niech $B \subset \mathbb{R}^2$ będzie obszarem takim, że

$$\int_B f(x, y) dx dy < \infty.$$

Pokazać, że $X|(X, Y) \in B$ ma gęstość

$$\frac{\int_{y:(x,y) \in B} f(x, y) dy}{\int_B f(x, y) dx dy}.$$

- 2.17 [Marsaglia [?]] Niech J ma funkcję prawdopodobieństwa $(p_m)_{m=0}^\infty$ z $p_m = 1/(ce^{m+1})$ i $c = 1/(e-1)$. Niech I ma funkcję prawdopodobieństwa $(q_n)_{n=1}^\infty$, gdzie $q_n = c/n!$. Zakładamy, że I, J oraz $U_1, \dots$ są niezależne. Pokazać, że $X = J + \min(U_1, \dots, U_I)$ ma rozkład wykładniczy $\text{Exp}(1)$. Powyższa własność rozkładu wykładniczego daje możliwość napisania algorytmu na generowanie wykładniczych zmiennych losowych bez użycia kosztownej funkcji log. Zastanowić się jak można szybko generować liczby losowe I, J . Wsk. Zauważyć, że I ma rozkład Poissona $\text{Poi}(e)$ przesunięty o 1, natomiast J ma rozkład geometryczny $\text{Geo}(e^{-1})$. Wsk. Zauważyć, że $J \sim \text{Geo}(e^{-1})$ oraz $c_{k+1} = 1/(k+1)$ ³
- 2.18 Uzasadnić dwie metody generowania rozkładu wielomianowego $M(n, \mathbf{a})$. Pierwsza metoda jest uogólnieniem metody ad hoc dla rozkładu dwumianowego (patrz podrozdział [??sec:adhoc??]). Polega na generowaniu niezależnych wektorów losowych przyjmujących wartości $(0, \dots, 1, \dots, 0)$ (jedynek na i -tym miejscu) z prawdopodobieństwem a_i i następnie zsumowanie liczby jedynek na każdej koordynacie. Natomiast druga metoda wykorzystuje unikalną własność rozkładu wielomianowego wektora $\mathbf{X} = (X_1, \dots, X_d)$: rozkład X_{j+1} pod warunkiem $X_1 = k_1, \dots, X_j = k_j$ jest dwumianowy $B(n - k_1 - \dots - k_j, a_{j+1}/(a_1 + \dots + a_j))$ i X_1 ma rozkład $B(n, a_1)$.
- 2.19 a) Pokazać, że jeśli $Z_1, \dots, Z_n$ są niezależnymi zmiennymi o rozkładzie normalnym $N(0,1)$ i $\|Z\| = \sum_{j=1}^n Z_j^2$, to punkt $X = (Z_1/\|Z\|, \dots, Z_n/\|Z\|)$ ma rozkład jednostajny na sferze $S_{n-1} = \{x = (x_1, \dots, x_n) : \sum_{j=1}^n x_j^2 = 1\}$.

³A jak szybko generować liczby losowe I, J ?

b) Jeśli ponadto punkt ze sfery X z punktu a) przemnożymy przez $U^{1/n}$, gdzie U i X są niezależne, to otrzymamy punkt losowy z kuli $B_n = \{x : \sum_{j=1}^n x_j^2 \leq 1\}$.

Zadania laboratoryjne

2.20 Podać wraz z uzasadnieniem jak generować liczbę losową X o rozkładzie $U[a,b]$. Napisać MATLABowska funkcję $\text{Unif}(a,b)$ na generowanie liczby losowej X .

2.21 W literaturze aktuarialnej rozważa się *rozkład Makehama* przyszłego czasu życia T_x dla x -latka z ogonem rozkładu

$$\mathbb{P}(T_x > t) = e^{-At - m(c^{t+x} - c^x)}.$$

Napisać procedurę $\text{Makeham}(A,m,c,x)$ generowania liczb losowych o takim rozkładzie. Przyjąć $A = 5 \cdot 10^{-4}$, $m = 7.5858 \cdot 10^{-5}$ oraz $c = \log 1.09144$ (tzw. G82). Wsk. Zinterpretować probabilistycznie jaką operacja prowadzi do faktoryzacji $e^{-At - m(c^{t+x} - c^x)} = e^{-At} e^{-m(c^{t+x} - c^x)}$. Zastanowić się czy rozkład z ogonem dystrybucyjny $e^{-m(c^{t+x} - c^x)}$ można przedstawić jako warunkowy.

2.22 Napisać procedurę $\text{Poi}(\lambda)$ generowania liczb losowych o rozkładzie $\text{Poi}(\lambda)$. Uzasadnić jej poprawność.

2.23 Napisać funkcję $\text{Gamma}(a,b)$ na generowanie liczby losowej o rozkładzie $\text{Gamma}(a,b)$. Wsk. Algorytm musi się składać z dwóch części: $a < 1$ oraz $a \geq 1$.

2.24 Przeprowadzić następujący eksperyment z MATLABowskimi generatorami. Obliczyć objętość kuli $k = 30$ wymiarowej o promieniu 1 i porównać z wynikiem teoretycznym na objętość takiej kuli:

$$V_k = \frac{2}{k} \frac{\pi^{k/2}}{\Gamma(k/2)}.$$

W zadaniu należy przeprowadzić symulacje punktu z $(0,1)^k$ i następnie sprawdzić czy leży on w kuli. Zastanowić się ile należy wziąć replikacji.

- 2.25 Porównać dwie procedury generowania rozkładu normalnego $N(0,1)$ z podsekcji `ss.box-muller`. Porównać szybkość tych procedur mierząc czas wykonania 100 000 replikacji dla każdej z nich.
- 2.26 Korzystając z zadania 2.18 porównać dwie procedury generowania rozkładu wielomianowego. Do generowania rozkładu dwumianowego wykorzystać procedurę ITR. Przeprowadzić eksperyment z $n = 10$ i $a_1 = 0.55, a_2 = \dots = a_{10} = 0.05$. Porównać szybkość tych procedur mierząc wykonanie 100 000 replikacji dla każdej z nich.
- 2.27 Napisać procedurę generowania liczb losowych z gęstością

$$f(x) = pe^{-x} + \frac{(1-p)}{0.2} \left(\frac{1}{1+(x/0.2)} \right)^{1.2}.$$

Zrobić wykładniczy i paradowski wykres kwantylowy dla 300 zmiennych losowych wygenerowanych z $p = 0.05$ i $p = 0.95$.

Projekt

- 2.28 Napisać procedurę MATLAB-owską generowania liczby losowej $N(0,1)$ (tablicy liczb losowych $N(0,1)$)
- (i) metodą eliminacji (`randnrej`),
 - (ii) 1-szą metodą Boxa-Mullera (`randnbmfirst`),
 - (iii) 2-gą metodą Boxa-Mullera (`randnbmsec`). Przeprowadzić test z pomiarem czasu 100 000 replikacji dla każdej z tych metod oraz MATLAB-owskiej `randn`. Użyć instrukcji `TIC` i `TOC`. Każdy eksperyment zacząć od `rand('state',0)`. Sporządzić raport z eksperymentów. Zrobić wykresy kwantylowe dla każdej z procedur z 1000 replikacji.
- 2.29 Używając rozważania z ćwiczenia 2.17 napisać algorytm na generowanie liczb losowych o rozkładzie wykładniczym $\text{Exp}(1)$. Zastanowić się jak optymalnie generować liczby losowe I, J z zadania 2.17. Przeprowadzić test z pomiarem czasu 100 000 replikacji dla tej metody i metody ITM ($-\log U$). Każdy eksperyment zacząć od `rand('state',0)`. Sporządzić raport z eksperymentów.

Zadania teoretyczne

4

2.1 Zaproponować algorytm obliczenia całki

$$\int_{-\infty}^{\infty} \psi(x) f(x) dx$$

(zakładamy, że jest ona absolutnie zbieżna), gdzie

$$f(x) = \begin{cases} e^{2x}, & -\infty < x < 0 \\ e^{-2x}, & 0 < x < \infty. \end{cases}$$

2.2 Przypuśćmy, że chcemy obliczyć I metodą Monte Carlo, gdzie $I = I' I''$ oraz wiemy, że $I' = \mathbb{E} Y'$, $I'' = \mathbb{E} Y''$. Niech $Y'_1, \dots, Y'_R$ będzie R niezależnych replikacji Y' oraz niezależnie od nich niech $Y''_1, \dots, Y''_R$ będzie R niezależnych replikacji Y'' . Pokazać, że następujące estymatory są nieobciążone dla I :

$$\hat{Z}_1 = \left(\frac{1}{R} \sum_{i=1}^R Y'_i \right) \left(\frac{1}{R} \sum_{i=1}^R Y''_i \right)$$

$$\hat{Z}_2 = \frac{1}{R} \sum_{i=1}^R Y'_i Y''_i.$$

Pokazać, że $\hat{Z}_1$ ma mniejszą wariancję. Wsk. Pokazać, że dla niezależnych zmiennych losowych X, Y mających drugie momenty mamy

$$\text{Var}(XY) = \text{Var} X \text{Var} Y + (\mathbb{E} X)^2 \text{Var} Y + (\mathbb{E} Y)^2 \text{Var} X.$$

2.3 Zanalizować przykład `ex.ZlyGeneratorState` pod kątem odpowiedniości liczby replikacji. Jak dobrać δ aby poza $k = 27$ na wykresie `fig:przykladZlyGenState.eps` trudno byłoby rozróżnić krzywą wysumulowaną od teoretycznej. Czy lepiej byłoby przyjąć $\delta = 0.05$.

- 2.4 [Igła Buffona a liczba π] W zadaniu Buffona o liczbie π oblicza się prawdopodobieństwo

$$p = \frac{2L}{\pi}$$

przecięcia jednej z równoległych prostych oddalonych od siebie o 1 przez igłę o długości L . Wiadomo, że przeprowadzając niezależne eksperymenty mamy nieobciążony estymator $\hat{p}$ prawdopodobieństwa p . Wtedy $\pi = (2L)/p$. Zaplanować obliczenie $f(I)$ z dokładnością $b = 0.01$ na poziomie $\alpha = 0.05$, gdy $f(x) = (2L)/x$ oraz $I = \mathbb{E} Y$, gdzie $Y = 1$, gdy igła przecina linię, natomiast $Y = 0$ w przeciwnym razie.

- 2.5 Ile należy zrobić replikacji na poziomie ufności $\alpha = 0.05$ aby błąd $b = \sigma/50$.

- 2.6 Wiadomo, że jeśli $Y_1, \dots, Y_n$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie $N(0,1)$ to

$$\frac{(n-1)\hat{S}^2/\sigma^2 - n + 1}{\sqrt{2n-2}} \xrightarrow{\mathcal{D}} N(0,1), \quad (2.1)$$

gdy $n \rightarrow \infty$; patrz [?], str. 27. Natomiast jeśli zmienne są o rozkładzie dowolnym, to

$$\sqrt{n} \left(\hat{S}_n^2 - \sigma^2 \right) \xrightarrow{\mathcal{D}} N(0, \mu_4 - \sigma^4), \quad (2.2)$$

gdzie $\mu_4 = \mathbb{E} (Y - \mathbb{E} Y)^4$ jest momentem centralnym rzędu 4. Ile należy zrobić prób aby wyestymować σ^2 z dokładnością do 5-ciu % na poziomie $\alpha = 0.05$. Rozważyc oba warianty przy użyciu (2.1), (2.2).

Zadania laboratoryjne

- 2.7 a. Napisać algorytm na wygenerowanie punktu losowego o rozkładzie jednostajnym $U(\Delta)$ w trójkącie Δ o wierzchołkach $(1, 1)$, $(-1, -1)$, $(1, -1)$. Policzyc średnią liczbę liczb losowych $U_1, U_2, \dots$ potrzebnych do wygenerowania jednej liczby V o rozkładzie $U(\Delta)$.
b. Napisać algorytm na obliczenia całki

$$\int_{\Delta} e^{-(x^2+y^2)} dx dy$$

zgrubną metodą MC.

- 2.8 Wyprodukować 10 000 przedziałów ufności postaci $(\hat{Y}_n \pm \frac{z_{1-\alpha/2}\hat{S}_n}{\sqrt{n}})$ dla $Y = \mathbf{1}(U \leq u)$, $n = 1000$ oraz $I = \mathbb{E}Y$ następnie sprawdzić jaki procent tych przedziałów pokryje $I = 0.5, 0.9, 0.95$ gdy odpowiednio $u = 0.5, 0.9, 0.95$. Wyjaśnić otrzymane wyniki.

Projekt

- 2.9 Obliczyć składkę netto z dokładnością do czterech miejsc po przecinku, dla 40-latka w ubezpieczeniu na całe życie, gdy techniczna stopa procentowa wynosi $r = 0.05$, przyszły czas życia $T = T_{40}$ ma rozkład Makehama (można użyć funkcji Makeham z zadania 2.21). Wzór na składkę netto gdy suma ubezpieczenia wynosi 1 zł jest

$$\pi = \frac{\mathbb{E} e^{-rT}}{\mathbb{E} \int_0^T e^{-rt} dt} .$$

Zadania teoretyczne

5

2.1 Niech $F \sim \text{Exp}(\lambda)$ i $Z_1 = F^{-1}(U)$, $Z_2 = F^{-1}(1 - U)$. Pokazać, że

$$\mathbb{E} Z_i = 1/\lambda, \quad \mathbb{E} Z_i^2 = 2/\lambda^2,$$

oraz

$$\text{corr}(Z_1, Z_2) = -0.6449.$$

Wsk. $\int_0^1 \log x \log(1 - x) dx = 0.3551.$

2.2 Obliczyć wartość oczekiwaną i wariancję rozkładu Pareto $\text{Par}(\alpha)$.

2.3 Niech $Z_1 = F^{-1}(U)$, $Z_2 = F^{-1}(1 - U)$. Oszacować korelację $\text{corr}(Z_1, Z_2)$ gdy F jest dystrybuantą:

- (a) rozkładu Poissona; $\lambda = 1, 2, \dots, 10$,
- (b) rozkładu geometrycznego $\text{Geo}(p)$.

2.4 a) Napisać algorytm na generowanie zmiennej losowej Y z gęstością

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1, \\ 0 & \text{w pozostałych przypadkach.} \end{cases}$$

Uzasadnić!

b. Przypuśćmy, że chcemy symulować $I = \mathbb{E} Y$, gdzie Y ma rozkład jak w części a. Ile należy zaplanować replikacji aby otrzymać I z dokładnością do jednego miejsca po przecinku, gdy

1. stosujemy zgrubny estymator Monte Carlo – CMC,
2. stosujemy metodę zmiennych antytetycznych.

Wsk: $\text{corr}(\text{rand}^{1/2}, (1 - \text{rand})^{1/2}) = -0.9245.$

2.5 Napisać wzór na estymator warstwowy, jeśli warstwy definiowane są przez zmienną X .

- 2.6 [Asmussen & Glynn [?]] Przedyskutować metodę redukcji wariancji dla obliczenia

$$\int_0^\infty (x + 0.02x^2) \exp(0.1\sqrt{1 + \cos x} - x) dx$$

metodą MC.

- 2.7 Dla przykładu [??ex.YS1??] obliczyć wariancję $\hat{Y}_n^{\text{IS}}$ gdy $\tilde{f}(x) = (4 - 2x)/3$, $0 \leq x \leq 1$.

- 2.8 [Madras [?]] (i) Chcemy symulować $I = \int_0^\infty k(x) dx$, gdzie

$$k(x) = \sin(x)(1+x)^{-2}e^{-x/5}$$

za pomocą estymatora istotnościowego przy losowaniu zmiennych o rozkładzie wykładniczym z parametrem λ . Pokazać, że dla pewnych wielkości λ wariancja estymatora istotnościowego jest nieskończona.

(ii) Pokazać, że jeśli $k(x) = \sin(x)(1+x)^{-2}$, to wariancja estymatora istotnościowego przy użyciu losowań o rozkładzie wykładniczym jest zawsze nieskończona. Zaprojektować dobrą metodą symulacji I .

(iii) Przypuśćmy, że całka I jest zbieżna warunkowo (tzn. nie jest zbieżna absolutnie). Pokazać, że wtedy każdy estymator istotnościowy ma nieskończoną średnią.

- 2.9 Niech $I = \mathbb{P}(L > x)$, gdzie

$$L = \sum_{j=1}^N D_n X_n,$$

$X_1, \dots, X_n$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie $N(0,1)$, niezależnymi od $D_1, \dots, D_n$. Niech P ma rozkład $\text{Beta}(1,19)$ z gęstością $(1-p)^{18}/19$, $0 < p < 1$. Pod warunkiem $P = p$ zmienne $D_1, \dots, D_n$ są niezależnymi próbami Bernoulliego z prawdopodobieństwem sukcesu p . Podać algorytm symulacji I warunkową metodą MC, z warunkowaniem względem $\sum_{j=1}^N D_j$. Wziąć x

$$x = 3\mathbb{E} L = 3N \exp P\mathbb{E} X = 3 \cdot 100 \cdot 0.05 \cdot 3 = 45.$$

(Asmussen&Glynn,p. 147).

- 2.10 W przypadku symulacji $I = \int_{[0,1]^d} g(x) dx$, gdy d jest duże, metoda warstw po każdej zmiennej będzie nieużyteczną ze względu na dużą liczbę warstw. Remedium jest wykorzystanie tzw. kwadratów łacińskich. Niech $U_i^{(j)}$ będą niezależnymi zmiennymi losowymi o rozkładzie $U(0,1)$ oraz $\pi_i = (\pi_i(1), \dots, \pi_i(n))$ będzie ciągiem niezależnych losowych permutacji $(1, \dots, n)$, $i = 1, \dots, d$. Dla $j = 1, \dots, n$

$$V^{(j)} = (V_1^{(j)}, \dots, V_d^{(j)}) \in [0, 1]^d,$$

gdzie

$$V_i^{(j)} = \frac{\pi_i(j) - 1 + U_i^{(j)}}{n}$$

Pokazać, że estymator

$$\hat{Y}_n^{LH} = \frac{1}{n} \sum_{j=1}^n g(V^{(j)})$$

jest estymatorem nieobciążonym I . Pokazać, że

$$\text{Var } \hat{Y}_n^{LH} = \frac{\sigma_Y^2}{n} + \frac{n-1}{n} \text{Cov}(g(V^{(1)}), g(V^{(2)}))$$

gdzie $\sigma_Y^2 = \text{Var } g(U_1, \dots, U_d)$.

Zadania laboratoryjne

- 2.11 Niech $k_1(x) = e^{\sqrt{x}}$ i $k_2(x) = (1 + \frac{\sqrt{x}}{50})^{50}$ oraz $X = U$. Można oszacować, że $|k_1(x) - k_2(x)| \leq 0.03$, a więc $d \leq 0.03$. Przeprowadzić eksperyment obliczenia I oraz b obiema metodami rozpatrywanymi w podrozdziale [??ss:jcn??].
- 2.12 Niech F będzie dystrybuantą rozkładu standardowego normalnego. Niech $p_i = 1/10$. Korzystając z tablic rozkładu normalnego wyznaczyć a_j ,

takie, że

$$\begin{aligned}a_0 &= -\infty \\a_1 &= F^{-1}(1/10) \\a_2 &= F^{-1}(2/10) \\&\vdots \\a_9 &= F^{-1}(9/10) \\a_{10} &= \infty.\end{aligned}$$

W pierwszym eksperymencie przeprowadzić 100 replikacji i narysować histogram z klasami $A_1, \dots, A_{10}$. Natomiast drugi eksperyment przeprowadzić przy losowaniu w warstwach losując w każdej 10 replikacji. Porównać otrzymane rysunki.

- 2.13 Pokazać, że estymator istotnościowy do obliczenia $I = \int_E k(x)f(x)$ zdefiniowany przez

$$k(x)f(x)/I$$

ma wariancję równą zero.

Zadania teoretyczne

- 2.1 W szcelinowym modelu ALOHA z $\lambda = 0.31$ oraz $h = 0.1$ policzyć funkcję

$$\phi(k) = \mathbb{E}[X_{n+1} - X_n | X_n = k] = \lambda - b_1(k)a_0 - b_0(k)a_1, \quad k = 0, 1, \dots,$$

gdy A_i mają rozkład Poissona. Zrobić wykres. Czy można na podstawie tego wykresu skomentować rys. [fig:alohaX] i [fig:alohaB].

- 2.2 Dla sieci z rys. [rys.siec] oszacować dwustronnie I i następnie przeprowadzić analizę symulacji dla zgrubnego estymatora.

Zadania laboratoryjne

- 2.3 Wygenerować $A_1, A_2, \dots$, w odcinku $[0, 1000]$, gdzie $A_0 = 0$ oraz $A_1 < A_2 < \dots$ są kolejnymi punktami w niejednorodnym procesie Poissona z funkcją intensywności $\lambda(t) = a(2 - \sin(\frac{2\pi}{24}t))$. Przyjąć $a = 10$. Niech $A(t)$ będzie liczbą punktów w odcinku $[0, t]$. Zastanowić się jaki ma rozkład $A(t)$ i znaleźć jego średnią. Policzyć z symulacji $A(1000)/1000$ – średnią liczbą punktów na jednostkę czasu, zwaną asymptotyczną intensywnością $\bar{\lambda}$. Porównać z

$$\int_0^{24} \lambda(t) dt / 24.$$

- 2.4 Kontynuacja zad. refzad.proc.Poissona. Wygenerować $\tau_1, \tau_2, \dots, \tau_{1000}$, gdzie $\tau_i = A_i - A_{i-1}$, $A_0 = 0$ oraz $A_1 < A_2 < \dots$ są kolejnymi punktami w niejednorodnym procesie Poissona z funkcją intensywności $\lambda(t) = a(2 - \sin(\frac{2\pi}{24}t))$. Przyjąć $a = 10$. Obliczyć średni odstęp między punktami $\hat{\tau} = \sum_{i=1}^{1000} \tau_i / 1000$.
- 2.5 Zbadać jak szybko $P(L(t) = 1)$ w alternującym procesie *on-off* zbiega do współczynnika gotowości. Rozpatrzyć kilka przypadków:

- $F_{\text{on}}, F_{\text{off}}$ są wykładnicze,
- $F_{\text{on}}, F_{\text{off}}$ są Erlanga odpowiednio $\text{Erl}(3, \lambda)$ i $\text{Erl}(3, \mu)$,
- $F_{\text{on}}, F_{\text{off}}$ są Pareto.

Przyjąć $\mathbb{E} T^{\text{on}} = 1$ i $\mathbb{E} T^{\text{off}} = 2$.

- 2.6 Obliczyć średnią $\mathbb{E} L(i)$ ($i = 1, \dots, 10$) w systemie M/M/1 z
- $\lambda = 1/2$ i $\mu = 1$,
 - $\lambda = 1$ i $\mu = 1$, jeśli $L(0) = 0$.
- 2.7 Przeprowadzić symulację dobroci *czystego protokołu ALOHA*. Bardziej dokładnie, przypuśćmy, że pakiety przybywają zgodnie z procesem Poissona z intensywnością λ i wszystkie są długości 1. W przypadku kolizji, każdy z pakietów jest retransmitowany po czasie wykładniczym z parametrem μ . Zakładamy, że wszystkie zmienne są niezależne. W modelu musimy obliczyć następujące zmienne. Na podstawie symulacji zastanowić się nad przepustowością γ przy tym protokole.
- 2.8 Napisać algorytm na symulację protokołu ETHERNET.
- 2.9 Obliczyć średni czas do awarii systemu składającego się z N elementów połączonych równolegle z jednym konserwatorem, jeśli czas życia elementu ma rozkład wykładniczy $\text{Exp}(1/2)$, natomiast czas naprawy przez konserwatora $\text{Exp}(1)$. Policzyc dla $N = 1, \dots, 10$. Porównać z średnim czasem do awarii systemu składającego się z N elementów połączonych równolegle, ale bez konserwatora.

Projekt

Projekt 3 Zbadać dobroć protokołu szczelinowa ALOHA ze względu na parametr λ . Przez symulację znaleźć wartości krytyczne dla λ i określić typowe zachowanie się protokołu. Zbadać inną modyfikację protokołu szczelinowa ALOHA, w której dopuszcza się, że użytkownicy mają też dodatkową wiedzę o stanie X_n , i wtedy w n -tej szczelinie przyjmować $h = 1/X_n$. Czy ten protokół może być stabilny, tzn. mający przepustowość $\gamma > 0$; jeśli tak to kiedy.

Projekt 4 Produkt jest składany z komponentów na dwóch stanowiskach. Po zakończeniu prac na pierwszym stanowisku jest dalej opracowywany na drugim, z którego wychodzi gotowy produkt. Produkcja trwa każdego dnia przez 8 godzin. Na początku dnia w magazynie jest przygotowane n_1 komponentów, które sukcesywnie są dostarczane na stanowisko pierwsze. Przed stanowiskiem drugim może oczekiwać co najwyżej k_1 , w przeciwnym razie praca na stanowisku pierwszym jest zablokowana. Ile należy przygotować komponentów n_1 i jak duży bufor k_1 aby z prawdopodobieństwem nie mniejszym 0.9 była zapewniona produkcja przez cały dzień, tj. 8 godzin. Przez aprzestanie produkcji rozumiemy, że zarówno magazyn jak i bufor przestają pracować. Jak duży należy przygotować magazyn $n_1 + k_1$. Przyjąć, że

- czas pracy na pierwszym stanowisku ma rozkład Erlanga $\text{Erl}(2,10)$,
- natomiast na drugim stanowisku jest mieszkanką rozkładów wykładniczych z gęstością $0.8 \times 9e^{-9x} + 0.2 \times 3e^{-3x}$,
- koszty magazynowania przed rozpoczęciem pracy czy w oczekiwaniu na prace na drugim stanowisku są liniowe od $n_1 + k_1$.

Projekt 5 Mamy dwa następujące warianty pracy 2-ch procesorów.
1. Dwa procesory pracują osobno. Zadania napływają do i -tego procesu zgodnie z procesem Poissona z intensywnością λ_i i mają rozkład rozmiaru $G^{(i)}$. Niech $D^{(i)}(t)$ będzie liczbą zadań obsłużonych przez i -ty procesor to chwili t . Zakładamy, że $\lambda_i < \int_0^\infty x dG^{(i)}(x)$, $i = 1, 2$. Wtedy przepustowość

$$\lim_{t \rightarrow \infty} D^{(i)}(t)/t = \lambda_i,$$

a więc dwa procesory, jeśli pracują osobno mają przepustowość $\lambda_1 + \lambda_2$.

2. Możemy tak zmienić konstrukcję, że w przypadku gdy jeden z procesów nie ma pracy, to drugi pracuje z szybkością 2.

Zbadać o ile efektywniejszy jest zparowany procesor, w zależności od intensywności wejścia i rozkładów rozmiarów zadań. Przyjąć $\lambda_1 = 1$.

Projekt 6 W banku jest c stanowisk obsługi. Klienci zgłaszają się zgodnie z procesem Poissona z intensywnością λ i ich czasy obsługi mają rozkład G . Zbadać następujące protokoły kolejowania.

1. Przed każdym okienkiem jest osobna kolejka. Klienci wybierają okienko w sposób losowy.
2. Przed każdym okienkiem jest osobna kolejka. Klienci wybierają kolejne okienko, tj. $1, 2, \dots, c, 1, 2, \dots$
3. Jest jedna kolejka i gdy nadchodzi czas klient idzie do akurat uwolnionego okienka.

Zadania teoretyczne

- 2.1 Uzasadnić zgodność estymatorów z podrozdziału [??whit.srinikant??]. Pokazać, że estymatory B_I oraz B_S są nieobciążone jeśli proces stanu systemu $L(t)$ jest stacjonarny.

Zadania laboratoryjne

- 2.2 Przeprowadzić następujące eksperymenty dla M/M/1. Napisać program kreślący realizację $L(t)$ dla $0 \leq t \leq 10$. Eksperyment przeprowadzić dla
- (a) $\lambda = 1, \mu = 1.5$ oraz $L(0) = 0$.
 - (b) $\lambda = 1, \mu = 1.5$ i $L(0) = 2$.
 - (c) $\lambda = 1, \mu = 1.5$ i $L(0) = 4$.
 - (d) $L(0) \sim \text{Geo}(\rho)$, gdzie $\rho = \lambda/\mu = 2/3$.
- 2.3 Kontynuacja zad. 2.2. Obliczyć $\hat{L}(t_{\max})$, gdzie $\hat{L}(t) = (\int_0^t L(s) ds)/t$ gdy $L(0) = 0$.
- (a) Przeprowadzić eksperyment przy $\lambda = 1, \mu = 1.5$ oraz $L(0) = 0$ dla $t_{\max} = 5, 10, 50, 100, 1000$. Policzyc dla każdej symulacji połowy długości przedziały ufności na poziomie 95%.
 - (b) Przeprowadzić eksperyment przy $\lambda = 1, \mu = 1.1$ oraz $L(0) = 0$ dla $t_{\max} = 5, 10, 50, 100, 1000$.
- 2.4 Kontynuacja zad. 2.2. Przez replikacje rozumiemy teraz pojedynczą realizację $L(t)$ dla $0 \leq t \leq 10$. Dla replikacji $L_i(t)$ zachowujemy wektor $\mathbf{l}_i = (L_i(1), L_i(2), \dots, L_i(10))$ i niech 1000 będzie liczbą replikacji. Obliczyć

$$\hat{\mathbf{l}} = \left(\frac{1}{1000} \sum_{j=1}^{1000} L_j(0.5), \frac{1}{1000} \sum_{j=1}^{1000} L_j(1), \dots, \frac{1}{1000} \sum_{j=1}^{1000} L_j(10) \right).$$

Zrobić wykres średniej długości kolejki w przedziale $0 \leq t \leq 10$.

- (a) Przeprowadzić eksperyment dla $L(0) = 0$.
- (b) Przeprowadzić eksperyment dla $L(0) \sim \text{Geo}(\rho)$.

- 2.5 Napisać algorytm na generowanie dla systemu M/G/ ∞ estymatora

$$\frac{1}{t_{\max} - t_{\min}} \int_{t_{\min}}^{t_{\max}} L(t) dt$$

(patrz wzór ([?eq:esttmintmax?])). Wsk. Zobacz algorytm w dodatku [ss.kolejki.n.l.k] z $t_{\min} = 0$. Przeprowadzić eksperymenty omówione w przykładzie [ex.mginftytmax] z różnymi $t_{\min}$.

Projekt

2.6 Napisać program symulujący realizację procesu liczby zadań $L(t)$ w systemie M/G/ ∞ . Przeprowadzić eksperyment z

(a) $\lambda = 5$ oraz $G = \text{Exp}(\mu)$ z $\mu = 1$.

(b) $\lambda = 5$ oraz rozmiary zadań mają rozkład Pareto z $\alpha = 1.2$ przemnożonym przez 0.2 (tak aby mieć średnią 1).

Przyjąć $L(0) = 0$, $L(0) = 10$ oraz $L(0)$ ma rozkład Poissona $\text{Poi}(5)$.

Dla wszystkich wypadków oszacować rozpędówkę.

Na jednym wykresie umieścić $\int_0^t L(s) ds/t$ dla $1 \leq t \leq 100$ z symulacji

(a) i (b). Przyjąć $L(0) = 0$.

3 Zadania

6

Zadania teoretyczne

- 3.1 Napisać algorytm na generowanie błędzenia przypadkowego na hiperkostce $\{0, 1\}^m$.
- 3.2 Niech $B = \sum_{i=1}^{\infty} i^{-2}$. Ta wielkość jest znana ($B = \pi^2/6$) ale to jest nam niepotrzebne. Zaprojektować przy użyciu metody Metropolis symulację rozkładu $\pi_i = 1/Bi^2$, $i = 1, 2, \dots$. Wsk. Sąsiedzi dla dowolnego $i > 1$ to $i - 1, i + 1$ oraz dla $i = 1$ sąsiadem jest tylko 2.

Zadania teoretyczne

3.1 Dla standardowego ruchu Browna $\{W(t), 0 \leq t \leq 1\}$ definiujemy proces

$$\{\tilde{W}^h(t), 0 \leq t \leq 1\}$$

który jest zgodny z procesem W w punktach $W(h), W(2h), \dots, W(1)$ gdzie $h = 1/N$ i interpolowany liniowo do realizacji ciągłych. Pokazać, że

$$\mathbb{E} \int_0^1 |\tilde{W}^h(t) - W(t)| = c/N^{1/2},$$

gdzie $c = \sqrt{\pi/32}$.

3.2 Niech $X(t) = x + \int_0^t a(s) ds + \int_0^t b(s) dW(s)$, $0 \leq t \leq T$. Pokazać, że jeśli $0 < t_1 < t_2 < \dots < t_n$ można symulować rekurencyjnie, zadając $X(0) = x$ oraz

$$X(t_{i+1}) = X(t_i) + \int_{t_i}^{t_{i+1}} a(s) ds + \sqrt{\int_{t_i}^{t_{i+1}} b^2(s) ds} Z_{i+1}$$

dla $i = 1, \dots, n$, i $Z_1, \dots, Z_n$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie $N(0,1)$.

Projekt

- 3.3 Według Asmussen&Glynn (2007). Zaprojektować obliczenie p metodą Monte Carlo, gdzie p jest intensywnością spłaty kredytu w następującym modelu.

Udzielono kredytu w wysokości k na okres T lat, który będzie spłacany ze stałą intensywnością p . Krótkoterminowa stopa kredytowa jest $r(t)$, $0 \leq t \leq T$. Jeśli $s(t)$ oznacza kwotę spłaconą do chwili t ; $s(0) = 0$ raz $q(t)$ oznacza ilość pieniędzy gdyby były trzymane na rachunku, to ignorując pozostałe koszty (administracja, podatki, zysk, itd) wielkość p wyznaczamy z równości $\mathbb{E} s(T) = \mathbb{E} q(T)$.

- a) Uzasadnić, że

$$\begin{aligned} ds(t) &= (p + s(t)r(t))dt \\ dq(t) &= q(t)r(t) \end{aligned}$$

- b) Jeśli przez $s_p(t)$ oznaczmy proces s obliczony przy intensywności spłaty p oraz $q_k(t)$ oznaczmy proces obliczony przy wartości początkowej $q(0) = k$, to pokazać, że

$$p = \frac{k \mathbb{E} q_1(T)}{\mathbb{E} s_1(T)}.$$

- c) Przyjąć, że $r(t)$ jest procesem CIR. W pierwszym wariancie założyć, że jest on stacjonarny, ze średnią równą długoterminowej stopie procentowej c . Do obliczeń przyjąć, że długoterminowa stopa procentowa $c = 7\%$, $T = 5$, oraz dobrać α, σ^2 tak aby realizacje przypominały te z wykresu ...

- d) Porównać z sytuacją, gdy $r(t) \equiv c$.

Skorowidz nazw

- n*-ty moment, 207
- średnia, 207
- średnia pęczkowa, 160
- średnia próbkowa, 227, 228
- błądzeniem przypadkowe
 - proste, symetryczne, 29
- Rozkład (dyskretny) jednostajny, 216
- rozkład ujemnie dwumianowy, 216
- zbieżność według prawdopodobieństwa, 212
- zbieżność z prawdopodobieństwem 1, 212
- algorytm, losowa permutacja, 26, ITM-Exp43, ITM-Par44, ITM-d45, DB47, ITR47, DP49, RM57, BM60, MB61
- algorytm , ITM43
- ALOHA
 - szczelinowa, 135
- alokacja proporcjonalna, 97
- asymptotyczna wariancja, 151
- asymptotyczne obciążenie, 152
- błądzenie przypadkowe, 243
- błądzenie przypadkowe
 - proste symetryczne, 244
- budżet, 83
- centralne twierdzenie graniczne, 81
- ciąg stacjonarny, 171
- CMC, 84
- CTG, 81
- czysty protokół ALOHA, 144
- długość rozpędówki, 159
- discrete event simulation, 124
- distribution
 - conditional, 209
 - Pareto, 218
- dni urodzin, 23
- dyspersja, 208
- dystrybuenta, 205
- dystrybuenta empiryczna, 15, 229
- efektywność, 84
- estymator, λ -trafił85, λ żenie86, 227, zgodny227, λ żony227
- estymator antytetyczny, 104
- estymator chybił trafił, 85
- estymator istotnościowy, 111
- estymator mocno zgodny, 80
- estymator nieobciążony, 80
- estymator zgodny, 227
- FCFS, 236
- funkcja partycji, 182
- funkcja charakterystyczna, 211
- funkcja tworząca momenty, 211
- function
 - moment generating, 211
- fundamentalny związek, 81, 151

- funkcję tworzącą, 211
- funkcja
 - kwantyl, 230
- funkcja intensywności, 122
- funkcja prawdopodobieństwa, 205
- gęstość, 206
- gęstością, 206
- generator liczb losowych, 7
- GLL, 7
- iloraz wiarygodność, 114
- kowariancja, 208
- LCFS, 236
- liczba losowa, 5
- logarytmicznie efektywny, 199
- mediana, 208
- metoda środka kwadratu, 6
- metoda delty, 86
- metoda eliminacji, 53, ciąga55
- metoda kongruencji liniowej, 7
- metoda warstw, 94
- metoda wspólnych liczb losowych, 105
- metoda zmiennych antytetycznych, 103
- metoda zmiennych kontrolnych, 107
- mieszanka, 206
- momenty zgłoszeń, 120
- niezależne zmienne losowe, 208
- obsługa z podziałem czasu, 236
- obsługa w losowej kolejności, 236
- obsługa w odwrotnej kolejności zgłoszeń
 - 236
- obsługa zgodna z kolejnością zgłoszeń,
 - 236
- obszar akceptacji, 54, 55
- odchylenie standardowe, 208
- ogon, 205
- ogon dystrybuanty, 42
- ograniczony błąd względny, 198
- Pareto distribution, 218
- pierwszy moment, 207
- poziom istotności, 80
- poziom ufności, 80
- próba prosta, 227
- prawo arcusa sinusa, 30
- proces liczący, 120
- proces odnowy, 124
- proces Poissona, 121, niejednorodny122
 - intensywność, 121
- proces Poissona , jednorodny121
- proces stacjonarny, 148
- przedział ufności , 228
- PS, 236
- Q-Q wykres, 229
- quantile function, 230
- regulamin obsługi, 236
- replikacja, 31
- rodzina zdarzeń rzadkich, 197
- rozkład
 - ciężki ogon, 215
 - lekki ogon, 215
- Rozkład Bernoulliego, 215
- rozkład Beta, 73
- rozkład beta, 217
- rozkład chi kwadrat, 217
- rozkład dwumianowy, 215
- Rozkład Erlanga, 217
- rozkład Frécheta, 218
- rozkład gamma, 217
- rozkład geometryczny, 216
- rozkład Gumbela, 218

- rozkład jednostajny, 217
- rozkład kratowy, 205
- rozkład logarytmiczno normalny, 218
- rozkład logarytmiczny, 216
- rozkład loggamma, 218
- rozkład logistyczny, 71
- rozkład Makehama, 75
- rozkład normalny, 217
- rozkład obcięty geometryczny, 216
- rozkład odwrotny gaussowski, 218
- rozkład Pascala, 216
- rozkład podwójny wykładniczy, 72
- Rozkład Poissna, 215
- rozkład Raleigha, 59
- rozkład statystyki ekstremalnej, 218
- rozkład trójkątny, 71
- rozkład Weibulla, 218
- rozkład wielomianowy, 74
- rozkład wielowymiarowy normalny, 219
- rozkład wykładniczy, 217
- rozkład zdegenrowany, 215
- rozkład ekstremalny Weibulla, 218
- słaba zbieżność, 213
- SIRO, 236
- symetrycznym błędzeniem przypad-
kowe
proste, 29
- symulacja asynchroniczna, 120
- symulacja synchroniczna, 120
- TAVC, 151
- test kolizji, 21
- test odstępów, 23
- transformata Laplace-Stieltjesa, 211
- twierdzenie
Gliwienko-Cantelli, 229
- uogólniona funkcja odwrotna, 41
- wariancja, 207
- wariancja próbkowa, 227
- wartość oczekiwana, 207
- warunek wielkiego załadowania, 164
- warunkową gęstością, 207
- warunkowa wartość oczekiwana, 209
- warunkowe prawdopodobieństwo con-
ditional, 209
- wektor losowy
absolutnie ciągły, 206
- współczynnik asymetrii, 208
- współczynnik zmienności, 208
- wzór Itô, 214
- załadowanie, 237
- zbiór ufności, 228
- zbieżność według rozkładu, 213
- zgrubny estymator Monte Carlo, 84
- zmienna losowa
absolutnie ciągła, 206
dyskretna, 205