Wykłady z symulacji stochastycznej i teorii Monte Carlo

Tomasz Rolski

Instytut Matematyczny, Uniwersytet Wrocławski

Wrocław 2009

Skrypt został opracowany na podstawie notatek do wykładu *Symulacja*, który odbył się w semestrze letnim 2009 r. Prezentowany materiał jest w trakcie opracowywania. Tekst może zawierać błędy. Układ skryptu jak i oznaczenia mogą ulegać zmianom. Autor będzie wdzięczny za wszelkie uwagi oraz listy błędów, ktore można przesłać e-mailem na adres: rolski@math.uni.wroc.pl.

Wersja drukowana w dniu 19 czerwca 2009 r.

Spis treści

Rozdział I

Wstęp

1 O skrypcie

Stochastyczne symulacje i *Teoria Monte Carlo* to bardzo często są wymiennymi terminami. Jednakże w opinii autora tego skryptu warto je rozróżniać. Pierwszy termin dotyczy netod generowania liczb losowych a właściwue pseudo-losowych, bo tylko takie mogą być genrowane na komputerze. Natomiast teoria monte Carlo zajmuje się teoretycznymi podstawami opracowania wyników oraz planowania symulacji. W przeciwieństwie do metod numerycznych, gdzie wypracowane algorytmy pozwalają kontrolować deterministycznie błędy, w przypadku obliczeń za pomocą metod stochastycznych dostajemy wynik losowy, i należy zrozumięc jak taki wynik imterpretować, co rozumiemy pod pojęciem błędu.

Skrypt jest pisany z myślą o wykładzie dla studentów Uniwersytetu Wrocławskiego, w szczególności studentów na kierunku matematyk, specjalnośc informatyczna.

2 Generatory liczb losowych

3 Symulacja zadań kombinatorycznych

Zaczniemy od przedyskutowanie algorytmu na generowanie losowej permutacji liczb $1, \ldots, N$. Celem jest wylosowani permutacji tych liczb takiej że kazdy wynik można otrzymać z prawdopdobieństwem $1/N!$. Zakładamy,

że mamy do dyspozycji generator liczb losowych $U_1, U_2, \ldots$, tj. ciąg niezależnych zmiennych losowych o jednakwym rozkładzie jednostajnym $\mathrm{U}(0,1)$. Generowanie losowej permutacji ciągu $1, \ldots, N$

Losowa permutacja

```
1. podstaw  t = N oraz A[i] = i dla i = 1, . . . , N;
2. generuj liczbę losową u pomiędzy  0 i 1;
3. podstaw  k = 1 + tu ; zamie\'n  A[k] z A[t];
4. podstaw  t = t - 1;
   jeśli t > 1, to powrót do kroku 2;
   w przeciwnym razie stop i A[1], . . . , A[N ]
   podaje losową  permutację.
```

Złożoność tego algorytmu jest $\mathrm{O}(N)$.
W MATLAB-ie generujemy losową permutacę za pomocą instrukcji

```
randperm
```

W rozdziale 2-gim książki Fellera [2] mamy zadanie o dniach urodzin. Na podastawie tego fragmentu można sformułować następujący zadanie na symulację przy użyciu komendy randmperm.

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 ten sam dzień urodzin w przeciągu r dni jeden od drugiego?

3. Przypuśćmy, że osoby pojawiają się kolejno. Jak długo trzeba czekac aby pojawiły sie dwie mające wspolne urodziny?

4 Proste zadania probabilistyczne

W rozdziale 3-cim książki Fellera [2] rozważa się rzuty symetryczną monetą. w n-tym rzucie (niezależnym od pozostałych) wygrywamy $+1$ gdy wypadnie orzeł ($=1$) oraz przegrywamy -1 gdy wypadnie reszka. Przypuścmy zaczynamy grać z pustym kontem; tj $S_0 = 0$. Niech S_n będzie wygraną po n rzutach. Możemy postawić różne pytania. Na przykład

1. Jak wyglądają oscylacje S_n.

2. Jakie jest prawdopodobieństwo $P(\alpha, \beta)$, że przy n rzutach będziemy nad kreską w przedziale pomiędzy $100\alpha\%$ a $100\beta\%$ procent czasu?

Przykładows symulacja prostego symetrycznego błądzenia przypadkowego. ciągu $1, \ldots, N$ można otrzymać za pomocą następującego prostego programu. (dem2-1.m)

Realizacja prostego błądzenia przypadkowego

```
N=100;
xi=2*floor(2*rand(1,N))-1;
A=triu(ones(N));
y=xi*A;
for i=2:N+1
  s(i)=y(i-1);
end
s(1)=0;
x=1:N+1;
plot(x,s)
```

5 Zagadnienie sekretarki

Jest to problem z dziedziny optymalnego zatrzymania. Rozwiązanie jego jest czasami zwane regułą 37%. Można go wysłowić następująco:

- Jest jeden etat do wypełnienia.
- Na ten 1 etat jest n kandydatów, gdzie n jest znane.
- Komisja oceniająca jest w stanie zrobić ranking kandydatów w ściślym sensie (bez remisów).
- Kandydaci zgłaszają się w losowym porządku. Zakłada się, że każde uporządkowanie jest jednakowo prawdopodobne.
- Po interview kandydat jest albo zatrudniony albo odrzucony. Nie można potem tego cofnąć.

- Decyzja jest podjęta na podstawie względnego rankingu kandydatów spotkanych do tej pory.
- Celem jest wybranie najlepszego zgłoszenia.

Okazuje się, że optymalna strategia, to jest maksymizująca prawdopodobieństwo wybrania najlepszego kandydata jest, należy szukać wśród następujących strategi: opuścić k kandydatów i następnie przyjąć pierwszego lepszego od dotychczas przeglądanych kandydatów. Jesli takiego nie ma to bierzemy ostatniego. Dla dużych n mamy $k \sim n/e$ kandydatów. Dla małych n mamy

n	1	2	3	4	5	6	7	8	9
k	0	0	1	1	2	2	2	3	3
P	1.000	0.500	0.500	0.458	0.433	0.428	0.414	0.410	0.406

To zadanie można prosto symulować jeśli tylko umiemy generować losową permutację. Możemy też zadać inne pytanie. Przypuśćmy, że kandydaci mają rangi od 1 do n, jednakże nie są one znane komisji zatrudniającej. Tak jak poprzednio komisja potrafii jedynie uporządkować kandydatów do tej pory starających sie o pracę. Możemy zadać pytanie czy optymalny wybór maksymizujący prawdopodobieństwo wybrania najlepszego kandydata jest również maksymizujący wartość oczekiwaną rangi zatrudnianej osoby. Na to pytanie można sobie odpowiedzieć przez symulację.

6 Uwagi bibligraficzne

Na rynku jest sporo dobrych podręczników o metodach stochastycznej symulacji i teorii Monte Carlo. Zacznijmy od najbardziej elementarnych, jak na przykład ksiązka Rossa [9]. Z pośród bardzo dobrych książek wykorzystujących w pełni nowoczesny aparat probabilistyczny i teorii procesów stochastycznych można wymienić książkę Asmussena i Glynna [1], Madrasa [6], Fishmana , Ripleya [8]. Wśród książek o symulacjach w szczególnych dziedzinach można wymienić książkę Glassermana [5] (z inżynierii finansowej), Fishmana [4] (symulację typu *discete event simulation*. W języku polskim można znaleźć książkę Zielińskiego [15].

Wykłady prezentowane w obecnym skrypcie były stymulowane przez notatki do wykładu prowadzonego przez J. Resinga [7] na Technische Universiteit Eindhoven w Holandii. W szczegolności autor skryptu dziękuje Jo-

hanowi van Leeuwaardenowi za udostępnienienie materiałów. W szczególności w niektórych miejscach autor tego skryptu wzoruje się na wykładach Johana van Leeuwardena i Marko Bono w Technicznym Instytucie Technologi w Eidhoven.

Zadania laboratoryjne

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

6.1 Uzasadnić, że procedura generowania losowej permutacji z wykładu daje losową permutację.

6.2 Napisać procedurę
1. z n obiektów wybieramy losowo k ze zwracaniem,
2. z n obiektów wybieramy k bez zwracania.
3. wybierającą losowy podzbiór k elementowy z $\{1, \ldots, n\}$.

6.3 Napisać procedurę n rzutów monetą $(\xi_1, \ldots, \xi_n)$.
Jeśli orzeł w i-tym rzucie to niech $\eta_i = 1$ jeśli reszka to $\eta_i = -1$. Niech $S_0 = 0$ oraz $S_k = \xi_1 + \ldots + \xi_k$. Obliczyć pierwszy moment i wariancję ξ_1.
Zrobić wykres $S_0, S_1, \ldots, S_{1000}$. Na wykresie zaznaczyc linie: $\pm\mathrm{Var}(\xi)\sqrt{n}$, $\pm 2\mathrm{Var}(\eta)\sqrt{n}$, $\pm 3\mathrm{Var}(\xi)\sqrt{n}$.
Zrobić zbiorczy wykres 10 replikacji tego eksperymentu.

6.4 Rzucamy N razy i generujemy błądzenie przypadkowe $(S_n)_{0 \le n \ge N}$. Napisać procedury do obliczenia:
1. Jak wygląda absolutna różnica pomiędzy maximum i minimum błądzenia przypadkowego gdy N rośnie?

6.5 Rozważamy ciąg $(S_n)_{0 \le n \ge N}$ z poprzedniego zadania. Przez prowadzenie w odcinku $(i, i+1)$ rozumiemy, że $S_i \ge 0 oraz S_{i+1} \ge 0$. Niech X będzie łączną długością prowadzeń w $(0, N)$.
1. Jakie jest prawdopodobieństwo, że jeden gracz będzie przeważał pomiędzy 50% a 55% czasu? Lub więcej niż 95% czasu?
3. Niech $N = 200$. Zrobić histogram wyników dla 500 powtórzeń tego eksperymentu.

6.6 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. Przypuśćmi, że osoby pojawiają się kolejno. Jak długo trzeba czekac

aby pojawiły sie dwie mające wspólne urodziny? Zrobić histogram dla 100 powtórzeń.

6.7 (Opcjonalne) Mamy 10 adresów, których odległości są zadane w następujący sposób. Zacząć od rand('state', 0). Przydzielić odległość kolejno wierszami a[ij] - odległość między adresem i oraz j. (W ten sposób wszyscy będą mieli taką samą macierz odległości).
Przez drogę wyznaczoną przez permutację π z dziesięciu liczb rozumiemy $(\pi(1), \pi(2)), \ldots (\pi(9), \pi(10))$. Wtedy odległość do pokonania jest $\sum_{j=2}^{10} a[\pi(i-1), a\pi(i)]$. Czy uda się odnaleźć najkrótszą trasę. Znaleźć najkrótszą trase po 100 losowaniach. Po 1000?

Projekt

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

Napisz raport w jaką stategię powinien przyjąc Bob. W szczególności w raporcie pwoinny się znaleźc odpowiedzi na następujące pytania.

(a) Bob decyduje się na ślub z pierwszą spotkaną partnerką. Oblicz prawdopodobieństwo P (teoretyczne), że ożeni się z najlepszą. Oblicz też oczekiwany rangę r

(b) Użyj symulacji aby określić P i r dla następującej strategii. Bob nigdy nie żeni się z pierwszą, żeni się z drugą jesli jest lepsza od pierwszej, w przeciwnym razie żeni się z trzecią.

(c) Rozważyć zadanie z 10-cioma partnerkami, które mają rango $1, 2, \ldots, 10$. Obliczyć teoretycznie prawdopodobieństwo P i oczekiwaną rangę r dla strategii jak w zadaniu (a).

(d) Zdefiniować zbiór strategii, które są adaptacją strategii z zadania (b) na 10 partnerek. Podać jeszcze inną strategię dla Boba. Dla każdej stategii obliczyć P i r.

Rozdział II

Symulacja zmiennych losowych

1 Metoda ITM

Naszym celem jest przeprowadzenie losowania w wyniku którego otrzymamy liczbę losową Z z dystrybuantą F mając do dyzpozycji zmienną losową U o rozkładzie jednostajnym $U[0,1]$.

Niech

$$F^{\leftarrow}(t) = \inf\{X : \ t \leq F(x)\}$$

W przypadku gdy $F(x)$ jest funkcją ściśle rosnącą, $F^{\leftarrow}(t)$ jest funkcją odwrotną $F^{-1}(t)$ do $F(x)$. Dlatego $F^{\leftarrow}(t)$ nazywa się *uogólnioną funkcją odwrotną*

Zbadamy teraz własności uogólnionej funkcj odwrotnej $F^{\leftarrow}(x)$. Jeśli F jest ściśle rosnąca to jest prawdziwa następująca równoważność: $u = F(x)$ wtedy i tylko wtedy gdy $F^{\leftarrow}(u) = u$. Ale można podać przykład taki, że $F^{\leftarrow}(u) = x$ ale $u > F(x)$. Natomiast zawsze są prawdziwe następujące własności. Jest prawdą, że zawsze $F^{\leftarrow}(F(x)) = x$, ale łatwo podać przykład, że $F(F^{\leftarrow}x) \neq x$. W skrypcie będziemy oznaczać *ogon dystrybuanty* $F(x)$ przez $\overline{F}(x) = 1 - F(x)$.

Lemat 1.1 (a) $u \leq F(x)$ *wtedy i tylko wtedy gdy* $F^{-1}(u) \leq u$.
(b) *Jeśli* $F(x)$ *jest funkcją ciągłą, to* $F(F^{\leftarrow}(x)) = x$.
(c) $\bar{F}^{-1}(x) = F^{-1}(1-x)$.

Fakt 1.2 *(a) Jeśli* U *jest zmienną losową, to* $F^{-1}(U)$ *jest zmienną losową z dystrybuntą* F.
(b) Jesli F *jest ciągłą i zmienna* $X \sim F$, *to* $F(X)$ *ma rozkład jednostajny U(0,1)*

Dowód (a) Korzystając z lemat 1.1 piszemy

$$\mathbb{P}(F^{\leftarrow}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x) \, .$$

(b) Mamy

$$\mathbb{P}(F(X) \geq u) = \mathbb{P}(X \geq F^{\leftarrow}(u))$$

Teraz korzystając z tego, że zmienna losowa X ma dystrybuantę ciągłą, a więc jest bezatomowa, piszemy

$$\mathbb{P}(X \geq F^{\leftarrow}(u)) = \mathbb{P}(X > F^{\leftarrow}(u)) = 1 - F(F^{\leftarrow}(u))$$

które na mocy lematu 1.1 jest równe $1 - u$. □

Korzystając z powyższgo faktu mamy następujący algorytm na generowanie zmiennej losowej o zadanym rozkładzie F

Algorytm ITM

```
% Dane $g(t)=F^{-1}(t)$. Otrzymujemy $Z\sim F$.
Z=g(rand);
```

Zauważmy też, że jeśli $U_1, U_2, \ldots$ jest ciągiem zmiennych losowych o jednakowym rozkładzie jenostajnym,, to $U_1', U_2' \ldots$, gdzie $U_j' = 1 - U_j$, jest też ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym. Oznaczmy przez

$$g^*(x) = g(1 - x) = F^{-1}(x).$$

Niestety jawne wzory na funkcję $g(x)$ czy $g^*(x)$ znamy w niewielu przypadkach, które omówimy poniżej.

Rozkład wykładniczy Exp(λ). W tym przypadku

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \exp(-\lambda x), & x \geq 0 \end{cases} \tag{1.1}$$

Jeśli $X \sim \mathrm{Exp}(\lambda)$, to $\mathbb{E}\, X = 1/\lambda$ oraz $\mathrm{Var}\, X = 1/\lambda$. Jest to zmienna z lekkim ogonem ponieważ

$$\mathbb{E}\, e^{sX} = \frac{\lambda}{\lambda - s}$$

jest określona dla $0 \leq s < \lambda$. Funkcja odwrotna

$$g^*(x) = -\frac{1}{\lambda}\log(x)\ .$$

Zmienne losowe wykładnicze generujemy następującym algorytmem.

Algorytm ITM-Exp.

```
% Dane $\lambda$. Otrzymujemy $Z\sim \mbox{Exp}(\lambda)$.
Z=-log(rand)$/\lambda$
```

Zajmiemy się teraz przykładem zmiennej ciężko-ogonowej.

Rozkład Pareto Par(α). W tym przypadku

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \frac{1}{(1+x)^\alpha}, & x \geq 0 \end{cases} \tag{1.2}$$

Jeśli $X \sim \mathrm{Par}(\alpha)$, to $\mathbb{E}\,X = 1/(\alpha-1)$ pod warunkiem, że $\alpha > 1$ oraz $\mathbb{E}\,X^2 = 2/((\alpha-1)(\alpha-2))$ pod warunkiem, że $\alpha > 2$. Jest to zmienna z cieżkim ogonem ponieważ

$$\mathbb{E}\,e^{sX} = \infty$$

dla dowolnego $0 \leq s < s_0$. Funkcja odwrotna

$$g^*(x) = x^{-1/\alpha} - 1.$$

Zauważmy, że rozkład Pareto jest tylko z jednym parametrem kształtu α. Dlatego jeśli chcemy miec szerszą klasę dopuszczającą dla tego samego parametru kształtu zadaną średnią m skorzystamy z tego, że $\mathbb{E}\,m(\alpha-1)X = m$. Oczywiście $\alpha > 1$. Można więc wprowadzić dwuparametrową rodzinę rozkładów Pareto $\mathrm{Par}(\alpha, c)$. Zmienna Y ma rozkład $\mathrm{Par}(\alpha, c)$ jeśli jej dystrybunata jest

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \frac{c^\alpha}{(c+x)^\alpha}, & x \geq 0 \end{cases} \tag{1.3}$$

Algorytm ITM-Par.

```
% Dane $\alpha$. Otrzymujemy $Z\sim \mbox{Par}(\lambda)$.
Z=rand${}^{(-1/\alpha)}-1$.
```

2 Metoda ITM oraz ITR dla rozkładów kratowych

Przypuśćmy, że rozkład jest kratowy skoncentrowany na $\{0, 1, \ldots\}$. Wtedy do zadania rozkładu wystarczy znać funkcję prawdopodobieństwa $\{p_k,\ k = 0, 1, \ldots\}$.

Fakt 2.1 *Niech U będzie liczbę losową. Zmienna*

$$Z = \min\{k : U \leq \sum_{i=0}^{k} p_i\}$$

ma funkcję prawdopodobieñstwa $\{p_k\}$.

Stąd mamy następujący algorytm na generowanie kratowych zmiennych losowych.

Algorytm ITM-d.

```
% Dane $p_k=p(k)$. Otrzymujemy $Z\sim \{p_k\}$.
U=rand;
Z=0;
S=p(0);
while U$>$S
  Z=Z+1;
  S=S+p(Z);
end
```

Fakt 2.2 *Liczba wywołań N w pętli* while *w powyższym algorytmie ma funkcję prawdopodobieństwa $\mathbb{P}(N = k) = p_{k-1},\ k = 1, 2, \ldots,$ skąd średnia liczba wywołań wynosi $\mathbb{E}\, N = \mathbb{E}\, Z + 1$.*

Rozkład geometryczny Geo(p) ma funkcję prawdopodobieństwa

$$p_k = (1 - p)p^k, \qquad k = 0, 1, \ldots$$

Zmienna $X \sim \text{Geo}(p)$ ma średnią $\mathbb{E}\, =$ oraz ... Z Faktu 1.2 mamy, że

$$X = \lfloor \log U / \log p \rfloor$$

ma rozkład Geo(p). Mianowicie

$$\begin{aligned} \mathbb{P}(X \geq k) &= \mathbb{P}(\lfloor \log U / \log p \rfloor \geq k) \\ &= \mathbb{P}(\log U / \log p \geq k) \\ &= \mathbb{P}(\log U) \leq (\log p)k) \\ &= \mathbb{P}(U \leq p^k) = p^k \ . \end{aligned}$$

W wielu przypadkach następująca funkcja

$$c_{k+1} = \frac{p_{k+1}}{p_k}$$

jest prostym wyrażeniem. Podamy teraz kilka przykładów.

- $X \sim \mathrm{B}(n, p)$. Wtedy

$$p_0 = (1-p)^n, \quad c_{k+1} = \frac{p(n-k)}{(1-p)(i+1)}.$$

- $X \sim \mathrm{Poi}(\lambda)$. Wtedy

$$p_0 = e^{-\lambda}, \quad c_{k+1} = \frac{\lambda}{k+1}.$$

- $X \sim \mathrm{NB}(r, p)$. Wtedy

$$p_0 = (1-p)^r, \quad c_{k+1} = \frac{(r+k)p}{k+1}.$$

W przypadki gdy znana jest jawna i zadana w prostej postaci funkcja c_{k+1} zamiast algorytmu ITM-d możemy używać następującego. Używamy oznaczenia c_{k+1} =c(k+1) i p_0 =p(0).

Algorytm ITR.

```
% Dane p(0) oraz c(k+1), k=0,1,... Otrzymujemy X z f. prawd. {p_k}.
S=p(0);
P=p(0);
X=0;
U=rand;
```

```
while (U>S)
  X=X+1;
  P=P*c(X);
  S=S+P;
end
```

3 Metody specjalne.

3.1 B(n, p)

Niech $X \sim \mathrm{B}(n, p)$. Wtedy

$$X =_{\mathrm{d}} \sum_{j=1}^{n} \xi_j,$$

gdzie $\xi_1, \ldots, \xi_n$ jest ciągiem prób Bernoulliego, tj. ciąg niezależnych zmiennych losowych $\mathbb{P}(\xi_j = 1) = p$, $\mathbb{P}(\xi_j = 0) = 1 - p$.

Algorytm DB

```
% Dane n,p. Otrzymujemy X z f. prawd. B(n,p)
X=X+1;
for i=0:n
  if (rand<= p)
    X=X+1;
  en
end
```

3.2 Poi(λ).

Teraz podamy algorytm na generowanie zmiennej $X \sim \mathrm{Poi}(\lambda)$ o rozkładzie Poissona. Niech $\tau_1, \tau_2, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie Exp(1). Niech

$$N = \#\{i = 1, 2, \ldots : \tau_1 + \ldots + \tau_i \leq \lambda\}. \tag{3.4}$$

Fakt 3.1

$$N =_{\mathrm{d}} Poi(\lambda).$$

Dowód Mamy

$$\{N \leq k\} = \{\sum_{j=0}^{k} \tau_j \geq \lambda\}.$$

Ponieważ

$$\mathbb{P}(\sum_{j=0}^{k} \tau_j > x) = e^{-x}(1 + x + \frac{x^2}{2!} + \ldots + \frac{x^k}{k!}),$$

więc

$$\mathbb{P}(N = 0) = \mathbb{P}(N \leq 0) = e^{-x},$$

oraz

$$\mathbb{P}(N = k) = \mathbb{P}(\{N \leq k\} \cap \{N \leq k-1\}^c) = \mathbb{P}(\{N \leq k\}) - \mathbb{P}(\{N \leq k-1\}) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

□

Wzór (3.4) można przepisać następująco:

$$X = \begin{cases} 0 & \tau_1 > \lambda \\ \max\{n \geq 1 : \tau_1 + \ldots + \tau_i \leq \lambda\} & \text{w przeciwnym razie.} \end{cases}$$

Na tej podstawie możemy napisać następujący algorytm.

Algorytm DP

```
% Dane $\lambda$. Otrzymujemy X$\sim\mbox{Poi}(\lambda)$.
X=0;
S=-1;
while (S<=lambda)
  Y=-log(rand);
  S=S+Y;
  X=X+1;
end
```

Algorytm DP możne jeszcze uprościć korzystając z następujących rachunków. Jeśli $\tau_1 \leq \lambda$ co jest równowazne $-\log U_1 \leq \lambda$ co jest równoważne $U_1 \geq$

$\exp(-\lambda)$, mamy

$$\begin{aligned} X &= \max\{n \geq 1 : \tau_1 + \ldots + \tau_i \leq \lambda\} \\ &= \max\{n \geq 1 : (-\log U_1) + \ldots + (-\log U_n \leq \lambda\} \\ &= \max\{n \geq 1 : \log \prod_{i=1}^{n} U_i \geq -\lambda\} \\ &= \max\{n \geq 1 : \prod_{i=1}^{n} U_i \geq e^{-\lambda}\} \ . \end{aligned}$$

3.3 Erl(n, λ)

Jeśli $\xi_1, \ldots, \xi_n$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie Exp(λ), to

$$X = \xi_1 + \ldots + \xi_n$$

ma rozkład Erlanga Erl(n, λ). Wiadomo, że jest to szczególny przypadek rozkładu Gamma(n, λ) z gęstością

$$f(x) = \frac{1}{\Gamma(n)} \lambda^n x^{n-1} e^{-\lambda x}, \qquad x \geq 0 \ .$$

4 Przykładowe symulacje

W dwóch wykresach na rys. 4.1 i 4.2 podajemy wyniki S_n/n, $n = 1, \ldots, 5000$ dla odpowiednio rozkład wykładniczego i Pareto ze średnią 1. Zmienne Pareto są w tym cely pomnożone przez 0.2. Warto zwrócić uwagę na te symulacje. Wdać, że dla zmiennych Pareto zbieżność jest dramatycznie wolniejsza.

Na rys. 4.3 podane sa wyniki symulacji, kolejno rozkładu normalnego N(1,1), wykładniczego Exp(1), oraz Pareto(1,2) przemnożonego przez 0.2. Polecamy zwrócić uwagę na charakter poszczególnych fragmentów na rys. 4.3. Chociaż wszystkie zmienne mają średnią 1, to w pierwszym fragmencie zmienne są mało rozrzucone, w drugiej części trochę bardzie natomiast duzy rozrzut można zaobserwować w części ostatniej. Jest to spowodowane cora cięższym ogonem rozkładu.

Rysunek 4.1: Rozkład wykladniczy Exp(1), 5000 replikacji

5 Metoda eliminacji i ilorazu jedostajnych

Wstępem do tego podrozdziału niech będzie nastepujaca obserwacja. Przypuścy, że A i B będą pozdbiorami $\mathbb{R}^n$ dla których istnieje objętość $\int_B d\boldsymbol{x} = |B|$ i $\int_A d\boldsymbol{x} = |A|$ oraz $A \subset B$. Jesli $\boldsymbol{X}$ ma rozkład jednostajny U(B), to $(X|X \in A)$ ma rozkład jednostajny U(A). Mianowicie trzeba skorzystać, że dla $D \subset A$

$$\mathbb{P}(X \in D|X \in A) = \frac{|D \cap |/|B|}{|A|/|B|} = \frac{|D|}{|A|}.$$

Stąd mamy następujący algorytm na generowanie wektora losowego $\boldsymbol{X} = (X_1, \ldots, X_n)$ o rozkładzie jednostajnym U(A), gdzie A jest ograniczonym podzbiorem $\mathbb{R}^n$. Mianowicie taki zbiór A jest zawarty w pewnym prostokącie $B = [a_1, b_1] \times \ldots \times [a_n, b_n]$. Latwo jest teraz wygenerować wektor losowy $(X_1, \ldots, X_n)$ o rozkładzie jednostajnym U(B) ponieważ wtedy zmienne X_i są niezależne o rozkładzie odpowiednio U$[a_i, b_i]$ $(i = 1, \ldots, n)$.

Algorytm U(A)

```
%na wyjściu X o rozkładzie jednostajnym w A.
1. for i=1:n
     Y(i)=(b(i)-a(i))*U(i)+b(i)
```

Rysunek 4.2: Rozkład Pareto(1.2), 5000 replikacji

```
   end
2. jeśli Y=(Y(1),...,Y(n) należy do A to X:= Y,
3. w przeciwnym razie idź do 1.
```

Metoda eliminacji pochodzi od von Neumanna. Zaczniemy od zaprezentowania jej dyskretnej wersji. Przypuśćmy, ze chcemy generować zmienną losową X z funkcją prawdopdobieństwa $(p_j, j \in E)$, podczas gdy znamy algorytm generowania liczby losowej Y z funkcją prawdopdobieństwa $(p_j^{'}, j \in E)$. Niech $c > 0$ będzie takie, że

$$p_j \leq cp_j^{'}, \qquad j \in E.$$

Oczywiście takie c zawsze znajdziemy gdy E jest skończona, ale tak nie musi być gdy E jest przeliczalna; polecamy czytelnikowi podać przykład. Oczywiście $c > 1$, jeśli funkcje prawdopodobieństwa są różne. Niech $U \sim \mathcal{U}(0,1)$ i definiujemy *obszar akceptacji* przez

$$\{c\,U\,p_Y^{'} \leq p_Y\}.$$

Zakładamy, że U i Y są niezależne. Podstawą naszego algorytmu będzie następujący fakt.

Fakt 5.1

$$\mathbb{P}(Y = j|A) = p_j, \qquad j \in E$$

oraz $\mathbb{P}(A) = 1/c$

Rysunek 4.3: Rozkład N(1,1) (5000 repl.), Exp(1) (5000 repl.), Pareto(1.2), (5000 repl.)

Dowód

$$\begin{aligned}
\mathbb{P}(Y=j|A) &= \frac{\mathbb{P}(\{Y=j\}\cap\{cUp_Y^{'}\le p_Y\})}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(Y=j, U\le p_j/(cp_j^{'}))}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(Y=j)\mathbb{P}(U\le p_j/(cp_j^{'}))}{\mathbb{P}(A)} \\
&= p_j^{'}\frac{p_j/(cp_j^{'})}{\mathbb{P}(A)} = p_j\frac{c}{\mathbb{P}(A)}\ .
\end{aligned}$$

Teraz musimy skorzystać z tego, że

$$1=\sum_{j\in E}\mathbb{P}(Y=j|A)=\frac{c}{\mathbb{P}(A)}.$$

□

Następujący algorytm podaje przepis na generowanie liczby losowej X z funkcją prawdopdobieństwa $(p_j, j\in E)$.

Algorytm RM-d

```
%zadana f. prawd. (p(j))
1. generuj Y z funkcją prawdopdobieństwa (p'(j));
2. jeśli c rand}p'(Y)\le p(Y), to podstaw X:=Y;
3. w przeciwnym razie idź do 2.
```

Widzimy, że liczba powtórzeń w pętli ma rozkład TGeo($\mathbb{P}(A)$)=TGeo($1/c$). A więc oczekiwana liczba powtórzeń wynosi $c/(c-1)$, a więc należy wybrać c jak najmniejsze.

Podamy teraz wersję ciągłą metody eliminacji. Przypuśćmy, ze chcemy generować liczbę losową X z gęstością f, podczas gdy znamy algorytm generowania liczby losowej Y z gęstością $g(x)$. Niech $c > 0$ będzie takie, że

$$f(x) \le cg(x), \qquad x \in \mathbb{R}.$$

Oczywiście $c > 1$, jeśli gęstości są różne. Niech $U \sim \mathcal{U}(0,1)$ i definiujemy *obszar akceptacji* przez

$$\{cUg(Y) \le f(Y)\}.$$

Zakładamy, że U i Y są niezależne. Podstawą naszego algorytmu będzie następujący fakt.

Fakt 5.2

$$\mathbb{P}(Y \le \xi | A) = \int_{-\infty}^{\xi} f(y)\, dy, \qquad \textit{dla wszystkich } x$$

oraz $\mathbb{P}(A) = 1/c$

Dowód Mamy

$$\begin{aligned} \mathbb{P}(Y \le \xi | A) &= \frac{\mathbb{P}(Y \le \xi, cUg(Y) \le f(Y))}{\mathbb{P}(A)} \\ &= \frac{\int_{-\infty}^{\xi} \frac{f(y)}{cg(y)}}{\mathbb{P}(A)} \\ &= \int_{-\infty}^{\xi} f(y)\, dy \frac{1}{c\mathbb{P}(A)}. \end{aligned}$$

Przechodząc z $\xi \to \infty$ widzimy, że $\mathbb{P}(A) = 1/c$, co kończy dowód. □

Następujący algorytm podaje przepis na generowanie liczby losowej X z gęstością $f(x)$.

Algorytm RM

```
%zadana gęstość f(x)
1. generuj Y z gęstością g(y);
2. jeśli c* rand * g(Y)<= f(Y) to podstaw X:=Y
3. w przeciwnym razie idź do 2.
```

Przykład 5.3 [Normalna zmienna losowa] Pokażemy, jak skorzystać z metody eliminacji aby wygenerować zmienną losową X o rozkładzie normalnym $\mathcal{N}(0,1)$. Najpierw wygenerujemy zmienną losową Z z gęstościa

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad 0 < x < \infty \, .$$

Niech $g(x) = e^{-x}, x > 0$ będzie gęstością rozkładu wykładniczego Exp(1). Należy znaleźć maksimum funkcji

$$\frac{f(x)}{g(x)} = \frac{2}{\sqrt{2\pi}} e^{x-\frac{x^2}{2}}, \qquad 0 < x < \infty \, ,$$

co się sprowadza do szukania wartości dla której $x - x^2/2$ jest największa w obszarze $x > 0$. Maksimum jest osiągnięte w punkcie 1 co pociąga

$$c = \max \frac{f(x)}{g(x)} = \sqrt{\frac{2e}{\pi}} \approx 1.32 \, .$$

Zauważmy, ze

$$\frac{f(x)}{cg(x)} = e^{x-x^2/2-1/2} = e^{-(x-1)^2/2} \, .$$

Teraz aby wygenerować zmienną losową X o rozkładzie $\mathcal{N}(0,1)$ musimy zauważyć, że należy wpierw wygenerować X z gęstościa $f(x)$ i następnie wylosować znak plus z prawdopodobieństwem $1/2$ i minus z prawdopdobieństwem $1/2$. To prowadzi do następującego algorytmu na generowanie liczby losowej o rozkładzie $\mathcal{N}(0,1)$.

Algorytm RM

```
%  na wyjściu X o rozkladzie N(0,1)
1. generuj Y z gęstością g(y);
2. generuj  I=2*floor(2*rand)-1;
3. jeśli U<= exp(-(Y-1)^2/2) to podstaw X:=Y;
4. w przeciwnym razie idź do 2.
5. Podstaw Z:=I*X;
```

6 Metody specjalne dla N(0,1).

Przedstawimy dwa warianty metody, zwanej Boxa-Mullera, generowania pary niezależnych zmiennyc losowych X_1, X_2 o jednakowym rozkładzie normalnym N(0,1). Pomysł polega na wykorzystaniu współrzędnych biegunowych. Zacznijmy od udowodnienia lematu.

Lemat 6.1 *Niech X_1, X_2 są niezależnymi zmiennymi losowymi o jednakowym rozkładzie normalnym N(0,1). Niech (R, Θ) będzie przedstawieniem punktu (X_1, X_2) we współrzędnych biegunowych. Wtedy R i Θ są niezależne, R ma gęstośc*

$$f(x) = xe^{-x^2/2}1(x > 0) \tag{6.5}$$

oraz Θ ma rozkład jednostajny $U(0, 2\pi)$.

Dowód Oczywiście, (X_1, X_2) można wyrazić z (R, Θ) przez

$$\begin{aligned} x_1 &= r\cos\theta, \\ x_2 &= r\sin\theta \end{aligned}$$

a X_1, X_2 ma gęstość $\exp(-(x_1^2+x_2^2)/2)$. Przejście do współrzędnych biegunowych ma jakobian

$$\left| \begin{array}{cc} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{array} \right| = r$$

a więc (R, Θ) ma gęstość

$$f_{(R,\Theta)}(r,\theta)\frac{1}{2\pi}e^{-\frac{r^2}{2}}r\,drd\theta, \qquad r > 0, 0 < \theta \leq 2\pi.$$

□

Zauważmy, że rozkład zadany wzorem (6.5) nazywa się *rozkładem Raleigha*

.

Pierwsza metoda generowania liczb losowych normalnych jest oparta na następującym fakcie.

Fakt 6.2 *Niech U_1 i U_2 będą niezależnymi zmiennymi loswymi o jednakowym rozkładzie jednostanym $U(0, 1)$. Niech*

$$\begin{aligned} X_1 &= (-2\log U_1)^{1/2}\cos(2\pi U_2), \\ X_2 &= (-2\log U_1)^{1/2}\sin(2\pi U_2)\ . \end{aligned}$$

Wtedy X_1, X_2 są niezależnymi zmiennymi losowymi o jednakowym rozkładzie normalnym N(0,1).

Dowód . Trzeba wygenerować zmienną Θ o rozkładzie jednostajnym (proste) oraz R z gęstościa $e^{-\frac{r^2}{2}}r$, $r > 0$. W tym celu zauważmy, że $Y = R^2$ ma rozkład wykładniczy Exp(1/2) ponieważ

$$\begin{aligned} \mathbb{P}(R^2 > x) &= \mathbb{P}(R > \sqrt{x}) \\ &= \int_{\sqrt{x}}^{\infty} re^{-r^2/2}\, dr = e^{-x/2}, \qquad x > 0. \end{aligned}$$

Zanim napiszemy algorytm musimy zauważyć następujący lemacik.

Lemat 6.3 *Jeśli U ma rozkład jednostajny U(0,1), to zmienna losowa $(-2\log U)^{1/2}$ ma rozkład Raleigha.*

Dowód Ponieważ $(-2\log U)$ ma rozkład wykładniczy Exp(1/2), więc jak wcześniej zauważylismy, pierwiastek z tej zmiennej losowej ma rozkład Raleigha.

Druga metoda opiera się na następującym fakcie.

Fakt 6.4 *Niech*

$$\begin{aligned} Y_1 &= \{-2\log(V_1^2 + V_2^2)\}^{1/2} \frac{V_1}{(V_1^2 + V_2^2)^{1/2}}, \\ Y_1 &= \{-2\log(V_1^2 + V_2^2)\}^{1/2} \frac{V_2}{(V_1^2 + V_2^2)^{1/2}}. \end{aligned}$$

i V_1, V_2 są niezależnymi zmiennymi losowymi o jednakowym rozkładzie jednostajnym U(-1,1). Wtedy

$$(X_1, X_2) =^{\mathrm{d}} ((Y_1, Y_2)|V_1^2 + V_2^2 \leq 1)$$

jest parą niezależnych zmiennych losowych o jednakowym rozkładzie normalnym N(0,1).

Dowód powyższego faktu wynika z następującego lematów. Niech K będzie kołem o promieniu 1.

Lemat 6.5 *Jeśli $(W_1, W_2) \sim U(K)$, i (R, Θ) jest przedstawieniem punktu (W_1, W_2) w współrzędnych biegunowych, to R i Θ są niezależne, R^2 ma rozkład jednostajny U(0,1) oraz Θ ma rozład jednostajny $U(0, 2\pi)$.*

Dowód Mamy

$$W_1 = R\cos\Theta, \qquad W_2 = R\sin\Theta\,.$$

Niech $B = \{a < R \le b, t_1 < \Theta \le t_2\}$. Wtedy

$$\begin{aligned}
\mathbb{P}((W_1, W_2) \in B) &= \frac{1}{\pi}\int_B 1((x,y) \in K)\, dx\, dy \\
&= \frac{1}{\pi}\int_{t_1}^{t_2}\int_a^b r\, dr\, d\theta \\
&= \frac{1}{2\pi}(t_2 - t_1) \times (b^2 - a^2) \\
&= \mathbb{P}(a < R \le b, t_1 < \Theta \le t_2).
\end{aligned}$$

Stąd mamy, ze R i Θ są niezależne, Θ o rozkładzie jednostajnym $\mathrm{U}(0, 2\pi)$ oraz R^2 o rozkładzie jednostajnym U(0,1).

7 Generowanie wektorów losowych $\mathrm{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$.

Przypomnijmy, że wektor losowy $\boldsymbol{X}$ o rozkładzie wielowymiarowym normalnym $\mathrm{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$ ma średnią $\boldsymbol{m}$ oraz macierz kowariancji $\boldsymbol{\Sigma}$ oraz dowolna kombinacja $\sum_{j=1}^n a_j X_j$ ma rozkład normalny. Oczywiście gdy mówimy o n-wymiarowym wektorze normalnym $\boldsymbol{X}$, to $\boldsymbol{m}$ jest n wymiarowym wektorem, $\boldsymbol{\Sigma}$ jest macierzą $n \times n$. Przypomnijmy też, że macierz kowariancji musi być nieujemnie określona. W praktyce tego wykładu będziemy dodatkowo zakładać, że macierz $\boldsymbol{\Sigma}$ jest niesingularna. Wtedy rozkład $\boldsymbol{X}$ ma gęstość zadaną wzorem

$$f(x_1, \ldots, x_n) = \frac{1}{\sqrt{2\pi \det(\boldsymbol{\Sigma})}} e^{-\frac{\sum_{j,k=1}^n c_{ij} x_j x_k}{2}}.$$

gdzie

$$\boldsymbol{C} = (c_{jk})_{j,k=1}^n = \boldsymbol{\Sigma}^{-1}$$

Na przykład dla $n = 2$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

gdzie $\rho = \mathrm{Corr}(X_1, X_2)$, $\sigma_i^2 = \mathrm{Var}\, X_i$ $(i = 1, 2)$. Aby wygerować taki dwumiarowy wektor normalny niech

$$Y_1 = \sigma_1(\sqrt{1 - |\rho|}Z_1 + \sqrt{|\rho|}Z_3, \quad Y_2 = \sigma_2(\sqrt{1 - |\rho|}Z_2 + \mathrm{sign}(\rho)\sqrt{|\rho|}Z_3\,,$$

gdzie Z_1, Z_2 są niezaleznymi zmiennymi losowymi o rozkładzie $N(0, 1)$. Teraz dostaniemy szukany wektor losowy $(X_1, X_2) = (Y_1 + m_1, Y_2 + m_2)$.

Dla dowolnego n, jeśli nie ma odpowiedniej struktury macierzy kowarniacji $\mathbf{\Sigma}$ pozwalającej na znalezienie metody ad hoc polecana jest nastepująca metoda. Niech $\boldsymbol{X} = (X_1, \ldots, X_n)$ oraz $\boldsymbol{A}$ takie $\mathbf{\Sigma} = (\boldsymbol{A})^T \boldsymbol{A}$.

Fakt 7.1 *Jeśli $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie normlanym N(0,1) to*

$$\boldsymbol{X} = \boldsymbol{Z}\boldsymbol{A} + \boldsymbol{m}$$

ma rozkład $N(\boldsymbol{m}, \mathbf{\Sigma})$.

Aby obliczyć 'pierwiastek" $\boldsymbol{A}$ z macierzy kowariancji $\mathbf{\Sigma}$ można skorzystać z gotowych metod numerycznych. Na przykład w MATLABie jest instrukcja chol($\mathbf{\Sigma}$), która używa metody faktoryzacji Cholevskiego.

Uwagi bibliograficzne [3]

Zadania teoretyczne

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

$$\begin{aligned} \mathbb{E}\, Z_i &= 1/\lambda \\ \mathbb{E}\, Z_i^2 &= 2/\lambda^2 \end{aligned}$$

oraz

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

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

7.2 Niech $Z_1 = F^{-1}(U)$, $Z_2 = F^{-1}(1-U)$. Obliczyć korelację $\mathrm{corr}\,(Z_1, Z_2)$ gdy F jest dystrybuantą:
(a) rozkładu Poissona; $\lambda = 1, 2, \ldots, 10$,
(b) rozkładu geometrycznego $\mathrm{Geo}(p)$.

7.3 Niech $F'(x)$ będzie warunkową dystrybuantą zdefiniowana przez

$$F'(x) = \begin{cases} 0 & x < a, \\ \frac{F(x)-F(a)}{F(b)-F(a)}, & a \le x < b, \\ 1 & x \ge b \end{cases}$$

gdzie $a < b$. Pokazać, że $X = F^{-1}(U')$, gdzie

$$U' = F(a) + (F(b) - F(a))U$$

ma dystrybuantę F'.

7.4 Podać procedurę generowania liczb losowych z gęstościa trójkątną Tri(a,b), z gęstością

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

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

7.5 Podać procedurę na generowanie liczby losowej z gęstością postaci

$$f(x) = \sum_{j=1}^{N} a_j x_j, \qquad 0 \le x \le 1,$$

gdzie $a_j \ge 0$. Wsk. Obliczyć gęstość $\max(U_1, \ldots, U_n)$ i następnie użyć metody superpozycji.

7.6 Niech F jest dystrybuantą rozkładu Weibulla $\mathrm{W}(\alpha, c)$,

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

Pokazać, że

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

ma rozkład $\mathrm{W}(\alpha, c)$.

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

7.8 a) Udowodnić, że $\mathbb{P}(\lfloor U^{-1} \rfloor = i) = \frac{1}{i(i+1)}$, dla $i = 1, 2, \ldots$.
b) Podać algorytm i obliczyć prawdopodobieństwo akceptacji na wygenerowanie metodą eliminacji dyskretnej liczby losowej X z funkcją prawdopodobieństwa $\{p_k, \ k = 1, 2, \ldots\}$, gdzie

$$p_k = \frac{6}{\pi^2} \frac{1}{k^2}, \qquad k = 1, 2, \ldots .$$

7.9 Podaj metodę generowania liczby losowej z gęstością:

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

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

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

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

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

Zadania laboratoryjne

7.11 Podać wraz z uzasadnieniem jak generować liczbę losową X o rozkładzie U[a,b]. Napisać funkcję Unif(a,b).

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

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

Napisać procedurę Makeham(A,m,c,x) generowania liczb losowych o takim rozkładzie. Wsk. Zinterpretować probabilistycznie jak operacja prowadzi do faktoryzacji $e^{-At-m(c^{t+x}-c^x)} = e^{-At}e^{-m(c^{t+x}-c^x)}$.

7.13 Napisać procedure Poi(lambda)) generowania liczb losowych o rozkładzie Poi(λ). Uzasadnić jej poprawność.

7.14 *Rozkład beta* Beta(α, β) ma gęstość

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

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 Beta(α, β) metodą eliminacji.

7.15 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. [Wsk. Można użyć instrukcji TIC and TOC.]

Rozdział III

Podstawy metod Monte Carlo

1 Estymatory Monte Carlo

Przypuśćmy, że chcemy obliczyć wartość I o której wiemy, że można przedstawić w postaci

$$I = \mathbb{E}\, Y$$

dla pewnej zmiennej losowej Y, takiej, że $\mathbb{E}\, Y^2 < \infty$. Ponadto wiemy jak generować zmienną losową Y. Zauważmy, że dla zadanego I istnieje wiele zmiennych losowych dla których $I = \mathbb{E}\, Y$, i celem tego rozdziału będzie zrozumienie, jak najlepiej dopbierać Y do konkretnego problemu. Przy okazji zobaczymy jak dobierać liczbę replikacji potrzebną do zagwarantowania żądanej dokładności.

1.1 Przypadek estymacji nieobciązonej

Przypuśćmy, że chcemy obliczyć wartość I o której wiemy, że można przedstawićc w postaci

$$I = \mathbb{E}\, Y$$

dla pewnej liczby losowej Y, takiej, że $\mathbb{E}\, Y^2 < \infty$. W tym podrozdziale będziemy przyjmować, że kolejne replikacje $Y_1, \ldots, Y_n$ będą niezależnymi replikacjami liczby losowej Y i będziemy rozważać *estymatory* szukanej wielkości I postaci

$$\hat{Y}_n = \frac{1}{n} \sum_{j=1}^{n} Y_j.$$

Zauważmy, że estymator $\hat{I}_n$ jest *nieobiążony* t.j.

$$\mathbb{E}\,\hat{Y}_n = Y$$

oraz *mocno zgodny* to znaczy z prawdopodobieństwem 1 zachodzi zbieżność dla $n \to \infty$

$$\hat{Y}_n \to Y$$

co pociąga ważną dla nas słabą zgodność, to znaczy, że dla każdego $b > 0$ mamy

$$\mathbb{P}(|\hat{Y}_n - Y| > b) \to 0.$$

Zauważmy, że warunek $|\hat{Y}_n - Y| \leq b$ oznacza, że błąd bezwzględny nie przekracza b. Stąd widzimy, że dla każdego b i liczby $0 < \alpha < 1$ istnieje n (nas interesuje najmniejsze naturalne) takie, że

$$\mathbb{P}(\hat{Y}_n - b \leq Y \leq \hat{Y}_n + b) \leq 1 - \alpha.$$

A więc z prawdopodobieństwem większym niż $1-\alpha$ szukana wielkość Y należy do przedziału losowego $[\hat{Y}_n - b, \hat{Y}_n + b]$ i dlatego α nazywa się *poziomem istotności* lub *poziomem ufności*. Inne zadania to: dla zadanej liczby replikacji n oraz poziomu ufności α można wyznaczyć błąd b, lub przy zadanym n i b wyznaczyć poziom ufności α. Wyprowadzimy teraz wzór na związek pomiędzy n,b, α. Ponieważ w metodach Monte Carlo n jest zwykle duże, będziemy mogli skorzystać z *centralnego twierdzenia granicznego* (CTG), które mówi, że jeśli tylko $Y_1, \ldots, Y_n$ są niezależne i o jednakowym rozkładzie oraz $\sigma_Y^2 = \operatorname{Var} Y_j < \infty$, to dla $n \to \infty$

$$\mathbb{P}\left(\sum_{j=1}^{n} \frac{Y_j - \mathbb{E}\,Y_1}{\sigma_Y \sqrt{n}} \leq x\right) \to \Phi(x) \tag{1.1}$$

gdzie $\sigma_Y = \sqrt{\operatorname{Var} Y_1}$ oraz $\Phi(x)$ jest dystrybuantą standardowego rozkładu normalnego. Niech $z_{1-\alpha/2}$ będzie $\alpha/2$–kwantylem rozkładu normalnego tj.

$$z_{1-\alpha/2} = \Phi^{-1}(1 - \frac{\alpha}{2}).$$

Stąd mamy, że

$$\lim_{n\to\infty} \mathbb{P}(-z_{1-\alpha/2} < \sum_{j=1}^{n} \frac{Y_j - \mathbb{E}\,Y_1}{\sigma_Y \sqrt{n}} \leq z_{1-\alpha/2}) = 1 - \alpha.$$

A więc

$$\mathbb{P}(-z_{1-\alpha/2} < \sum_{j=1}^{n} \frac{(Y_j - \mathbb{E}\, Y)}{\sigma_Y \sqrt{n}} \leq z_{1-\alpha/2}) \approx 1 - \alpha \, .$$

Przekształcając otrzymujemy równoważną postać

$$\mathbb{P}(\hat{Y}_n - z_{1-\alpha/2} \frac{\sigma_Y}{\sqrt{n}} \leq I \leq \hat{Y}_n + z_{1-\alpha/2} \frac{\sigma_Y}{\sqrt{n}}) \approx 1 - \alpha \, .$$

Stąd mamy *fundamentalny związek*

$$b = z_{1-\alpha/2} \frac{\sigma_Y}{\sqrt{n}} \, .$$

Jesli nie znamy σ_Y^2, a to jest typowa sytuacja, to możemy zastąpić tę wartość przez wariancję próbkową

$$\hat{S}_n^2 = \frac{1}{n-1} \sum_{j=1}^{n} (Y_i - \hat{Y}_n)^2.$$

Estymator $\hat{S}_n^2$ jest nieobciążony, tj. $\mathbb{E}\, \hat{S}_n^2 = \sigma_Y^2$ oraz zgodny. Jest również prawdą, że

$$\mathbb{P}\left(\sum_{j=1}^{n} \frac{(Y_j - \mathbb{E}\, Y)}{\hat{S}_n \sqrt{n}} \leq x \right) \to \Phi(x)$$

A więc

$$\mathbb{P}(-z_{1-\alpha/2} < \sum_{j=1}^{n} \frac{(Y_j - \mathbb{E}\, Y)}{\hat{S}_n \sqrt{n}} \leq z_{1-\alpha/2}) \approx 1 - \alpha \, .$$

Przekształcając otrzymujemy równoważną postać

$$\mathbb{P}(\hat{Y}_n - z_{1-\alpha/2} \frac{\hat{S}_n}{\sqrt{n}} \leq I \leq \hat{Y}_n + z_{1-\alpha/2} \frac{\hat{S}_n}{\sqrt{n}}) \approx 1 - \alpha \, .$$

Stąd możemy napisać związek

$$b = z_{1-\alpha/2} \frac{\hat{S}_n}{\sqrt{n}}$$

W teorii Monte Carlo przyjęło rozważać się $\alpha = 0.05$ skąd $z_{1-0.025} = 1.96$ i tak będziemy robić w dalszej części wykładu. Jednakże zauważmy, że istnieje

mozliwość, na przykład na poziomie 99%. Wtedy odpowiadająca $z_{1-0.005} = 2.58$. Polecamy czytelnikowi obliczenie odpowiednich kwantyli dla innych poziomów istotności korzystając z tablic rozkładu normalnego z Dodatku ??? Mamy więc fundamentalny związek na poziomie istotności 0.05

$$b = \frac{1.96\sigma_Y}{\sqrt{n}}. \tag{1.2}$$

A więc aby osiągnąć dokładność b należy przeprowadzić liczbę $1.96^2\sigma_Y^2/b^2$ prób.

Jak już wspomnieliśmy, niestety w praktyce niedysponujemy znajomością σ_Y. Estymujemy więc σ_Y z wzoru:

$$\hat{S}_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_i - \hat{Y}_n)^2.$$

Można to zrobić podczas pilotażowej symulacji, przed zasadniczą symulacją mającą na celu obliczenie Y. Wtedy powinniśmy napisać wynik symulacji w postaci przedziału ufności

$$(\hat{Y}_n - \frac{z_{1-\alpha/2}\hat{S}_n}{\sqrt{n}}, \hat{Y}_n + \frac{z_{1-\alpha/2}\hat{S}_n}{\sqrt{n}})$$

lub w formie $(\hat{Y}_n \pm \frac{z_{1-\alpha/2}\hat{\sigma}_n}{\sqrt{n}})$.

Podamy teraz wzór rekurencjyjny obliczanie $\hat{Y}_n$ oraz $\hat{S}_n^2$.

Algorytm

$$\hat{Y}_n = \frac{n-1}{n}\hat{Y}_{n-1} + \frac{1}{n}Y_n$$

i

$$\hat{S}_n^2 = \frac{n-2}{n-1}\hat{S}_{n-1}^2 + \frac{1}{n}(Y_n - \hat{Y}_{n-1})^2$$

dla $n = 2, 3, \ldots$ i

$$\hat{Y}_1 = Y_1 \qquad \hat{S}_1^2 = 0\ .$$

Przykład 1.1 [Asmussen & Glynn] Przypuśćmy, że chcemy obliczyć $I = \pi = 3.1415\ldots$ metodą Monte Carlo. Niech $Y = 4\mathbf{1}(U_1^2 + U_2^2 \leq 1)$, gdzie U_1, U_2 są niezależnymi jednostajnymi U(0,1). Wtedy $\mathbf{1}(U_1^2 + U_2 \leq 1)$ jest zmienną Bernoulliego z średnią $\pi/4$ oraz wariancją $\pi/4(1-\pi/4)$. Tak więc

$$\mathbb{E}\,Y = \pi$$

oraz wariancja Y równa się

$$\operatorname{Var} Y = 16\pi/4(1-\pi/4) = 2.6968.$$

Tak więc aby przy zadanym poziomie istotności dostać jedną cyfrę wiodącą, tj. 3 należy zrobić $n \approx 1.96^2 \cdot 2.7^2/0.5^2 = 112$ replikacji, podczas gdy aby dostać drugą cyfrę wiodąca to musimy zrobić $n \approx 1.96^2 \cdot 2.7^2/0.05^2 = 11\ 200$ replikacji, itd.

Wyobraźmy sobie, że chcemy porównać kilka eksperymentów. Poprzednio nie braliśmy pod uwagę czasu potrzebnego na wygenerowanie jednej replikacji Y_j. Przyjmijmy, że czas można zdyskretyzować i $n = 0, 1, 2, \ldots$, i że dany jest łączny czas m (zwany *budżetem*) którym dysponujemy dla wykonania całego eksperymentu. Niech średnio jedna replikacja trwa czas t. A więc średnio można wykonąc w ramach danego budżetu jedynie m/t replikacji. A więc mamy

$$\hat{Y}_n = \frac{1}{n}\sum_{j=1}^{n} Y_j \approx \frac{1}{m/t}\sum_{j=1}^{m/t} Y_j.$$

Aby porównać ten eksperyment z innym, w którym posługujemy się liczbami losowymi $Y_1', \ldots$, takimi, że $Y = \mathbb{E}\, Y'$, gdzie na wykonanie jednej replikacji Y_j' potrzebujemy średnio t' jednostek czasu, musimy porównać wariancje

$$\operatorname{Var} \hat{Y}_n, \qquad \operatorname{Var} \hat{Y}_n'.$$

Rozpisując mamy odpowiedno $(t\operatorname{Var} Y)/m$ i $t'\operatorname{Var} Y')/m$. W celu wybrania lepszego estymatora musimy więc porównywać wielkości $t\operatorname{Var} Y$, którą zwiemy *efektywnością* estymatora.

Zgrubna metoda Monte Carlo Przypuśćmy, że szukaną wielkość możemy zapisać jako

$$\int_B k(x)\, dx,$$

gdzie $|B| < \infty$. Wtedy przyjmujemy $Y = k(V)$, gdzie V ma rozkład jednostajny na B. Niech $Y_1, \ldots, Y_n$ będą niezależnymi zmiennymi losowymi. Wtedy

$$\hat{Y}^{\mathrm{CMC}} = \hat{Y}_n^{\mathrm{CMC}} = \frac{1}{n}\sum_{j=1}^{n} Y_j$$

jest *zgrubnym estymatorem Monte Carlo*(crude Monte Carlo) i stąd skrót CMC.

Przykład 1.2 Niech $k : [0,1] \to \mathbb{R}$ będzie funkcją całkowalną z kwadratem $\int_0^1 k^2(x)\,dx < \infty$. Niech $Y = k(U)$. Wtedy $\mathbb{E}\,Y = \int_0^1 k(x)\,dx$ oraz $\mathbb{E}\,Y^2 = \int_0^1 k^2(x)\,dx$. Stąd

$$\sigma_Y = \sqrt{\int_0^1 k^2(x)\,dx - (\int_0^1 k(x)\,dx)^2}.$$

W szczególności rozważmy

$$k(x) = \frac{e^x - 1}{e - 1}.$$

Wtedy oczywiście mamy

$$\begin{aligned} \int_0^1 k(x)\,dx &= \frac{1}{e-1}(\int_0^1 e^x\,dx - 1) = \frac{e-2}{e-1} = 0.418 \\ \int_0^1 k^2(x)\,dx &= ... \end{aligned}$$

Stąd $\sigma_Y^2 = 0.286^2 = 0.0818$ a więc $n = 0.31423/b^2$.

Metoda chybił–trafił. Zwróćmy uwagę, że do estymacji możemy używać innych estymatorów. Na przykład często proponuje się estymator *chybił-trafił* (Hit 0r Miss) (HoM). Pomysł polega na przedstawieniu $I = \mathbb{E}\,Y$, gdzie

$$Y = \mathbf{1}(k(U_1) > U_2).$$

Jest bowiem oczywiste, że

$$\mathbb{E}\,Y = \int_0^1 \int_{k(u_2)>u_1} du_1\,du_2 = \int_0^1 k(u_2)\,du_2.$$

Wtedy dla $U_1, \ldots, U_{2n}$ obliczamy replikacje $Y_1, \ldots, Y_n$ skąd obliczamy estymator

$$\hat{Y}_n^{\mathrm{HoM}} = \frac{1}{n}\sum_{j=1}^{n} Y_j.$$

zwanym estymatorem *chybił trafił*. Wracając do przykładu 1.2 możemy łatwo policzyć wariancję $\operatorname{Var} Y = I(1-I) = 0.2433$. Stąd mamy, że $n = 0.9346/b^2$, czyli dla osiągnięcia tej samej dokładności trzeba około 3 razy więcej replikacji Y-tów. W praktyce też na wylosowanie Z w metodzie chybił trafił potrzeba dwa razy więcej liczb losowych.

Jak widać celem jest szukanie estymatorów z bardzo małą wariancją.

Przypadek obciążony Przedyskutujemy teraz jak symulowac $\mathbb{E}\, f(Y)$, jeśli wiadomo jak generować replikację Y. Zwróćmy uwagę, że $\hat{Z}_n = \frac{1}{n}\sum_{j=1}^{n} f(Y_j)$ nie musi być nieobciążony. Niech więc

$$\delta_n = \mathbb{E}\, \hat{Z}_n - \mathbb{E}\, f(Y)$$

będzie obciążeniem oraz są spełnione warunki

Rozdział IV

Techniki redukcji wariancji

W tym rozdziale zajmiemy się konstrukcją i analizą metod prowadzących do redukcji estymatora wariancji. Jak zauważyliśmy w porzednim rozdziale, metoda MC polega na tym, że aby obliczyć powiedzmy I znajdujemy taką zmienną losową Y, że $I = \mathbb{E}\, Y$. Wtedy $\hat{Y} = \sum_{j=1}^{n} Y_j/n$ jest estymatorem MC dla I, gdzie $Y_1, \ldots$ są niezależnymi replikacjami Y. Wariancja $\operatorname{Var} \hat{Y} = \operatorname{Var} Y/n$, a więc widzimy, że musimy szukać takiej zmiennej losowej Y dla której $\mathbb{E}\, Y = I$, i mającej możliwie mała wariancję. Ponadto musimy wziąść pod uwagę możliwość łatwego generowania repliacji Y. W tym rozdziale zajmiemy się następującymi technikami:

1. Metoda warstw
2. Metoda zmiennych antytetycznych
3. Metoda wspólnych liczb losowych
4. Metoda zmiennych kontrolnych
5. Warunkowa metoda MC
6. Metoda losowania istotnościowego

1 Metoda warstw

Jak zwykle chcemy obliczyć $I = \mathbb{E}\, Y$. Przestrzeń zdarzeń Ω rozbijamy na m *warstw* $\Omega_1, \ldots, \Omega_m$. Celem będzie redukcja zmienności na warstwach jak tylko to będzie możliwe. Dobrze będzie rozpatrzyć następujący przykład.

Przykład 1.1 Jeśli chcemy obliczyć metodą Monte Carlo całkę $\mathbb{E}\, g(U) = \int_0^1 g(x)\, dx$ to możemy wprowadzić warstwy $\Omega_j = \{\omega : (j-1)/m \le U < j/m\}$. W tym przypadku znamy $p_j = \mathbb{P}(\Omega_j) = 1/m$.

Niech $J(\omega) = j$ jeśli $\omega \in \Omega_j$. Niech Y_j będzie zmienną losową o rozkładzie $Y|J = j$, tj.

$$\mathbb{P}(Y_j \in B) = \mathbb{P}(Y \in B|J = j) = \frac{\mathbb{P}(\{Y \in B\} \cap \Omega_j)}{\mathbb{P}(\Omega_j)}.$$

Zakładamy niezależność $(Y_j)_j$ i J. Wtedy $\mathbb{E}\, Y_J = I$.

Oznaczmy prawdopodobieństwo wylosowanie warstwy $p_j = \mathbb{P}(\Omega_j)$ (to przyjmujemy za znane) oraz $y_j = \mathbb{E}\, Y_j$ (oczywiście to jest nieznane bo inaczej nie trzeba by nic obliczać). Ze wzoru na prawdopodobieństwo całkowite

$$\mathbb{E}\, Y = p_1 y_1 + \ldots + p_m y_m\ .$$

Niech teraz n będzie ogólną liczbą replikacji, którą podzielimy na odpowiednio n_j replikacji Y_j w warstwie j. Te replikacje w warstiwe j są $Y_1^j, \ldots, Y_{n_j}^j$. Oczywiście $n = n_1 + \ldots + n_m$. Niech $\hat{Y}^j = \frac{1}{n_j}\sum_{i=1}^{n_j} Y_i^j$ będzie estymatorem y_j i definiujemy estymator warstwowy

$$\hat{Y}^{\mathrm{str}} = p_1 \hat{Y}^1 + \ldots + p_m \hat{Y}^m\ .$$

Zauważmy, że tutaj nie piszemy indeksu z liczbą replikacji. Przypuścmy, ze wiadomo jak generowac poszczególne Y_j. Niech $\sigma_j = \mathrm{Var}\, Y_j$ będzie wariancją warstwy j. Zauważmy, że estymator warstwowy jest nieobciążony

$$\begin{aligned}\mathbb{E}\, \hat{Y}^{\mathrm{str}} &= p_1 \mathbb{E}\, \hat{Y}^1 + \ldots + p_m \mathbb{E}\, \hat{Y}^m \\ &= \sum_{i=1}^{m} p_i \frac{\mathbb{E}\, Y \mathbf{1}_{\Omega_i}}{p_i} = \mathbb{E}\, Y = I.\end{aligned}$$

Wtedy wariancja estymatora warstwowego (pamiętamy o niezależności poszczególnych replikacji)

$$\hat{\sigma}^2_{\mathrm{str}} = n\mathrm{Var}\, \hat{Y}^{\mathrm{str}} = p_1^2 \frac{n}{n_1}\sigma_1^2 + \ldots + p_m^2 \frac{n}{n_m}\sigma_m^2\ .$$

Zastanowimy się co się stanie gdy z $n \to \infty$ tak aby n_j/n były zbieżne do powiedzmy q_j.

Fakt 1.2

$$\frac{\sqrt{n}}{\hat{\sigma}^2_{\rm str}}(\hat{Y}^{\rm str} - I) \stackrel{\rm d}{\to} N(0,1) \; .$$

Dowód Zauważmy, że

$$\frac{\sqrt{n_j}}{\sigma_j}\left(\hat{Y}^j - y_j\right) \stackrel{\rm d}{\to} {\rm N}(0,1) \; .$$

Teraz

$$\begin{aligned}\sqrt{n}(\sum_{j=1}^m p_j\hat{Y}^j - I) &= \sqrt{n}(\sum_{j=1}^m p_j\hat{Y}^j - \sum_{j=1}^m p_j y_j) \\ &= \sqrt{n}\sum_{j=1}^m \frac{\sigma_j}{\sqrt{n_j}}\left[\frac{\sqrt{n_j}}{\sigma_j}\left(\hat{Y}^j - y_j\right)\right] \\ &\stackrel{\rm d}{\to} {\rm N}(0, \sum_{j=1}^m p_j^2 \frac{\sigma_j^2}{(n_j/n)}).\end{aligned}$$

Przyjmiemy teraz, że $q_j = p_j$. Wtedy

$$\lim_{n\to\infty} \hat{\sigma}^2_{\rm str} = \sum_{j=1}^m p_j\sigma_j^2 \; .$$

Z drugiej strony pamiętając, że $Y =_{\rm d} Y_J$

$${\rm Var}\, Y = {\rm Var}\, Y_J = \sum_{j=1}^m p_j \mathbb{E}\,(Y_j - I)^2 \geq \sum_{j=1}^m p_j \mathbb{E}\,(Y_j - y_j)^2$$

ponieważ funkcja od x równa $\mathbb{E}\,(Y_j - x)^2$ jest minimalna dla $x = \mathbb{E}\, Y_j$.

Okazuje się, że optymalny dobór liczby replikacji n_j (to znaczy na warstwach mozna wyliczyć w nastepującym twierdzeniu.

Twierdzenie 1.3 *Niech n będzie ustalone i $n_1 + \ldots + n_m = n$. Wariancja $p_1^2\frac{n}{n_1}\sigma_1^2 + \ldots + p_m^2\frac{n}{n_m}\sigma_m^2$ estymator warstwowego $\hat{Y}^{\rm str}$ jest minimalna gdy*

$$n_j = \frac{p_j\sigma_j}{\sum_{j=1}^m p_j\sigma_j} n \; .$$

Dowód Musimy zminimizować

$$V(n_1,\ldots,n_m) = p_1^2\frac{n}{n_1}\sigma_1^2 + \ldots + p_m^2\frac{n}{n_m}\sigma_m^2$$

przy warunku $n = n_1 + \ldots + n_m$ i $n_j \geq 0$. Jeśli podstawimy

$$n_j = \frac{p_j\sigma_j}{D}$$

gdzie $D = n\sum_{j=1}^m p_j\sigma_j$ to $V(n_1,\ldots,n_m) = nD^2$. Ale z nierówności Cauchy-Szwarca

$$\begin{aligned} nD^2 &= \frac{1}{n}\left(\sum_{j=1}^m p_j\sigma_j\right)^2 \\ &= \frac{1}{n}\left(\sum_{j=1}^m \sqrt{n_j}(p_j\sigma_j/\sqrt{n_j})\right)^2 \\ &\leq \frac{1}{n}\left(\sum_{j=1}^m n_j\right)\left(\sum_{j=1}^m (p_j^2\sigma_j^2/n_j)\right)^2 . \end{aligned}$$

Niestety ten wzór jest mało użyteczny ponieważ nie znamy zazwyczaj wariancji σ_j, $j = 1,\ldots,m$.

Przykład 1.4 Obliczenie prawdopodobieństwa awarii sieci. Sieć jest spójnym grafem składającym się z węzłów, z których dwa z s,t są wyróżnione. Węzły są połaczone krawędziami ponumerowanymi $1,\ldots,n$. Każda krawędź może być albo sprawna – 0 lub nie – 1. W przypadku niesprawnej krawędzi będziemy mówili o przerwaniu. A więc stan krawędzi opisuje wektor $e_1,\ldots,e_n$, gdzie $e_i = 0$ lub 1. Sieć w całości może być sprawna – 0 lub nie (t.j. z przerwą) – 1, jeśli jest połączenie od s do t. Formalnie możemy opisać stan systemu za pomocą funkcji $k : \mathcal{S} \to \{0,1\}$ określonej na podzbiorach $1,\ldots,n$ Mówimy, że sieć jest niesprawna i oznaczamy to przez 1 jeśli nie ma połaczenia pomiędzy węzłami s i t, a 0 w przeciwnym razie. To czy sieć jest niesprawna (lub sprawna) decyduje podzbiór B niesprawnych krawędzi. Podamy teraz przykłady. Będziemy przez $|B|$ oznaczać liczbę elementów zbioru B, gdzie $B \subset \{1,\ldots,n\}$.

- *Układ szeregowy* jest niesprawny gdy jakakolwiek krawędź ma przerwanie. A więc $k(B) = 1$ wtedy i tylko wtedy gdy $|B| \geq 1$. przykłady.

- *Układ równoległy* jest niesprawny jeśli wszystkie elementy są niesprawne. A więc $(B) = 1$ wtedy i tylko wtedy gdy $|B| = n$.

Teraz przypuśćmy, że każda krawędź może mieć przerwanie z prawdopodobieństwem q niezależnie od stanu innych krawędzi. Niech X będzie zbiorem (losowym) krawędzi z przerwaniami, tzn. $X \subset \{1, \ldots, n\}$. Zauważmy, że X jest dyskretną zmienną losowa z funkcją prawdopodobieństwa

$$\mathbb{P}(X = B) = p(B) = q^{|B|}(1-q)^{n-|B|},$$

dla $B \subset \{1, \ldots, n\}$. Naszym celem jest obliczenie prawdopodobieństwa, że sieć jest niesprawna. A więc

$$p = \mathbb{E}\, k(X) = \sum_{B \subset \{1,\ldots,n\}} k(B)p(B)$$

jest szukanym prawdopodobieństwem, że sieć jest niesprawna. Z powyższego widzimy, że możemy używać estymatora dla p postaci

$$\hat{X}_n = \frac{1}{n}\sum_{i=1}^{n} k(X_i),$$

gdzie $X_1, \ldots, X_n$ są niezależnymi replikacjami X. Ponieważ typowo prawdopodobieństwo I jest bardzo małe, więc aby je obliczyć musimy zrobić wiele replikacji. Dlatego jest bardzo ważne optymalne skonstruowanie dobrego estymatora dla I. Można zaproponować estymator warstwowy, gdzie j-tą warstwą $\mathcal{S}(j)$ jest pozdbiór $\mathcal{S}$ mający dokładnie j boków. Zauważmy, że

$$p_j = \mathbb{P}(X \in \mathcal{S}(j)) = \mathbb{P}(\text{dokładnie } j \text{ boków ma przerwę}) = \binom{|\mathcal{S}|}{j} q^j (1-q)^{|\mathcal{S}|-j}\ .$$

Każdy $A \subset \mathcal{S}$ ma jednkowe prawdopodobieństwo $q^j(1-q)^{|\mathcal{S}|-j}$ oraz

$$\mathbb{P}(X = A | X \in \mathcal{S}(j)) = 1/\binom{|\mathcal{S}|}{j} \qquad A \subset \mathcal{S}\ .$$

Aby więc otrzymać estymator warstwowy generujemy np_j replikacji na każdej warstwie $\mathcal{S}(j)$ z rozkładem jednostajnym.

1.1 Nierówność

Twierdzenie 1.5 *Jesli $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi, i ϕ oraz ψ funkcjami niemalejącymi n zmiennych, to*

$$\mathbb{E}\left[\phi(X_1, \ldots, X_n)\psi(X_1, \ldots, X_n)\right] \geq \mathbb{E}\left[\phi(X_1, \ldots, X_n)\right]\mathbb{E}\left[\psi(X_1, \ldots, X_n)\right]$$

przy założeniu że wartości oczekiwane są skończone.

Dowód Dowód jest indukcyjne względem liczby zmiennych $n = 1, 2, \ldots$. Niech $n = 1$. Wtedy

$$(\psi(x) - \psi(y))(\phi(x) - \phi(y)) \geq 0,$$

skąd dla niezależnych zmiennych losowych X, Y o jednakowych rozkładach jak X_1

$$\mathbb{E}\,(\psi(X) - \psi(Y))(\phi(X) - \phi(Y)) \geq 0$$

Teraz rozpisujemy

$$\mathbb{E}\,\psi(X)\phi(X) - \mathbb{E}\,\psi(X)\phi(Y) - \mathbb{E}\,\psi(Y)\phi(X) + \mathbb{E}\,\psi(Y)\phi(Y)\ .$$

Korzystając, że X i Y są niezalezne o jednakowym rozkładzie jak X_1 mamy

$$\mathbb{E}\left[\phi(X_1)\psi(X_1)\right] \geq \mathbb{E}\left[\phi(X_1)\right]\mathbb{E}\left[\psi(X_1)\right]\ .$$

Teraz zakładamy, że twierdzenie jest prawdziwe dla $n - 1$. Stąd

$$\mathbb{E}\left[\phi(X_1, \ldots, X_{n-1}, x_n)\psi(X_1, \ldots, X_{n-1}, x_n)\right] \geq \mathbb{E}\left[\phi(X_1, \ldots, X_{n-1}, x_n)\right]\mathbb{E}\left[\psi(X_1, \ldots, X_{n-1}, x_n)\right],$$

dla wszystkich x_n. Teraz podstawiając $x_n = X_n$ mamy

$$\mathbb{E}\left[\phi(X_1, \ldots, X_{n-1}, X_n)\psi(X_1, \ldots, X_{n-1}, X_n)\right] \geq \mathbb{E}\left[\phi(X_1, \ldots, X_{n-1}, X_n)\right]\mathbb{E}\left[\psi(X_1, \ldots, X_{n-1}, X_n)\right],$$

□

Wniosek 1.6 *Jesli funkcja $\psi(x_1, \ldots, x_n,\ 0 \leq x_1, \ldots, x_n$ jest monotoniczną każdego argumentu, to dla niezależnych zmiennych $U_1, \ldots, U_n$ o rozkładzie jednostajnym*

$$\mathrm{Cov}\left(\psi(U_1, \ldots, U_n), \psi(1 - U_1, \ldots, 1 - U_n)\right) \leq 0\ .$$

1.2 Zmienne antytetyczne

Rozpatrzmy teraz $2n$ zmiennych losowych $Y_1, Y_2, \ldots, Y_{2n-1}, Y_{2n}$, takich, że

- pary $(Y_i, Y_{i+1})_i$, $i = 1, \ldots, n$ są niezależne o jednakowym rozkładzie,
- Zmienne losowe $(Y_j)_{j=1}^{2n}$ mają ten sam rozkład brzegowy.

A więc jesli rozważamy zgrubny estymator $\hat{Y}^{\mathrm{CMC}}$ generowany przez niezależne replikacje to wariancja wynosi $\mathrm{Var}\, Y/n$ Zauważmy, że jesli będziemy rozpatrywać estymator

$$\hat{Y}^2_{\mathrm{anth}} = \frac{\sum_{j=1}^{2n} Y_j}{2n}$$

to jego wariancja

$$\begin{aligned}
\frac{\mathrm{Var}\,(\hat{Y}^2_{\mathrm{anth}})}{2n} &= \frac{\mathrm{Var}\,(\frac{Y_1+Y_2}{2})}{n} \\
&= \frac{1}{4n}(2\mathrm{Var}\,(Y) + 2\mathrm{cov}(Y_1, Y_2)) \\
&= \frac{1}{4n}(2\mathrm{Var}\,(Y) + 2\mathrm{Var}\,(Y_1)\mathrm{Corr}(Y_1, Y_2)) \\
&= \frac{1}{2n}\mathrm{Var}\,(Y)(1 + \mathrm{Corr}(Y_1, Y_2))\ .
\end{aligned}$$

A więc jeśli $\mathrm{Corr}(Y_1, Y_2))$ jest ujemne to redukujemy wariancję.

1.3 Wspólne liczby losowe

Przypuśćmy, że chcemy sprawdzić różnicę $d = \mathbb{E}\, k_1(X) - \mathbb{E}\, k_2(X)$. Można to zrobic na dwa sposoby. Najpierw estymujemy $\mathbb{E}\, k_1(X)$ a następnie, niezależnie $\mathbb{E}\, k_2(X)$ i wtedy liczymy d. To nie zawsze jest optymalne, szczególnie gdy różnica d jest bardzo mała. Inny sposób jest zapisanie $d = \mathbb{E}\,(k_1(X) - k_2(X)$ i następnie estymowania d. W pierwszym przypadku, $Y_1 = k_1(X_1)$ jest niezależne od $Y_2 = k_2(X_2)$, gdzie X_1, X_2 są niezależnymi replikacjami X. A więc wariancja $\mathrm{Var}\,(k_1(X_1) - k_2(X_2)) = \mathrm{Var}\,(k_1(X)) + \mathrm{Var}\,(k_2(X)$. W drugim przypadku $\mathrm{Var}\,(k_1(X) - k_2(X)) = \mathrm{Var}\,(k_1(X)) + \mathrm{Var}\,(k_2(X) - 2\mathrm{Cov}\,(k_1(X), k_2(X))$. Jesli więc $\mathrm{Cov}\,(k_1(X), k_2(X)) > 0$ to istotnie zmienjszymy wariancję. Tak będzie w przypadku gdy obie funkcje są monotoniczne. Rozpatrzmy następujący przykład. Niech $k_1(x) = e^{\sqrt{x}}$ i $k_2(x) = (1 + \frac{\sqrt{x}}{50})^{50}$. Można oszacować, że $|k_1(x) - k_2(x)| \leq 0.03$, a więc $d \leq 0.03$. Przy piewszej

metodzie wariancja Y_1 wynosi $0.195 + 0.188 = 0.383$ natomiast przy drugiej że jest mniejsza od 0.0009.

Rozpatrzmy teraz następujący problem. Przypuśćmy, że mamy m zadań do wykonania. Czasy zadań są niezależnymi zmiennymi losowymi o jednakowym rozkładzie z dystrybuantą F. Mamy do dyspozycji c serwerów. Możemy te zadania wykonać na dwa sposoby. Albo według najdłuższego zadania (Longest Processing Time First - LPTF) , lub najkrótszego zadania (Shortest Processing Time First - SPTF). Decyzje podejmuje się w momentach zakończenia poszczególnych zadań. Celem jest policzenie $\mathbb{E}\, C^{\mathrm{SPTF}} - \mathbb{E}\, C^{\mathrm{LPTF}}$, gdzie C jest czasem do zakończnia ostatniego zadania.

Przykład Powiedzmy, że $N = 5$ i $c = 2$ oraz zadania są wielkości $3, 1, 2, 4, 5$. Wtedy postępując wg SPTF zaczynamy z zadaniami wielkości 1 i 2 i po jednej jednostce czasu resztowe wielkości tych zadań są $3, 0, 1, 4, 5$ a więc mamy teraz zadanie z $N = 4$ i $c = 2$. Postępując tak dalej widzimy, że $C^{\mathrm{SPTF}} = 9$. postępując wg LPTF zaczynamy z zadaniami wielkości 4 i 5 i po 4-ch jednostkach czasu resztowe wielkości tych zadań są $3, 1, 2, 0, 1$ a więc mamy teraz zadanie z $N = 4$ i $c = 2$. Postępując tak dalej widzimy, że $C^{\mathrm{LPTF}} = 8$. Przypuśćmy, że rozmiary zadań są niezależne o jednakowym rozkładzie wykładniczym.

2 Metoda zmiennych kontrolnych

Zacznijmy, że jak zwykle aby obliczyć I znajdujemy taką zmienną losową Y, że $I = \mathbb{E}\, Y$. Wtedy $\hat{Y}^{\mathrm{CMC}} = \sum_{j=1}^{n} Y_j/n$ jest estymatorem MC dla I, gdzie $Y_1, \ldots, Y_n$ są niezależnymi replikacjami Y. Wariancja tego estymatore wynosi $\mathrm{Var}\, \hat{Y} = \mathrm{Var}\, Y/n$. Teraz przypuśćmy, że symulujemy jednocześnie drugą zmienną losową X (więc mamy niezależne replikacje par $(Y_1, X_2), \ldots, (Y_n, X_n)$) dla której znamy $\mathbb{E}\, X$. A więc mamy niezależne replikacje par $(Y_1, X_2), \ldots, (Y_n, X_n))$ Niech $\hat{X} = \sum_{j=1}^{n} X_j/n$ i definiujemy

$$\hat{Y}^{CV} = \hat{Y}^{\mathrm{CMC}} + c(\hat{X} - \mathbb{E}\, X).$$

Obliczmy wariancję nowego estymatora:

$$\mathrm{Var}\, (\hat{Y}^{CV}) = \mathrm{Var}\, (Y + cX)/n\ .$$

Wtedy $\mathrm{Var}\,(\hat{Y}^{CV})$ jest minimalna gdy c,

$$c = -\frac{\mathrm{Cov}\,(Y,X)}{\mathrm{Var}\,(X)}$$

i wtedy minimalna wariancja wynosi $\mathrm{Var}\,(Y)(1-\rho^2)/n$, gdzie $\rho = \mathrm{corr}\,(Y,X)$. A więc wariancja jest zredukowana przez czynnik $1 - rho^2$)

W praktyce nie znamy $\mathrm{Cov}\,(Y,X)$ ani $\mathrm{Cov}\,(X)$ i musimy te wielkości symulacjć stosując wzory

$$\hat{c} = -\frac{\hat{S}^2_{YX}}{\hat{S}^2_X}$$

gdzie

$$\begin{aligned} \hat{S}^2_X &= \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \hat{Y})^2 \\ \hat{S}^2_{YX} &= \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \hat{Y})(X_j - \hat{X}). \end{aligned}$$

Przykład 2.1 Kontynuujemy przykład III.1.1. Liczyliśmy tam π za pomocą estymatora $Y = 4\ \mathbf{1}(U_1^2 + U_2^2 \le 1)$. Jako zmienną kontrolną weźmiemy $X = \mathbf{1}(U_1 + U_2 \le 1)$, które jest mocno skorelowane z Y oraz $\mathbb{E}\,X = 1/2$. Możemy policzyć $1 - \rho^2 = 0.727$. Teraz weźmiemy inną zmienną kontrolną $X = \mathbf{1}(U_1 + U_2 > \sqrt{2})$. Wtedy $\mathbb{E}\,X = (2-\sqrt{2})^2/2$ i $1 - \rho^2 = 0.242$. Zauważmy (to jest zadanie) że jeśli weźmiemy za zmienną kontrolną X lub $a + bX$ to otrzymamy taką samą redukcję wariancji.

Przykład 2.2 Chcemy obliczyć całkę $\int_0^1 g(x)\,dx$ metodą MC. Możemy położyć $Y = g(U)$ i jako kontrolną zmienną przyjąć $X = f(U)$ jeśli znamy $\int_0^1 f(x)\,dx$.

3 Warunkowa metoda MC

Pomysłem jest zastąpienia estymatora Y takiego $I = \mathbb{E}\,Y$ esymatorem $\phi(X)$ $\phi(x) = \mathbb{E}\,[Y|X = x]$. Mamy bowiem

$$\mathbb{E}\,\phi(X) = \mathbb{E}\,Y$$

oraz

$$\mathrm{Var}\,(Y) = \mathbb{E}\,\phi(X) + \mathrm{Var}\,\phi(X)\ ,$$

skąd mamy

$$\mathrm{Var}\,(Y) \geq \mathrm{Var}\,\phi(X)\ .$$

Mianowicie, jeśli X jest wektorem z gęstością f_X, to

$$\mathrm{Var}\,(\phi(X)) = \int (\mathbb{E}\,[Y|X=x])^2\, f_X(x)\,dx - I^2$$

i

$$\mathbb{E}\,[Y^2|X=x] \geq (\mathbb{E}\,[Y|X=x])^2\ .$$

Stąd

$$\begin{aligned}\mathbb{E}\,Y^2 &= \int \mathbb{E}\,[Y^2|X=x]\, f_X(x)\,dx \\ &\geq \int (\mathbb{E}\,[Y|X=x])^2\, f_X(x)\,dx.\end{aligned}$$

Estymatorem warunkowym $\hat{Y}^{\mathrm{Cond}}$ nazywamy

$$\hat{Y}^{\mathrm{Cond}} = \sum_{j=1}^{n} Z_j$$

gdzie $Y_1^{\mathrm{cond}}, \ldots, Y_n^{\mathrm{cond}}$ są replikacjami $Y^{\mathrm{cond}} = \phi(X)$.

Przykład 3.1 Kontynuujemy przykład III.1.1. Liczyliśmy tam π za pomocą estymatora $Y = 4\mathbf{1}(U_1^2 + U_2^2 \leq 1)$. Przyjmując za $X = U_1$ mamy

$$\phi(x) = \mathbb{E}\,[Y|X=x] = 4\sqrt{1-x^2}.$$

Estymator zdefiniowany przez $Y^{\mathrm{Cond}} = 4\sqrt{1-U_1^2}$ redukuje wariancję 3.38 razy, ponieważ jak zostało podane w przykładzie 4.1, mamy $\mathrm{Var}\,(Y^{\mathrm{Cond}}) = 0.797$ a w przykładzie III.1.1 policzyliśmy, że $\mathrm{Var}\,(Y) = 2.6968$.

Przykład 3.2 Niech $I = \mathbb{P}(\xi_1 + \xi_2 > y)$, gdzie ξ_1, ξ_2 są niezależnymi zmiennymi losowymi z jednakową dystrybuantą F, która jest znana, a dla której nie mamy analitycznego wzoru na F^{*2}. Oczywiście za Y możemy wziąść

$$Y = \mathbf{1}(\xi_1 + \xi_2 > y)$$

a za $X = \xi_1$. Wtedy

$$Y^{\mathrm{Cond}} = \overline{F}(y - \xi_1)$$

Przykład 3.3 Będziemy w pewnym sensie kontynuować rozważania z przykładu 3.2. Niech $S_n = \xi_1 + \ldots + \xi_n$, gdzie $\xi_1, \xi_2, \ldots, \xi_n$ są niezależnymi zmiennymi losowymi z jednakową gęstością f. Estymator warunkowy można też używać do estymowania gęstości $f^{*n}(x)$ dystrybuanty $F^{*n}(x)$, gdzie jest dana jawna postać f. Wtedy za estymator warunkowy dla $f^{*n}(y)$ można wziąść

$$Y^{\text{cond}} = f(y - S_{n-1}) \, .$$

4 Losowanie istotnościowe

Przypuścmy, że chcemy obliczyć I, które możemy przedstawić w postaci

$$I = \int_E k(x) f(x) \, dx = \mathbb{E}\, k(Z)$$

dla funkcji $k(x)$ takiej, że $\int (k(x))^2 f(x) \, dx < \infty$. Zakładamy, że Z określona na przestrzeni probabilistycznej $\Omega, \mathcal{F}, \mathbb{P})$ jest o wartościach w E. Zauważmy, że gęstość jest określona na E, gdzie E może być podzbiorem $\mathbb{R}$ lub $\mathbb{R}^k$. Niech $\tilde{f}(x)$ będzie gęstością na E, o której będziemy zakładać, że jest ściśle dodatnia na E i niech X będzie zmienną losową z gęstością $\tilde{f}(x)$. Wtedy

$$I = \int_E \frac{k(x)}{\tilde{f}(x)} \tilde{f}(x) \, dx = \mathbb{E}\, \frac{k(X)}{\tilde{f}(X)} = \mathbb{E}\, k_1(X),$$

gdzie $k_1(x) = k(x)/\tilde{f}(x)$. Zauważmy, że to samo można zapisać formalnie inaczej. Mianowicie możemy przypuścić (i to może być ściśle udowodnione), że Z jest określona na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ (ta sama Ω i $\mathcal{F}$), i $\tilde{\mathbb{P}}(Z \le x) = \int_0^x \tilde{f}(y) \, dy$. Teraz definiujemy *estymator istotnościowy*

$$\hat{Y}_n^{\text{IS}} = \frac{1}{n} \sum_{j=1}^{n} \frac{k(X_i)}{\tilde{f}(X)},$$

gdzie $X_1, \ldots, X_n$ są niezależnymi replikacjami liczby losowej X. Oczywiście $\hat{Y}_n^{\text{IS}}$ jest nieobciążony z wariancją

$$\text{Var}\,(\hat{Y}_n^{\text{IS}}) = \frac{1}{n} \text{Var}\, \frac{k(X)}{\tilde{f}(X)} = \frac{1}{n} \int_E \left(\frac{k(x)}{\tilde{f}(x)} \right)^2 \tilde{f}(x) \, dx.$$

Okazuje się, że odpowiednio wybierając $\tilde{f}(x)$ można istotnie zmniejszyć wariancję. Niestety, jak pokazuje następujący przykład, można też istotnie pogorszyć.

Przykład 4.1 Funkcja ćwiartki koła jest zadana wzorem

$$k(x) = 4\sqrt{1-x^2}, \qquad x \in [0,1],$$

i oczywiście

$$\int_0^1 4\sqrt{1-x^2}\,dx = \pi.$$

W tym przykładzie $Z = U$ jest o rozkładzie jednostajnym na $E = [0,1]$. Wtedy zgrubny estymator jest dany przez

$$\hat{Y}_n^{\mathrm{CMC}} = \frac{1}{n}\sum_{j=1}^{n} 4\sqrt{1-U_j^2}$$

i ma wariancję

$$\operatorname{Var}\hat{Y}_n^{\mathrm{CMC}} = \frac{1}{n}\Big(\int_0^1 16(1-x^2)\,dx - \pi^2\Big) = \frac{0.797}{n}.$$

Niech teraz $\tilde{f}(x) = 2x$. Wtedy estymator istotnościowy

$$\hat{Y}_n^{\mathrm{IS}} = \frac{1}{n}\sum_{i=1}^{n} \frac{4\sqrt{1-X_i^2}}{2X_i},$$

gdzie $X_1, \ldots, X_n$ są niezależnymi replikacjami liczby losowej X z gęstością $2x$ ma wariancję

$$\operatorname{Var}(\hat{Y}_n^{\mathrm{IS}}) = \frac{1}{n}\operatorname{Var}\frac{k(X)}{\tilde{f}(X)}.$$

Teraz policzymy wariancję

$$\operatorname{Var}\frac{k(X)}{\tilde{f}(X)} = \mathbb{E}\left(\frac{4\sqrt{1-X^2}}{2X}\right)^2 - I^2 = 8\int_0^1 \frac{1-x^2}{x}\,dx - I^2 = \infty.$$

Teraz zastanowmy się jak dobrać gęstość $\tilde{f}$ aby otrzymać estymator istotnościowy ale z mniejszą wariancją niż dla estymatora CMC. Rozpatrzmy teraz inny estymator $\hat{Y}$ gdy weźmiemy jako gęstość

$$\tilde{f}(x) = (4-2x)/3.$$

Wtedy

$$\hat{Y}_n^{\mathrm{IS}} = \frac{1}{n}\sum_{i=1}^{n} \frac{4\sqrt{1-X_i^2}}{(4-2X_i)/3},$$

i

$$\mathrm{Var}\,\frac{k(X)}{\tilde{f}(X)} = \int_0^1 \frac{16(1-x^2)}{(4-2x)/3} - \pi^2 = 0.224$$

a więc ponad 3 razy lepszy od estymatora CMC.

Przypatrzmy się wzorowi na wariancję

$$\int_0^1 \left(\frac{k(x)}{\tilde{f}(x)} - I\right)^2 \tilde{f}(x)dx.$$

Przykład 4.2 Obliczenie prawdopodobieństwa awarii sieci. Sieć jest spójnym grafem składającym się z węzłów, z których dwa z s, t są wyróżnione. Węzły są połaczone krawędziami ponumerowanymi $1, \ldots, n$. Każda krawędź może być albo sprawna – 0 lub nie – 1. W przypadku niesprawnej krawędzi będziemy mówili o przerwaniu. A więc stan krawędzi opisuje wektor $e_1, \ldots, e_n$, gdzie $e_i = 0$ lub 1. Sieć w całości może być sprawna – 0 lub nie (t.j. z przerwą) – 1, jeśli jest połączenie od s do t. Formalnie możemy opisać stan systemu za pomocą funkcji $k : \mathcal{S} \to \{0, 1\}$ określonej na podzbiorach $1, \ldots, n$ Mówimy, że sieć jest niesprawna i oznaczamy to przez 1 jeśli nie ma połaczenia pomiędzy węzłami s i t, a 0 w przeciwnym razie. To czy sieć jest niesprawna (lub sprawna) decyduje podzbiór B niesprawnych krawędzi. Podamy teraz przykłady. Będziemy przez $|B|$ oznaczać liczbę elementów zbioru B, gdzie $B \subset \{1, \ldots, n\}$.

- *Układ szeregowy* jest niesprawny gdy jakakolwiek krawędź ma przerwanie. A więc $k(B) = 1$ wtedy i tylko wtedy gdy $|B| \geq 1$. przykłady.

- *Układ równoległy* jest niesprawny jeśli wszystkie elementy są niesprawne. A więc $(B) = 1$ wtedy i tylko wtedy gdy $|B| = n$.

Teraz przypuśćmy, że każda krawędź może mieć przerwanie z prawdopodobieństwem q niezależnie od stanu innych krawędzi. Niech X będzie zbiorem (losowym) krawędzi z przerwaniami, tzn. $X \subset \{1, \ldots, n\}$. Zauważmy, że X jest dyskretną zmienną losowa z funkcją prawdopodobieństwa

$$\mathbb{P}(X = B) = p(B) = q^{|B|}(1-q)^{n-|B|},$$

dla $B \subset \{1,\ldots,n\}$. Naszym celem jest obliczenie prawdopodobieństwa, że sieć jest niesprawna. A więc

$$p = \mathbb{E}\, k(X) = \sum_{B\subset\{1,\ldots,n\}} k(B)p(B)$$

jest szukanym prawdopodobieństwem, że sieć jest niesprawna. Z powyższego widzimy, że możemy używać estymatora dla p postaci

$$\hat{X}_n = \frac{1}{n}\sum_{i=1}^{n} k(X_i),$$

gdzie $X_1,\ldots,X_n$ są niezależnymi replikacjami X. Ponieważ typowo prawdopodobieństwo I jest bardzo małe, więc aby je obliczyć musimy zrobić wiele replikacji. Dlatego jest bardzo ważne optymalne skonstruowanie dobrego estymatora dla I.

Zanalizujemy teraz estymator istotnościowy z $\tilde{Y}(B) = \beta^{|B|}(1-\beta)^{n-|B|}$. Mamy

$$I = \sum_{B\subset\{1,\ldots,n\}} k(B)p(B) = \sum_{B\subset\{1,\ldots,n\}} \frac{k(B)p(B)}{\tilde{p}(B)}\tilde{p}(B).$$

Teraz zauważmy, że $k(B) = 1$ wtedy i tylko wtedy gdy sięc jest niesprawna a w pozostałych przypadkach równa się zero. A więc

$$I = \sum_{B:k(B)=1} \frac{p(B)}{\tilde{p}(B)}\tilde{p}(B),$$

gdzie

$$\frac{p(B)}{\tilde{p}(B)} = \frac{q^{|B|}(1-q)^{n-|B|}}{\beta^{|B|}(1-\beta)^{n-|B|}} = \left(\frac{1-q}{1-\beta}\right)\left(\frac{q(1-\beta)}{\beta(1-q)}\right)^{|B|}.$$

Teraz rozwiązując nierówność

$$\frac{q(1-\beta)}{\beta(1-q)} \le 1$$

mamy warunek $\beta \ge q$. Niech teraz d będzie najmniejszą liczbą krawędzi z przerwaniem, powodująca niesprawność całej sieci. Na przykład dla sieci szeregowej b wynosi 1, a dla równoległej n. Dla sieci z Rys. ??? mamy $b = 3$. A więc

$$\left(\frac{1-q}{1-\beta}\right)^{n}\left(\frac{q(1-\beta)}{\beta(1-q)}\right)^{|B|} \le \left(\frac{1-q}{1-\beta}\right)\left(\frac{q(1-\beta)}{\beta(1-q)}\right)^{d}, \qquad |B| \ge d.$$

Teraz pytamy dla jakiego $\beta \geq q$ prawa strona jest najmniejsza. Odpowiedź na to pytanie jest równoważna szukaniu minimum funkcji $\beta^d(1-\beta)^{n-d}$. Przyrównując pochodną do zera mamy

$$d\beta^{d-1}(1-\beta)^{n-d} - (n-d)\beta^d(1-\beta)^{n-d-1} = 0.$$

Stąd $\beta = d/n$. A więc dla $|B| \geq d$ mamy

$$\left(\frac{1-q}{1-\beta}\right)^n \left(\frac{q(1-\beta)}{\beta(1-q)}\right)^{|B|} \leq \left(\frac{1-q}{1-\frac{d}{n}}\right) \left(\frac{q(1-\frac{d}{n})}{\frac{d}{n}(1-q)}\right)^d = \delta.$$

Stąd

$$\operatorname{Var} \hat{Y}_n^{\mathrm{IS}} \leq \frac{\delta}{n} \sum_{B:k(B)=1} p(B) = \frac{\delta}{n} I.$$

Ponieważ I zazwyczaj jest małe

$$\operatorname{Var} \hat{Y}_n = \frac{I(1-I)}{n} \approx \frac{I}{n}.$$

A więc estymator IS ma $1/\delta$ razy mniejszą wariancję. W szczególności dla sieci z Rys. ... mamy $\delta = 0.0053$. Czyli używając estymatora IS potrzebne będzie $\sqrt{1/\delta} \approx 14$ razy mniej replikacji.

Zadania teoretyczne

4.1 Dla przykładu 4.1 obliczyć wariancję $\hat{Y}_n^{IS}$ gdy $\tilde{f}(x) = (4 - 2x)/3$, $0 \leq x \leq 1$.

Rozdział V

Symulacje stochastyczne w badaniach operacyjnych

1 Modele kolejkowe

1.1 Proces Poissona

Zajmiemy się teraz opisem i następnie sposobami symulacji losowego rozmieszczenia punktów. Tutaj będziemy zajmowali się punktami na $\mathbb{R}_+$, które modelują momenty zgłoszeń. Jest jeden z elementów potrzebny do zdefiniowania procesu kolejkowego, lub procesu ryzyka. Jeśli zgłoszenia są pojedyncze to najprostszy sposób jest zadania tak zwanego procesu liczącego $N(t)$. Zakładamy, że $N(0) = 0$. Jest to proces losowy, który jest niemalejący, kawałkami stały z skokami jendostkowymi. Ponadto przyjmujemy, że jako funkcja czasu t jest prawostronnie ciągły. Interpretacją zmiennej losowej $N(t)$ jest liczba punktów do chwili t lub w w odcinku $(0, t]$. Zauważmy, że $N(t+h)-N(t)$ jest liczbą punktów w odcinku $(t, t + h]$. Przez proces Poissona o intensywności λ rozumiemy proces liczący, taki że

1. liczby punktów w rozłącznych odcinkach są niezależne,
2. liczba punktów w w odcinku $(t, t + h]$ ma rozkład Poissona z średnią λh.

Zauważmy, że

$$\mathbb{P}(N(t, t + h] = 1) = \lambda h e^{-\lambda h} = \lambda h + o(h),$$

i dlatego parametr λ nazywa się intensywnością procesu Poissona. Inna interperetacja to $\mathbb{E}\, N(t, t+h] = \lambda h$, czyli λ jest średnią liczbą punktów na jednoskę czasu. Średnia jest niezależna od czasu t. Przypomnijmy, że w podrozdziale II.2 zauważyliśmy, natępujący fakt. Niech $\tau_1, \tau_2, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie Exp(1) oraz

$$N = \#\{i = 1, 2, \ldots : \tau_1 + \ldots + \tau_i \leq \lambda\}.$$

Wtedy $N =$ ma rozkład Poissona z średnią λ. Zauważmy, że jeśli $\tau_1, \tau_2, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie Exp(λ), to

$$N = \#\{i = 1, 2, \ldots : \tau_1 + \ldots + \tau_i \leq 1\}.$$

Okazuje się, że tak naprawdę kolejne sumy niezależnych zmiennych losowych o jednakowym rozkładzie wykładniczym definiują coś więcej, czyli proces Poissona. A więc i-ty punkt ma współrzędną $t_i = \tau_1 + \ldots + \tau_i$, natomiast proces liczący można zapisać

$$N(t) = \#\{i = 1, 2, \ldots : \tau_1 + \ldots + \tau_i \leq t\}\ .$$

Podamy teraz procedurę matlabowską generowania procesu Poissona w odcinku (0,tmax].

Procedura poipro

```
% poipro simulate  Poisson process
% of intensity lambda.
%
% Inputs: tmax - simulation interval
%         lambda - arrival intensity
%
% Outputs: n - number of arrivals in (0,tmax],
%          arrtime - time points of arrivals, if n=0 then arrtime='empty'

arrtime=-log(rand)/lambda;
if (arrtime>tmax)
  n=0; arrtime='empty'
  else
  i=1;
```

```
  while (arrtime(i)<=tmax)
    arrtime = [arrtime, arrtime(i)-log(rand)/lambda];
    i=i+1;
  end
  n=length(arrtime);           \% arrival times t1,...,tn
end
```

Proces Poissona nie zawsze oddaje dobrze przybycia do systemu, ze względu na to że instnesywność jest stała. Istrnieją potrzeby modelowania zgłoszeń gdzie intensywność zależy od momentu t. Podamy teraz co to jest niejednorodny proces Poissona z *funkcją intensywności* $\lambda(t)$. Jeśli intensywność jednorodnego procesu Poissona λ jest $\mathbb{P}(N(dt) = 1) = \lambda\, dt$, to w przypadku procesu niejednorodnego $\mathbb{P}(N(dt) = 1) = \lambda(t)\, dt$. Formalnie definiujemy niejednorodny proces Poissona z funkcją intensywności $\lambda(t)$ jako proces liczący, taki że

1. liczby punktów w rozłącznych odcinkach są niezależne,
2. liczba punktów w odcinku $(t, t + h]$ ma rozkład Poissona z średnią $\int_t^{t+h} \lambda(s)\, ds$.

Okazuje się, że jesli w interesującym nas odcinku $(0, T_{\max}]$ $\lambda(t) \leq \lambda_{\max}$ to proces niejednorodny można otrzymać z procesu Poissona o intensywności $\lambda_{\max}$ przez niezależne przerzedzanie punktów jednorodnego procesu Poissona; jeśli jest punkt w punkcie t to zostawiamy go z prawdopodobieństwem $\lambda(t)/\lambda_{\max}$. Mamy więc schemat dwukrokowy. Najpierw generukemy proces Poissona o stałej intensywności $\lambda_{\max}$ a następnie przerzedzamy każdy punkt tego proces, niezależnie jeden od drugiego, z prawdopdobieństwami zależnymi od położenia danego punktu.

Podamy teraz procedurę matlabowską generowania procesu Poissona w odcinku (0,tmax].

Procedura nhpoipro

```
% nhpoipro simulate  nonhomogeneous Poisson process
% of intensity lambda(t).
%
% Inputs: tmax - simulation interval
%         lambda - arrival intensity function lambda(t)
```

```
%
% Outputs: n - number of arrivals in (0,tmax],
%          arrtime - time points of arrivals, if n=0 then arrtime='empty'

arrt=-log(rand)/lambdamax;
if (arrt>tmax)
  n=0; arrtime='empty';
  else
  i=1;
  while (arrt(i)<=tmax)
    arrtime=arrt;
    arrt = [arrt, arrt(i)-log(rand)/lambdamax];
    i=i+1;
  end
  m=length(arrtime);
  b= lambda(arrtime)/lambdamax;
  c= rand(1,m);
  arrtime=nonzeros(arrtime.*(c<b))
  n=length(arrtime);
end
```

2 Podstawowe modele teorii kolejek

Teoria kolejek jest działem procesów stochastycznych. Zajmuje się analizą systemów, do których wchodzą/wpływają zadania/pakiety/itd mające być obsługiwane, i które tworzą kolejkę/wypełniają bufor, i które doznają zwłoki w obsłudze. Zajmiemy sie teraz przeglądem podstawowych modeli i ich rozwiązań.

2.1 Standardowe pojęcia i systemy

Zadania zgłaszają się w momentach $A_1 < A_2 < \ldots$ i są rozmiarów $S_1, S_2, \ldots$. Odstępy między zgłoszeniami oznaczamy przez $\tau_i = A_i - A_{i-1}$. Zadania są obsługiwane według jednego z regulaminów jak na przykład:

- zgodnie z kolejnościa zgłoszeń (FCFS -First Come First Seved)

- obsługa w odwrotnej kolejności zgłoszeń (LCFS - Last come First Served)
- w losowej kolejności (SIRO),
- z podziałem czasu (PS), tj jesli w systemie jest n zadań to serwer poświęca każdemy zadaniu $1/n$ swojej mocy.

Zakładamy tutaj, że zadania są natychmiast obsługiwane jeśli tylko serwer jest pusty, w przeciwnym razie oczekuje w buforze. W przypadku większej liczby serwerów będziemy przyjmować, że są one jednakowo wydolne, to zanczy nie odgrywa roli w przypadku gdy kilka jest wolnych, który będzie obsługiwał zadanie. Bufor może być nieskończony lub z ograniczoną liczba miejsc. W tym ostatnim przypadku, jeśli bufor w momencie zgłoszenia do systemu jest pełny to zadanie jest stracone. Można badać:

- $L(t)$ – liczba zadań w systemie w chwili t,
- $L_q(t)$ – liczba zadań oczekujących w buforze,
- W_n – czas czekania n-tego zadania,
- D_n – czas odpowiedzi dla i-tego zadania (czas od momentu zgłoszenia do momentu wyjścia zadania z systemu)
- załadowanie $V(t)$ (workload) w chwili t, co jest sumą wszystkich zadań które zostały do wykonania (wliczając to co zostało do wyknania dla już obsługiwanych).

W przypadku systemów ze stratami, bada się prawdopodobieństwo, że nadchodzące zadanie jest stracone (blocking probability).

Przez $\bar{l}$, $\bar{w}$, $\bar{d}$, $\bar{v}$ będziemy oznaczać odpowiednio w warunkach stabilności: średnią liczbę zadań w systemie, średni czas czekania, średni czas odpowiedzi, średnie załadowanie.

M/M/1 Najłatwiejszym do pełnej analizy jest system gdy odstępy między zgłoszeniami (τ_i) są niezależnymi zmiennymi losowymi o jednakowym rozkładzie wykładniczym z parametrem λ. To założenie odpowiada, ze zgłoszenia są zgodne z procesem Poissona z intensywanością λ. Czasy obsług sa niezależne o jednakowym rozkładzie μ. Wtedy, między innymi, wiadomo że w przypadku $\lambda < \mu$ (lu równoważnie $\rho < 1$, gdzie $\rho = \lambda/\mu$), niezależnie od wyżej wymienionej dyscypliny obsługi:

- $p_k = \lim_{t\to\infty} \mathbb{P}(L(t) = k) = (1-\rho)\rho^k$ $(k = 0, 1, 2, \ldots)$ oraz $\bar{l} = \lim_{t\to\infty} \mathbb{E}\, L(t) = \frac{\lambda}{\mu-\lambda}$,
- $\bar{l}_q = \lim_{t\to\infty} \mathbb{E}\, L_q(t) = \frac{\lambda^2}{\mu(\mu-\lambda)}$.

Natomiast w przypadku dyscypliny FCFS mamy

- $\bar{w} = \bar{\lim}_{n\to\infty} \mathbb{E}\, W_n = \frac{\lambda}{\mu(\mu-\lambda)}$,
- $\bar{d} = \lim_{n\to\infty} \mathbb{E}\, D_n = \frac{1}{\mu-\lambda}$.

Ponadto mamy z prawdopobieństwem 1 zbieżność

- $\lim_{t\to\infty} \int_0^t L(s)\, ds/t = \frac{\lambda}{\mu-\lambda}$
- $\lim_{t\to\infty} \int_0^t L_q(s)\, ds/t = \frac{\lambda^2}{\mu(\mu-\lambda)}$

a w przypadku dyscypliny FCFS mamy

- $\lim_{k\to\infty} \sum_{j=1}^k W_j/k = \frac{1\lambda}{\mu(\mu-\lambda)}$
- $\lim_{k\to\infty} \sum_{j=1}^k d_j/k = \frac{1}{\mu-\lambda}$

M/M/c Proces zgłoszeń zadań oraz ich rozmiary są takie same jak w M/M/1. Natomiast w tym systemie jest c serwerów obsługujących z tą samą prędkościa. Podobnie jak dla M/M/1 możemy pokazać, że, jesli obługa jest zgodna z kolejnościaą zgłoszeń, to, przy założeniu $\rho < c$

$$\bar{w} = \lim_{n\to\infty} \mathbb{E}\, W_n = \frac{p_W}{\mu(c-\rho)}$$

i mamy również z prawdopdobieństwem 1 zbieżność

$$\lim_{k\to\infty} \sum_{j=1}^k W_j/k = \bar{w}\ .$$

Mamy rówież dla dyscypliny FCFS ale i również pewnej klasy dyscyplin obsługi

$$p_k = \lim_{t\to\infty} \mathbb{P}(L(t) = k) \quad = \quad \begin{cases} \frac{\rho^k}{k!} p_0, & k = 0, \ldots, c \\ \frac{\rho^k}{c! c^{k-c}} p_0, & k > c. \end{cases}$$

gdzie

$$p_0 = \left[\sum_{k=0}^{c} \frac{\rho^k}{k!} + \frac{\rho^{c+1}}{c!(c-\rho)} \right]^{-1} .$$

oraz z prawdopodobieństwem 1

$$\lim_{t\to\infty} \int_0^t L(s)\, ds/t = \ .$$

M/M/∞ Proces zgłoszeń zadań oraz ich rozmiary są takie same jak w M/M/1. Jednakże w tym systemie jest nieograniczony zasób serwerów a więc żadne zadanie nie czeka i nie ma sensu mówić o czasie czekania który jest zerowy i czasie odpowiedzi, który równy jest rozmiarowi zadania. Wtedy mamy, niezaleznie od wielkości ρ

- $\lim_{t\to\infty} \mathbb{P}(L(t) = k) = \frac{\rho^k}{k!} e^{-\rho}$ oraz $\lim_{t\to\infty} \mathbb{E}\, L(t) = \rho$,

M/G/∞ Proces zgłoszeń zadań jest taki sam jak w M/M/∞, jednakże rozmiary kolejnych zadań $(S_j)_j$ są niezależne o jednakowym rozkładzie G. Musimy założyć, że $\mathbb{E}\, S_1 = \mu^{-1} < \infty$. W tym systemie jest nieograniczony zasób serwerów a więc żadne zadanie nie czeka i nie ma sensu mówić o czasie czekania który jest zerowy i czasie odpowiedzi, który równy jest rozmiarowi zadania. Wtedy mamy, niezależnie od wielkości $\rho = \lambda \mathbb{E}\, S < \infty$

- $\lim_{t\to\infty} \mathbb{P}(L(t) = k) = \frac{\rho^k}{k!} e^{-\rho}$ oraz
- $\lim_{t\to\infty} \mathbb{E}\, L(t) = \rho$,

Proces $L(t)$ będzie stacjonarny jeśli jako $L(0)$ weźmiemy zmienną losową o rozkładzie Poissona Poi(ρ). Wtedy funkcja kowariancji [1]

$$\begin{aligned} R(t) &= \mathbb{E}\,[(L(s) - \rho)(L(s+t) - \rho)] && (2.1) \\ &= \lambda \int_t^\infty (1 - G(s))\, ds\ . && (2.2) \end{aligned}$$

oraz jeśli w chwili $t = 0$ nie ma żadnych zadań

$$\rho - \mathbb{E}\, L(t) = \lambda \int_t^\infty (1 - G(s))\, ds\ .$$

W szczególności dla M/D/∞, tj. gdy rozmiary zadań są deterministyczne mamy $R(t) = \rho(1 - \mu t)$ dla $t \le \mu^{-1}$ oraz $R(t) = 0$ dla $t > \mu^{-1}$.

[1] Cytowanie za Whittem [12]

M/M/c ze stratami Proces zgłoszeń zadań oraz ich rozmiary są takie same jak w M/M/1. Natomiast w tym systemie jest c serwerów obsługujących z tą samą prędkościa, ale nie ma bufora. Jeśli przychodzące zadanie zastaję wszystkie serwery zajęte, to jest stracone. Można pokazać istnienie granicy

$$p_k = \lim_{t\to\infty} \mathbb{P}(L(t) = k) = \frac{\rho^k/k!}{\sum_{j=0}^{c} \rho^j/j!}, \qquad k = 0, \ldots, c$$

gdzie $\rho = \lambda/\mu$, i teraz ta granica istnieje zawsze niezależnie od wielkości ρ. W szczególności mamy też

$$\lim_{t\to\infty} \frac{\int_0^t \mathbf{1}(L(s) = c)\, ds}{t} = p_c \, .$$

Co ciekawsze, jeśli

$$I_k = \begin{cases} 1 & k\text{-te zadanie jest stracone} \\ 0 & \text{w przeciwnym razie} \end{cases}$$

to z prawdopodbieństwem 1 istnieje granica

$$p_{\text{block}} \;=\; \lim_{k\to\infty} \sum_{j=1}^{k} I_j/k = \frac{\rho^c/c!}{\sum_{j=0}^{c} \rho^j/j!}. \tag{2.3}$$

Prawdopodobieństwo p_{block} jest stacjonarnym prawdopdobieństwem, że przychodzące zadanie jest stracone. Odróżnijmy p_{block} od p_c, które jest stacjonarnym prawdopodobieństwem, że system jest pełny. Tutaj te prawdopodobieństwa się pokrywają, ale tak nie musi byc, jeśli zadania zgłaszają się do systemu nie zgodnie z procesem Poissona. Okazuje się, że tzw. wzór Erlanga (2.3) jest prawdziwy dla systemów M/G/c ze stratami, gdzie rozmiary kolejnych zadań tworzą ciąg niezaleznych zmiennych losowych z jednakową dystrybuantą G, gdzie $\mu^{-1} = \int_0^\infty x\, dG(x)$ jest wartością oczekiwaną rozmiaru zadania.

GI/G/1 Jest to system taki, że $(\tau_k)_k$ ($A_1 = \tau_1, A_2 = \tau_1 + \tau_2, \ldots$) są niezaleznymi zmiennymi losowymi o jednakowym rozkładzie F a rozmiary zadań $(S_k)_k$ tworzą ciąg niezależnych zmiennych losowych o jednakowym rozkładzie G. Przypuśćmy, że w chwili $t = 0$ system jest pusty, i pierwsze zadanie przychodzi w momencie A_1. Wtedy oczywiście czas czekania tego zadania wynosi $W_1 = 0$. Dla systemów jednokanałowych prawdziwa jest rekurencja

$$W_{k+1} = (W_k + S_k - \tau_{k+1})_+. \tag{2.4}$$

Można rozpatrywać alternatywne założenia początkowe. Na przykład, przypuścić, że w chwili $A_0 = 0$ mamy zgłoszenie zadania rozmiaru S_0, które będzie musiało czekać przez czas W_0. W takim przypadku rozpatruje się rózne załozenia o rozkładzie W_0, S_0 i A_1, ale zawsze trzeba pamiętać niezależności ażeby poniższe fakty były prawdziwe. Natomiast rekurencja (2.4) jest niezależna od wszlkich założeń probabilistycznych. Przy założeniach jak wyżej W_k tworzy łańcuch Markowa i jeśli $\mathbb{E}\, S_1 < \mathbb{E}\, \tau_1$, to istnieją granice $\lim_{k\to\infty} \mathbb{P}(W_k \le x)$ oraz, jeśli $\mathbb{E}\, S_1^2 < \infty$

$$\bar{w} = \lim_{k\to\infty} \mathbb{E}\, W_k \, .$$

Ponadto z prawdopodobieństwem 1

$$\bar{w} = \lim_{k\to\infty} \sum_{j=1}^{k} W_j/k \, .$$

W przypadku, gdy τ_1 ma rozkład wykładniczy z $\mathbb{E}\, \tau_1 = 1/\lambda$ to

$$\bar{w} = \frac{\lambda \mathbb{E}\, S_1^2}{2(1 - \lambda/\mathbb{E}\, S_1)} \, . \tag{2.5}$$

Wzór (2.5) jest zwany wzorem Pollaczka-Chinczyna. Jeśli S ma rozkład wykładniczy z parametrem μ, to wtedy powyższy wzór pokrywa się odpowiednim wzorem dla M/M/1.

W szczególności jeśli rozmiary zadań są niezależne o jednakowym rozkładzie wykładniczym, to oznaczamy taki system przez GI/M/1. Jeśli proces zgłoszeń zadań jest procesem Poissona, to wtedy piszemy M/G/1. Jeśli zarówno proces wejściowy jest Poissonowski jak i rozmiary zadań są wykładnicze, to piszemy M/M/1.

GI/G/c Jest to system taki, że odstępy między zgłoszeniami są $(\tau_k)_k$ $(A_1 = \tau_1, A_2 = \tau_1 + \tau_2, \ldots)$ a rozmiary zadań $(S_k)_k$. Jest c serwerów, i zadania oczekujące w buforze do obługi zgodnie z kolejnością zgłoszeń. Zamiast rekurencji (2.4) mamy następującą. Oznaczmy $\boldsymbol{V}_n = (V_{n1}, V_{n2}, \ldots, V_{nc})$, gdzie $0 \le V_{n1} \le V_{n2} \le \ldots \le V_{nc}$. Składowe tego wektora reprezentują dla n-go zadania (uporządkowane) załadowanie każdego z serwerów. A więc przychodząc idzie do najmniej załadowanego z załadowaniem V_{n1} i wobec tego jego załadowanie tuż po momencie zgłoszenia wyniesie $V_{n1} + S_n$. Po

czasie τ_{n+1} (moment zgłoszenia $n+1$-szego) załadowania wynoszą $V_{n1}+S_n-\tau_{n+1}, V_{n2}-\tau_{n+1}, \ldots, V_{nc}-\tau_{n+1}$. Oczywiście jesli są ujemne to kładziemy 0, tj $(V_{n1}+S_n-\tau_{n+1})_+, (V_{n2}-\tau_{n+1})_+, \ldots, (V_{nc}-\tau_{n+1})_+$. Teraz musimy wyniki uporządkować. Wprowadzamy więc operator $\mathcal{R}(\boldsymbol{x})$, który wektor $\boldsymbol{x}$ z $\mathbb{R}^c$ przedstawia w formie uporządkowaej tj. $x_{(1)} \leq \ldots \leq x_{(c)}$. A więc dla systemu z c serwerami mamy rekurencję

$$\begin{aligned}&(V_{n+1,1}, V_{n+1,2}, \ldots, V_{n+1,c})\\ &\quad = \quad \mathcal{R}((V_{n1}+S_n-\tau_{n+1})_+, (V_{n2}-\tau_{n+1})_+, \ldots, (V_{nc}-\tau_{n+1})_+ \ .\end{aligned}$$

Przypuśćmy, że w chwili $t=0$ system jest pusty, i pierwsze zadanie przychodzi w momencie A_1. Wtedy oczywiście czas czekania tego zadania wynosi $V_1 = V_2 = \ldots = V_c = 0$. Rekurencja powyższa jest prawdziwa bez założeń probabilitycznych. Jesli teraz założymy, że ciąg (τ_n) i (S_n) są niezależne i odpowiednio o jednakowym rozkładzie F i G, oraz jeśli $\rho = \frac{\mathbf{E}\,S}{\mathbf{E}\,T} < c$, to system się stabilizuje i istnieje rozkład stacjonarny.

3 Metoda dyskretnej symulacji od zdarzenia po zdarzenie

Tutaj zajmiemy się symulacjami systemów, które ewoluują w czasie, zmieniając swój stan jedynie w momentach dyskretnych. Będziemy więc symulować przebieg stanów systemu od zdarzenia do zdarzenia. Tak metoda nazywa sie w języku angielskim *discrete event simulation*.

Przypuścmy, że chcemy symulować $\boldsymbol{n}$ – *stan systemu* do momentu $T_{\max}$ - horyzont symulacji. Jedyne stochastyczne zmiany systemu są możliwe w pewnych momentach $A_1, A_2, \ldots$. Pomiędzy tymi momentami system zachowuje się deterministycznie. Przypuścmy, że mamy N typów zdarzeń, z każdą związana aktywność $\mathcal{A}_i$ $(i=1,\ldots,N)$. Pomysł polega na utworzeniu listy zdarzeniowej LZ w której zapisujemy momenty następnych po momencie obecnego zdarzenie; zdarzenie $\mathcal{A}_i$ będzie w momencie $t_{\mathcal{A}_i}$. Algorytm znajduje najbliższe zdarzenie, podstawia nową zmienną czasu, wykonuje odpowiednie aktywności, itd.

Ogólny algorytm dyskretnej symulacji po zdarzeniach

`INICJALIZACJA`

```
podstaw zmienną czasu równą 0;
ustaw parametry początkowe "stanu systemu" i liczników;
ustaw początkową listę zdarzeniową LZ;

PROGRAM

podczas gdy zmienna czasu < tmax;
wykonaj
   ustal rodzaj następnego zdarzenia;
   przesuń zmienną czasu;
   uaktualnij {\it stan systemu} + liczniki;
   generuj następne zdarzenia + uaktualnij listę zdarzenową LZ
koniec

OUTPUT
Obliczyć estymatory + przygotować dany wyjściowe;
```

3.1 Symulacja kolejka z jednym serwerem

Metodę zilustrujemy na przykładzie algorytmu symulacji kolejki M/M/1. Będziemy potrzebować

- zmienna czasu – t
- liczniki:
  NA – liczba zgłoszeń do czasu t (włącznie)
  NO – liczba odejść (wykonywanych zadań) do czasu t (włącznie)
- n – zmienna systemowa; liczba zadań w chwili t,

Przez $T_{\max}$ oznaczamy horyzont czasowy symulacji. Po $T_{\max}$ nie będziemy rejestrować już wejść ale jedynie wyjścia. Dla systemu M/M/1 mamy dwa zdarzenia: *zgłoszenie*, *odejście*. Poszczególne elementy algorytmu zależą od stanu *listy zdarzeniowej* LZ, tj. od wielkości tA,tO - czasów następnego odpowiednio zgłoszenia i odejścia. Możliwe stany listy zdarzeniowej LZ to tA,tO, tA,tO, $\emptyset$. Będziemy też kolekcjonować następujące dane (output variables):

- t(i),n(i) - zmienne otrzymane (output variables) – moment kolejnego i-tego zdarzenia oraz stan systemu w tym momencie,
- (A(i),O(i)) - zmienne otrzymane (output variables) – moment zgłoszenia oraz moment wykonania i-tego zadania,
- tp – czas od tmax do ostatniego odejścia,
- tmax – horyzont czasowy symulacji.

Ustalamy horyzont czasowy symulacji tmax. Po $T_{\max}$ nie będziemy rejestrować już wejść ale jedynie wyjścia. Niech F będzie dystrybuantą odstępu τ_i między zgłoszeniami oraz G dystrybuanta rozmiaru zadania. Zauważmy:

- Na podstawie (t(j),n(j) możemy wykreślić realizację $L(s)$, $0 \leq s \leq tmax$.
- Na podstawie A(i),O(i) mozemy znaleźć ciąg czasów odpowiedzi $D_i = O(i) - A(i)$

Algorytm dla M/M/1

```
POCZĄTEK (gdy w chwili 0 system jest pusty).

utwórz wektory/macierze:
   DK -- macierz dwu-wierszowa,
   A -- wektor
   D -- wektor.

Losujemy X \sim Exp(lambda).
  Jeśli X>tmax, to NA=0 i  LZ=\emptyset.
  Jeśli X=< tmax,
   to podstawiamy
   t=n=NA=N0=0, tA=X i LZ=tA.
```

[Zauważmy, że jeśli $n = 0$ to odpowiada mu przypadek LZ={tA}. Następna część zależy od stanu LZ.]

```
PRZYPADEK 1. LZ=tA.

dołącz do DK: t:=tA, n:=n+1,
resetuj: NA:=NA+1
dołącz do A:   A(NA)=t
generuj: zm. los. X o rozkładzie  Exp(mu) do najbliższego zgłoszenia,
generuj: zm. los. S o rozkładzie Exp(mu) (rozmiar następnego zadania),
resetuj: jeśli t+X>tmax to: tO=t+S; nowy LZ={t_O}
resetuj: jesli t+X=< tmax, to tO=t+S, tA=t+X; nowy LZ={tA,tO}.

PRZYPADEK 2. LZ={tA,tO}, gdzie tA=< tO.

dołącz do DK: t=tA, n=n+1
resetuj NA=NA+1
dołącz do A: A(NA)=t
generuj: X\sim Exp(mu)
resetuj: jeśli t+X>tmax to podstaw tO; nowy LZ={tO}
resetuj: jesli t+X=< tmax,
to  podstaw tO, tA=t+X; nowy LZ={tA,tO}

PRZYPADEK 3. LZ={tA,tO}, gdzie tA> tO.
```

```
dołącz do DK: t=t0, n=n-1
resetuj N0=N0+1
dołącz do D: 0(N0)=t
jeśli n>0: jeśli to generuj: Y\sim G i resetuj  t0=t+Y,
nowa lista zdarzeniowa LZ={tA,t0}
jeli n=0 to nowa  LZ={tA}

PRZYPADEK 4:  LZ={t0}

Procedura wg przyp.~3, z tą różnicą, że jeśli zresetowana wartość n=0
to resetujemy listę zdarzeniową LZ=\emptyset

PRZYPADEK 5: Lista zdarzeniowa LZ=\emptyset

KONIEC -- skończ symulacje.
```

Teraz zrobimy kilka uwag o rozszerzeniu stosowalności algorytmu.

Dowolne warunki początkowe Modyfikujemy POCZĄTEK następująco:

```
POCZĄTEK (gdy w chwili 0 w systemie  jest k>0 zadań).

utwórz wektory/macierze:
  DK -- macierz dwu-wierszowa,
  A -- wektor
  D -- wektor.
Losujemy X\sim Exp(lambda oraz Y\sim Exp(mu).
Jeśli X>tmax, to NA=0 i  LZ=\emptyset.
Jeśli X=< tmax,
   to podstawiamy
   n=k,  tA=X, tD=Y i LZ={tA,tD}.
```

Możemy też założyć, że w chwili $t = 0$ jest k zadań z prawdopodobieństwem p_k. Wtedy należy wylosować najpierw k z prawdopodobieństwem p_k.

Jeśli F jest dowolne, natomiast $G = \mathrm{Exp}(\mu)$ to musimy oprócz założenia o liczbie zadań, że w chwili 0 w systemie jest k zadań, musimy też założyc znajomość tA.

```
POCZĄTEK (gdy w chwili 0 w systemie  jest $k>0$ zadań
                i czas do wpłynięcia nowego jest $t_A$ ).

utwórz wektory/macierze:
  DK -- macierz dwu-wierszowa,
  A -- wektor
  D -- wektor.
Losujemy  Y\sim Exp(mu).
Jeśli tA>tmax, to NA=0 i  LZ=\emptyset.
Jeśli tA=< tmax,
  to podstawiamy
  n=k,  tD=Y i LZ=tA,tD}.



POCZĄTEK (gdy w chwili 0 w systemie  jest k=0 zadań
                 i czas do wpłynięcia nowego jest tA ).

utwórz wektory/macierze:
  DK -- macierz dwu-wierszowa,
  A -- wektor
  D -- wektor.
Jeśli tA>tmax, to NA=0 i  LZ=\emptyset.
Jeśli tA=< tmax,
   to podstawiamy
   n=0 i LZ={tA}.
```

3.2 Własność braku pamięci

Niektóre algorytmy istotnie wykorzystują tzw. *brak pamięci* rozkładu wykładniczego czasu obsługi i opuszczenie tego założenia będzie wymagało istotnych modyfikacji. Natomiast możemy założyć, że są znane momenty napływania zadań $A_1 < A_2 < \ldots$. Tak jest na przykład gdy zgłoszenia są zgodne z niejednorodnym Procesem Poissona. Taj jest na przykład gdy odstępy między zgłoszeniami są niezależne o jednakowym rozkładzie (niewykładniczym). W każdym zdarzeniu (niech będzi ono w chwili t), gdy dotychczasowy algorytm mówił, że należy generować X i podstawić $t_A = t + X$, należy podstawić najwcześniejszy moment zgłoszenia zadania po chwili t.

Przy konstrukcji algorytmu dla systemów z jednym serwerem wykorzystywaliśmy tak zwaną własność *braku pamięci* rozkładu wykładniczego. Własność można sformułować następująco: dla:

$$\mathbb{P}(X > s+t | X > s) = \mathbb{P}(X > t), \qquad s, t \geq \ .$$

Łatwo pokazać, że własnośc jest prawdziwa dla zmiennej losowej o rozkładzie wykładniczym $X \sim$Exp(λ). Okazuje się również, że rozkład wykładniczy jest jedynym rozkładem o takiej własności.

Będzie też przydatny następujący lemat.

Lemat 3.1 *Jeśli $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowym oraz $X_i \sim$ Exp(λ_i), to zmienne*

$$T = \min_{i=1,\ldots,n} X_i, \qquad I = \arg \min_{i=1,\ldots,n} X_i$$

są niezależne oraz $T \sim$ Exp$(\lambda_1 + \ldots\ldots + \lambda_n)$ i $\mathbb{P}(J = j) = \lambda_j/(\lambda_1 + \ldots\ldots + \lambda_n)$.

Dowód Mamy

$$\begin{aligned}
\mathbb{P}(T > t, J = j) &= \mathbb{P}(\min_{i=1,\ldots,n} X_i > t, \min_{i \neq j} X_i > X_j) \\
&= \int_0^\infty \mathbb{P}(\min_{i=1,\ldots,n} X_i > t, \min_{i \neq j} X_i > x) \lambda_j e^{-\lambda_j x} \, dx \\
&= \int_t^\infty \mathbb{P}(\min_{i \neq j} X_i > x) \lambda_j e^{-\lambda_j x} \, dx \\
&= \int_t^\infty e^{-\sum_{i \neq j} \lambda_i x} \lambda_j e^{-\lambda_j x} \, dx \\
&= e^{-\sum_i \lambda_i t} \frac{\lambda_j}{\sum_i \lambda_i}
\end{aligned}$$

4 Uwagi bibliograficzne

Metoda symulacji po zdarzeniach (discrete event simulation) ma bogatą literaturę. W szczególności można polecić podręcznika Fishmana [4].

Zadania laboratoryjne

4.1 Przeprowadzić następujące eksperymenty dla M/M/1. Napisać program kreślący realizację $L(t)$ dla $0 \le t \le 10$. 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 $\rho = \lambda/\mu = 2/3$.

4.2 Kontynuacja zad. 1. Obliczyć $\widehat{L}(t_{\max})$, gdzie $\widehat{L}(t) = (\int_0^t L(s)\,ds)/t$ gdy $L(0) = 0$.
(a) Przeprowadzić eksperyment przy $\lambda = 1, \mu = 1.5$ oraz $L(0) = 0$ dla $t_{\max} = 5, 10, 50, 100, 1000$. Policzyć dla każdej symulacji połowy długości przedziały ufności na poziomie 95%.
(b) Przeprowadzić eksperyment przy $\lambda = 1, \mu = 1.1$ oraz $L(0) = 0$ dla $t_{\max} = 5, 10, 50, 100, 1000$. Policzyć

4.3 Kontynuacja zad. 1. Przez replikacje rozumiemy teraz pojedynczą realizację $L(t)$ dla $0 \le t \le 10$. Dla replikacji $L_i(t)$ zachowujemy wektor $\boldsymbol{l}_i = (L_i(1), L_i(2), \ldots, L_i(10))$ i niech 1000 będzie liczbą replikacji. Obliczyć

$$\widehat{\boldsymbol{l}} = (\frac{1}{1000}\sum_{j=1}^{1000} L_j(0.5), \frac{1}{1000}\sum_{j=1}^{1000} L_j(1), \ldots, \frac{1}{1000}\sum_{j=1}^{1000} L_j(10)) \,.$$

Zrobić wykres średniej długości kolejki w przedziale $0 \le t \le 10$.
(a) Przeprowadzić eksperyment dla $L(0) = 0$.
(b) Przeprowadzić eksperyment dla $L(0) \sim \text{Geo}(\rho)$.

4.4 Wygenerować $A_1, A_2, \ldots$, w odcinku $[0, 1000]$, gdzie $A_0 = 0$ oraz $A_1 < A_2 < \ldots$ 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 $N(t)$ i znaleźć jego średnią. Policzyć z symulacji $N(1000)/1000$ – średnią liczbą punktów na jednostkę czasu, zwaną asymptotyczną intensywnością $\bar{\lambda}$. Porównać z

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

4.5 Kontynuacja zad. 4. Wygenerować $\tau_1, \tau_2, \ldots, \tau_{1000}$, gdzie $\tau_i = A_i - A_{i-1}$, $A_0 = 0$ oraz $A_1 < A_2 < \ldots$ są kolejnymi punktami w niejednorodnym procesie Poissona z funckcją intensywności $\lambda(t) = a(2 - \sin(\frac{2\pi}{24}t))$. Przyjąć $a = 10$. Obliczyć średni odstęp między punktami $\hat{\tau} = \sum_{j=1}^{1000} \tau_i/1000$.

4.6 Napisać program symulujący realizację procesu liczby zadań $L(t)$ w systemie M/G/∞. Przeprowadzić eksperyment z
(a) $\lambda = 5$ oraz $G = \mathrm{Exp}(\mu)$ 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).
Na jednym wykresie umieścić $\int_0^t L(s)\,ds/t$ dla $1 \leq t \leq 100$ z symulacji (a) i (b). Przyjąć $L(0) = 0$.

Rozdział VI

Planowanie symulacji procesów stochastycznych

Przypuśćmy, że chcemy obliczyć wartość I, gdy wiadomo, że z prawdopodobieństwem 1

$$I = \lim_{t\to\infty} \frac{\int_0^t X(s)\,ds}{t}$$

dla pewnego procesu stochastycznego $(X(s))_{s\geq 0}$. Wtedy naturalnym estymatorem jest

$$\hat{X}(t) = \frac{\int_0^t X(s)\,ds}{t},$$

który jest (mocno) zgodnym estymatorem I. Jednakże zazwyczaj $\hat{X}$ nie jest nieobciązonym estymatorem, tzn. $\mathbb{E}\,\hat{X} \neq I$.

1 Planowanie symulacji dla procesów stacjonarnych i procesów regenerujących się

1.1 Procesy stacjonarne

Rozważmy proces stochastyczny $(X(t))_{t\geq 0}$ dla którego $\mathbb{E}\,X(t)^2 < \infty$, $(t \geq 0)$. Mówimy, że jest on *stacjonarny* jeśli spełnione są następujące warunki.

(1) Wszystkie rozkłady $X(t)$ są takie same. W konsekwencji wszystkie $\mathbb{E}\,X(t)$ są równe jeśli tylko pierwszy moment $m = \mathbb{E}\,X(0)$ istnieje.

(2) Dla każdego $0 < h$, łączne rozkłady $(X(t), X(t+h))$ są takie same. W konsekwencji wszystkie kowariancje $R(h) = \mathrm{Cov}\,(X(t), X(t+h)$ są równe jeśli tylko drugi moment istnieje.

(k) Dla każdego $0 < h_1 < \ldots < h_{k-1}$, łączne rozkłady $(X(t), X(t + h_1, \ldots, X(t + h_{k-1}))$ są takie same.

Procesy stacjonarne mają następującą własność. Z prawdopodobieństwem 1

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t X(s)\, ds$$

istnieje i jeśli granica jest stałą, to musi równać się $I = \mathbb{E}\, X(0)$. Do tego potrzebne jest aby było spełnione tzw założenie ergodyczności. Łatwo jest podać przykład procesu który nie jest ergodyczny. Weźmy na przykład proces $X(t) = \xi$, gdzie ξ jest zmienną losową. Wtedy w trywialny sposób

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t X(s)\, ds = \xi\ .$$

Jesli są spełnione ponadto tzw. *warunki mieszania*, to wtedy dla pewnej stałej $\bar{\sigma}^2$ (to nie jest $\mathrm{Var}\, X(t)$) proces

$$\frac{\int_0^t X(s)\, ds - tI}{\bar{\sigma}\sqrt{t}} \tag{1.1}$$

dla $t \to \infty$, zbiega według rozkładu do $\mathcal{N}(0,1)$. Zastanówmy się, ile wynosi $\bar{\sigma}^2$. Najpierw przepiszemy (1.1) w wygodnej dla nas formie

$$\frac{\int_0^t X(s)\, ds - tI}{\bar{\sigma}\sqrt{t}} = \frac{\sqrt{t}}{\bar{\sigma}}(\hat{X}(t) - I).$$

Chcemy aby $(\hat{X}(t) - I)$ było ustandaryzowane (przynajmniej asymptotycznie). Oczywiście $\mathbb{E}\,(\hat{X}(t) - I) = 0$, natomiast aby wariancja była równa 1 musimy unormować

$$\frac{\hat{X}(t) - I}{\sqrt{\mathrm{Var}\, \hat{X}(t)}}\ .$$

Zbadamy teraz zachowanie asymptotyczne mianownika.

Fakt 1.1 *Jeśli* $(X(t)_{t\geq}$ *jest procesem stacjonarnym,* $\mathbb{E}\,X(t)^2 < \infty$ *oraz* $\int_0^\infty |R(v)|\,dv < \infty$*, to zmienna losowa* $\hat{X}(t)$ *ma wartość oczekiwaną* I *oraz*

$$\lim_{t\to\infty} t\mathrm{Var}\,(\hat{X}(t)) = \bar{\sigma}^2 = 2\int_0^t R(s)\,ds\ .$$

Zanim pokażemy dowód zrobimy kilka uwag. Wielkość $\bar{\sigma}^2 = 2\int_0^\infty R(s)\,ds$ zwie się TAVC (od *time average variance constant*). Ponieważ

$$\frac{\hat{X}(t)-I}{\sqrt{\mathrm{Var}\,\hat{X}(t}} = \frac{\sqrt{t}}{\bar{\sigma}}(\hat{X}(t)-I)\left(\frac{\bar{\sigma}}{\sqrt{t}}\right)/\sqrt{\mathrm{Var}\,\hat{X}(t}$$

oraz

$$\left(\frac{\bar{\sigma}}{\sqrt{t}}\right)/\sqrt{\mathrm{Var}\,\hat{X}(t} \to 1$$

skąd mamy CTG w postaci

- $\frac{\sqrt{t}}{\bar{\sigma}}(\hat{X}(t)-I)$ zbiega wg rozkładu do $\mathcal{N}(0,1)$.

Dowód Faktu 1.1. Mamy

$$\mathbb{E}\int_0^t X(s)\,ds = \int_0^t \mathbb{E}\,X(s)\,ds = mt\ .$$

Dalej bez straty ogólności możemy założyć $m=0$. Wtedy

$$\begin{aligned}
&\mathrm{Var}\,(\int_0^t X(s)\,ds)\\
&= \mathbb{E}\,(\int_0^t X(s)\,ds)^2\\
&= \mathbb{E}\,[\int_0^t\int_0^t X(w)X(v)\,dw\,dv]\\
&= \int_0^t\int_0^t \mathbb{E}\,X(w)X(v))\,dw\,dv\\
&= \int_{(w,v)\in[0,t]^2:w\leq v} \mathbb{E}\,X(w)X(v))\,dw\,dv + \int_{(w,v)\in[0,t]^2:w<v} \mathbb{E}\,X(w)X(v))\,dw\,dv
\end{aligned}$$

Teraz dla $v \leq w$ mamy $\mathbb{E}\, X(w)X(v)) = R(w-v)$ i stąd

$$\begin{aligned}\int_{(w,v)\in[0,t]^2:w\leq v} \mathbb{E}\, X(w)X(v)\, dw\, dv &+ \int_{(w,v)\in[0,t]^2:w<v} \mathbb{E}\, X(w)X(v)\, dw\, dv \\ &= 2\int_0^t dv \int_0^v R(v-w)\, dv \\ &\sim t\, 2\int_0^\infty R(w)\, dw\ .\end{aligned}$$

□

Podsumujmy. Podobnie jak w rozdziale III.1 możemy napisać *fundamentalny związek*

$$b = z_{1-\epsilon/2}\frac{\bar{\sigma}}{\sqrt{t}}\ .$$

pomiędzy błędem b, poziomem ufności ϵ oraz horyzontem symulacji t. Natomiast przedział ufności na poziome ϵ wynosi

$$\left(\hat{X}(t) - \frac{z_{1-\epsilon/2}\bar{\sigma}}{\sqrt{t}}, \hat{X}(t) + \frac{z_{1-\epsilon/2}\bar{\sigma}}{\sqrt{t}}\right)\ .$$

lub w formie

$$(\hat{X}(t) \pm \frac{z_{1-\epsilon/2}\bar{\sigma}}{\sqrt{t}}).$$

Niestety w typowych sytuacjach nie znamy $\bar{\sigma}^2$ a więc musimy tą wielkość wyestymować z symulacji.

Przykład 1.2 Dla systemu M/M/1 jest prawdą, że

$$\bar{\sigma}^2 = \frac{2\rho(1+\rho)}{(1-\rho)^4}\ .$$

Stąd widać, że $\bar{\sigma}^2$ gwałtownie rośnie gdy $\rho \to 1$, czyli w *warunkach wielkiego obciążenia*. Wzór powyższy jest szczególnym przypadkiem wzoru otrzymanego dla procesów narodzin i śmierci bez obliczania funkcji $R(t)$; patrz Whitt [13].

2 Przypadek symulacji obciążonej

Będziemy teraz analizowali przypadek symulacji liczby I dla estymatora $\hat{X}$ spełniającego:

1. $I = \lim_{t\to\infty} \hat{X}(t) = I$ (z prawdopodobieństwem 1),
2. $\frac{\sqrt{t}}{\bar{\sigma}}(\hat{X}(t) - I)$ zbiega wg rozkładu do $\mathcal{N}(0,1)$, gdzie

$$t\mathrm{Var}\,(\int_0^t X(s)\,ds) \to \bar{\sigma}^2,$$

gdy $t \to \infty$ dla pewnej skończonej liczby $\bar{\sigma}^2 > 0$.

Ponieważ nie zakładamy stacjonarności, więc estymator $\hat{X}(t)$ jest zgodny ale nie jest nieobciążony. Powstaje więc problem analizy jak szybko $b(t) = \mathbb{E}\,\hat{X}(t) - I$ zbiega do zera. Zanalizujemy przypadek gdy $\beta = \int_0^\infty (\mathbb{E}\,X(s) - I)\,ds$ zbiega absolutnie, tzn . $\int_0^\infty |\mathbb{E}\,X(s) - I|\,ds < \infty$. Wielkość

$$\beta = \lim_{t\to\infty} t(\mathbb{E}\,\hat{X}(t) - I)$$

nazywamy *asymptotycznym obciążeniem*. Wtedy

$$\begin{aligned}
&\mathbb{E}\,\hat{X}(t) - I \\
&\quad = \frac{\int_0^t (\mathbb{E}\,(X(s)) - I)}{t} \\
&\quad = \frac{\int_0^\infty (\mathbb{E}\,(X(s)) - I)\,ds - \int_t^\infty (\mathbb{E}\,(X(s)) - I)}{t} \\
&\quad = \frac{\beta}{t} + o(1/t)\ .
\end{aligned}$$

Ponieważ $\sqrt{\mathrm{Var}\,\hat{X}(t)} \sim \bar{\sigma}/\sqrt{t}$ a więc zmienność naszego estymatora jest rzędu $\mathrm{O}(t^{-1/2})$ czyli większa niż jego obciążenie.

Przykład 2.1 Aby uzmysłowić czytelnikowi trudność jakie można napotkać podczas symulacji rozpatrzymy przypadek symulacji średniej liczby zadań w systemie M/G/∞. Będziemy rozpatrywać liczbę zadań $L(t)$ dla

1. $\lambda = 5$, i $S =_{\mathrm{d}} \mathrm{Exp}(1)$, $T_{\max} = 1000$,
2. $\lambda = 5$, i $S =_{\mathrm{d}} 0.2X$, gdzie $X =_{\mathrm{d}} \mathrm{Par}(1.2)$,$T_{\max} = 1000$,
3. $\lambda = 5$, i $S =_{\mathrm{d}} 0.2X$, gdzie $X =_{\mathrm{d}} \mathrm{Par}(1.2)$,$T_{\max} = 10000$,
4. $\lambda = 5$, i $S =_{\mathrm{d}} 0.2X$, gdzie $X =_{\mathrm{d}} \mathrm{Par}(1.2)$,$T_{\max} = 100000$,

5. $\lambda = 5$, i $S =_{\rm d} 1.2X$, gdzie $X =_{\rm d}$ Par(2.2),$T_{\max} = 1000$,

6. $\lambda = 5$, i $S =_{\rm d} 2.2X$, gdzie $X =_{\rm d}$ Par(3.2),$T_{\max} = 1000$,

We wszystkich tych przykładach $\rho = 5$. A więc we wszystkich przypadkach powinniśmy otrzymać $\bar{l} = 5$. Na rysunkach odpowiednio 2.1-2.6 pokazana jest wysymulwana realizcja $\hat{L}(t)$ w przedziale $1 \leq t \leq T_{\max}$.

Możemy zauważyć, że symulacja 1-sza bardzo szybko zbiega do granicznej wartości 5. Natomiast z 2-gą jest coś dziwnego i dlatego warto ten przypadek bliżej przeanalizować. Policzymy najpierw $\bar{\sigma}^2$ ze wzoru (V.2.1). Mamy

$$R(t) = 5 \int_t^\infty (1+s)^{-1.2}\, ds = \frac{5}{0.2}(1+t)^{-0.2}$$

skąd mamy, że

$$\bar{\sigma}^2 = 2 \int_0^\infty R(t)\, dt = \infty\ .$$

Nasze założenie o bezwzględnej całkowalności $R(t)$ nie jest spełnione. Nie można zatem stosować fundamentalnego związku, chociaż oczywiście ten estymator zbiega do 5. Zobaczymy to na dalszych przykładach na rys. 2.3,2.4 (symulacja 3-cia i 4-ta). W przypadku sumulacji 5-tej mamy $\alpha = 2.2$ i wtedy $S = 1.2X$, gdzie $X =_{\rm d}$Par(2.2). Stąd

$$R(t) = 5 \int_t^\infty (1+\frac{s}{1.2})^{-2.2}\, ds = 5(1+\frac{s}{1.2})^{-1.2}\,.$$

Widzimy więc, ze $\bar{\sigma}^2 = 2\int_0^\infty R(t)\, dt = 60$, skąd mamy że aby $b = 0.5$ (jedna cyfra wiodąca) wystarczy $T_{\max} = 922$. Należy ostrzec, że nie wzielismy pod uwagę rozpędówki, przyjmując do obliczeń, że proces $L(t)$ jest stacjonarny. Jesli $\alpha = 3.2$ i wtedy $S = 2.2X$, gdzie $X =_{\rm d}$Par(3.2). Stąd

$$R(t) = 5 \int_t^\infty (1+\frac{s}{2.2})^{-3.2}\, ds = 5(1+\frac{s}{2.2})^{-2.2}\,.$$

Widzimy więc, ze $\bar{\sigma}^2 = 2\int_0^\infty R(t)\, dt = 18.3$, skąd mamy że aby $b = 0.5$ (jedna cyfra wiodąca) wystarczy $T_{\max} = 281$. Natomiast aby $b = 0.05$ to musimy mięc $T_{\max}$ większe od 28100.

Rysunek 2.1: $\lambda = 5$, Exp(1), $T_{\max} = 1000$; $\bar{l} = 5.0859$

Rysunek 2.2: $\lambda = 5$, Par(1.2) $\times$ 0.2,$T_{\max} = 1000$; $\bar{l} = 3.2208$

Rysunek 2.3: $\lambda = 5$, Par(1.2) $\times$ 0.2,$T_{\max} = 10000$; $\bar{l} = 3.9090$

Rysunek 2.4: $\lambda = 5$, Par(1.2) $\times$ 0.2, $T_{\max} = 100000$; $\bar{l} = 4.8395$

Rysunek 2.5: $\lambda = 5$, Par(2.2) $\times$ 1.2, $T_{\max} = 1000$; $\bar{l} = 4.8775$

Rysunek 2.6: $\lambda = 5$, Par(3.2) $\times$ 2.2, $T_{\max} = 1000$; $\bar{l} = 5.1314$

2.1 Symulacja gdy X jest procesem regenerującym się.

Będziemy zakładać następującą strukturę procesu stochastycznego X. Przypuśćmy, że istnieją chwile losowe $0 \leq T_1 < T_2 < \ldots$ takie, że pary zmiennych losowych $((Z_i, C_i))_{i\geq 1}$ są niezależne i począwszy od numeru $i \geq 2$ niezależne, gdzie

$$Z_0 = \int_0^{T_1} X(s)\, ds, Z_1 = \int_{T_1}^{T_2} X(s)\, ds, \ldots$$

oraz $C_0 = T_1, C_1 = T_2 - T_1, \ldots$. Takie procesy nazywamy procesami *regenerującymi się*. Wtedy mamy następujący rezultat (patrz Asmussen & Glynn [1]):

Fakt 2.2 *Z prawdopodobieństwem 1, dla* $t \to \infty$

$$\hat{X}(t) \to I = \frac{\mathbb{E}\int_{T_1}^{T_2} X(s)\, ds}{\mathbb{E}\, C_1}, \tag{2.2}$$

oraz

$$t^{1/2}(\hat{X}(t) - I) \to_{\mathrm{d}} \mathcal{N}(0, \bar{\sigma}^2). \tag{2.3}$$

Wtedy TAVC wynosi

$$\bar{\sigma}^2 = \frac{\mathbb{E}\,\zeta_1^2}{\mathbb{E}\, Z_1}$$

gdzie

$$\zeta_k = Z_k - IC_k\ .$$

Patrząc na postać (2.2) widzimy, że można też uzyc innego estymatora, który jest krewniamiem estymatora $\hat{X}$. Mianowicie możemy wziąść

$$\hat{I}_k = \frac{Z_1 + \ldots + Z_k}{C_1 + \ldots + C_k}.$$

Jest oczywiste, że z prawdopdobieństwem 1

$$\hat{I}_k \to I$$

ponieważ

$$\frac{Z_1 + \ldots + Z_k}{C_1 + \ldots + C_k} = \frac{(Z_1 + \ldots + Z_k)/k}{(C_1 + \ldots + C_k)/k} \to \frac{\mathbb{E}\, Z_1}{\mathbb{E}\, C_1}\ .$$

Wtedy mamy następujące twierdzenie (patrz Asmussen & Glynn, Prop. 4.1).

Twierdzenie 2.3 *Dla* $k \to \infty$

$$k^{1/2}(\hat{I}_k - I) \to_{\mathrm{d}} \mathcal{N}(0, \eta^2),$$

gdzie

$$\eta^2 = \frac{\mathbb{E}\,\zeta_1^2}{\mathbb{E}\,C_1}$$

i

$$\zeta_1 = Z_1 - IC_1\ .$$

Stąd asymptotyczny $100(1-\epsilon)\%$ *przedział ufności jest dany przez*

$$\hat{I}_k \pm z_{1-\epsilon/2}\hat{\eta}^2/\sqrt{k}$$

gdzie

$$\hat{\eta}^2 = \frac{1}{k-1}\sum_{l=1}^{k}(C_l - \hat{I}_k Z_l)^2/(\frac{1}{k}\sum_{l=1}^{k} C_l)\ .$$

Przykład 2.4 Proces $L(t)$ w M/M/1 jest regenerujący ponieważ

$$T_1 = \inf\{t > 0 : L(t-) > 0, L(t) = 0\}$$

oraz

$$T_{n+1} = \inf\{t > T_n : L(t-) > 0, L(t) = 0\}$$

definiuje momenty regeneracji procesu $L(t)$. Wynika to z własności braku pamięci rozkładu wykładniczego. Po osiągnięciu przez $L(t)$ zera, niezależnie jak długo czekaliśmy na nowe zgłoszenie, czas do nowego zgłoszenia ma rozkład wykładniczy Exp(λ).

2.2 Przepisy na szacowanie rozpędówki

W wielu przypadkach proces po *pewnym czasie* jest już w przybliżeniu w warunkach stacjonarności. Następujący przykład ilustruje ten pomysł w sposób transparentny.

Przykład 2.5 Z teorii kolejek wiadomo, że jeśli w systemie M/M/1 wystartujemy w chwili $t = 0$ z rozkładu geometrycznego Geo(ρ), to proces liczby zadań $L_1(t)$ jest stacjonarny. Dla uproszczenia wystartujmy proces niestacjonarny z $L_2(0) = 0$. Zakładamy przypadek $\rho < 1$. Wtedy proces $L_1(t)$

osiąga stan 0 w skończonym czasie. Ponieważ oba procesy mają tylko skoki ± 1, więc mamy że procesy $L_1(t)$ i $L_2(t)$ muszą się spotkać, tzn.

$$T = \inf\{t > 0 : L_1(t) = L_2(t)\}$$

jest skończone z prawdopodobieństwem 1. Od tego momentu procesy ewoluują jednakowo. To pokazuje, że od moment T proces $L_2(t)$ jest stacjonarny. Jednakże jest to moment losowy.

Bardzo ważnym z praktycznych względów jest ustalenie poczatkowego okresu s który nie powinniśmy rejestrować. Ten okres nazywamy *długością rozpędówki* lub krótko *rozpędówką* (warm-up period). W tym celu definiujemy nowy estymator

$$\hat{X}(s,t) = \frac{\int_s^t X(v)\,dv}{t-s}\ .$$

Jedna realizacja – metoda średnich pęczkowych Dzielimy odcinek symulacji $[0,t]$ na l pęczków jednakowej długości t/l i zdefiniujmy *średnie pęczkowe*

$$\hat{X}_k(t) = \hat{X}((k-1)t/l, kt/l) = \frac{1}{t/l}\int_{(k-1)t/l}^{kt/l} X(s)\,ds,$$

gdzie $k = 1, 2, \ldots$. Jeśli t/l jest odpowiednio duże, to można przyjąć, że

(P1) $\hat{X}_k(t)$ $(k = 1, \ldots, l)$ ma w przybliżeniu rozkład normalny $\mathcal{N}(I, \bar{\sigma}^2 l/t)$; $(k = 1, \ldots, l)$,

(P2) oraz $\hat{X}_k(t)$ $(k = 1, \ldots, l)$ są niezależne.

Ponieważ

$$\hat{X}(t) = \frac{\hat{X}_1(t) + \ldots + \hat{X}_l(t)}{l}$$

więc dla $t \to \infty$

$$l^{1/2}\frac{\hat{X}(t) - I}{\hat{s}_l(t)} \to_d T_{l-1}$$

gdzie T_{l-1} zmienną losową o rozkładzie t-Studenta z $l-1$ stopniami swobody i

$$\hat{s}_l^2(t) = \frac{1}{l-1}\sum_{i=1}^{l}(\hat{X}_i(t) - \hat{X}(t))^2\ .$$

A więc jeśli $t_{1-\epsilon/2}$ jest $1-\epsilon/2$ kwantylem zmiennej losowej T_{l-1}, to asymptotyczny przedział ufności dla I jest

$$\hat{Y}(t) \pm t_{1-\epsilon/2}\frac{\hat{s}_l(t)}{\sqrt{l}} \ .$$

Asmussen & Glynn sugerują aby $5 \leq l \leq 30$.

Przypuśćy, na chwilę, że proces $X(t)$ jest stacjonarny (będzie to nasza hipoteza zerowa H_0) i taki, że $\frac{\sqrt{t}}{\bar{\sigma}}(\hat{X}(t-)I)$ zbiega wg rozkładu do $\mathcal{N}(0,1)$, gdzie $t\mathrm{Var}\,(\int_0^t X(s)\,ds) \to \bar{\sigma}^2$ gdy $t \to \infty$ dla pewnej skończonej liczby $\bar{\sigma}^2 > 0$. Przyjmując powyższe jako hipotezę zerową możemy testować następujcą hipotezę. Formalizując mamy

H_0. Zmienne $\hat{X}_k(t)$ $(k = 1, \ldots, l)$ są niezależne o jednakowym rozkładzie normalnym $\mathcal{N}(I, \bar{\sigma}^2 l/t)$; $(k = 1, \ldots, l)$.

Wybieramy $l_1, l_2 \geq 0$ takie, że $l_1 + l_2 = l$. Niech $\hat{Y}_1$, $\hat{Y}_2$ będą odpowiednio średnimi arytmetycznymi z $Y_1, \ldots, Y_{l_1}$ i $Y_{l_1+1}, \ldots, Y_l$ oraz $\hat{S}_1 = \frac{\sum_{i=1}^{l_i}(Y_i - \hat{Y}_1)^2}{l_1 - 1}$, $\hat{S}_2 = \frac{\sum_{i=l_1+1}^{l}(Y_i - \hat{Y}_2)^2}{l_2 - 1}$, wariancjami próbkowymi. Wtedy przy założeniu hipotezy zerowej $\hat{S}_i$ mają rozkład chi-kwadrat $\chi^2_{l_i-1}$ z $l_i - 1$ stopniami swobody natomiast $\hat{S}_1/\hat{S}_2$ ma rozkład F_{l_1-1,l_2-1} Snedecora.

A więc efekty symulacji podczas rozpędówki nie musimy rejestrować a dopiero po tym czasie zaczynamy rejestrować i następnie opracować wynik jak dla procesu stacjonarnego. Problemem jest, że nie wiemy co to znaczy po *pewnym czasie*. Dlatego też mamy następujący przepis.

Metoda niezależnych replikacji Fishman [3, 4] zaproponował następującą graficzną metodę szacowania rozpędówki. Zamiast symulowanie jednej realizacji $X(t)$ $(0 \leq t \leq T_{\max})$ symulujemy teraz n realizacji $X^i(t)$, $(0 \leq t \leq T_{\max})$, gdzie $i = 1, \ldots, n$. Każdą zaczynamy od początku ale z rozłącznymi liczbami losowymi. Niech

$$\hat{X}^i(s,t) = \frac{\int_s^t X^i(v)\,dv}{t-s} \ .$$

będzie opóźnioną średnią dla i-tej replikacji realizacji procesu $(i = 1, \ldots, n)$. Niech

$$\hat{X}^{\cdot}(t) = \frac{1}{n}\sum_{i=1}^{n} X^i(t), \qquad 0 \leq t \leq T_{\max}$$

oraz

$$\hat{X}^{\cdot}(t, T_{\max}) = \frac{1}{n} \sum_{i=1}^{n} \frac{\int_t^{T_{\max}} X^i(v)\, dv}{T_{\max} - t}, \qquad 0 \le t \le T_{\max}\ .$$

Jeśli proces X jest stacjonarny to obie krzywe $\hat{X}^{\cdot}(t)$ i $\hat{X}^{\cdot}(t, T_{\max})$ powinny mieć podobne zachowanie. Pomijamy fluktuacje spowodowane niedużym n dla $\hat{X}^{\cdot}(t)$ oraz krótkością odcinka dla $\hat{X}^{\cdot}(t, T_{\max})$. Dlatego porównując na jednym wykresie

$$(\hat{X}^{\cdot}(t), \hat{X}^{\cdot}(t, T_{\max})),\ 0 \le t \le T_{\max}$$

możemy oszacować jakie s należy przyjąć.

2.3 Efekt stanu początkowego i załadowania systemu

Na długośc rozpędówki mają istotny wpływ przyjęty stan początkowy symulacji oraz załadowanie systemu.

Przykład 2.6 Tytułowy problem zilustrujemy za pracą Srikanta i Whitta [10] o symulacji prawdopodobieństwa straty w systemie M/M/c ze stratami. ZAdania wpływają z intensywnością λ natomiast ich śreni rozmiar wynosi μ^{-1}. Jak zwykle oznaczamy $\rho = \lambda/\mu$. Dla tego systemu oczywiście znamy to prawdopodobieństwo jak i wiele innych charakterystyk, co pozwoli nam zrozumieć problemy jakie mogą występować przy innych symulacjach. Zacznijmy od przeglądu kandydatów na estymatory. Jeśli oznaczymy przez

$$B_{\mathrm{N}}(t) = \frac{B(t)}{A(t)},$$

gdzie $B(t)$ jest liczba straconych zadań do chwili w odcinku $[0, t]$ natomiast $A(t)$ jest liczbą wszystkich zadań (straconych i zaakceptowanych) które wpłynęły do chwili t. Estymator ten nazwiemy *estymatorem naturalnym*. Innym estymatorem jest tak zwany *prosty estymator*

$$B_{\mathrm{S}}(t) = \frac{B(t)}{\lambda t},$$

lub *niebezpośredni estymator* (indrirect estimator)

$$B_{\mathrm{I}}(t) = 1 - \frac{\hat{L}(t)}{\rho}.$$

Podamy teraz krótkie uzasadnienia dla tych estymatorów. Czy są one zgodne? Oczywiście ponieważ dla B_{N} mamy

$$\begin{aligned} p_{\mathrm{block}} &= \lim_{k\to\infty} \sum_{j=1}^{k} I_j/k = \frac{\rho^c/c!}{\sum_{j=0}^{c} \rho^j/j!} \\ &= \lim_{t\to\infty} \frac{B(t)}{A(t)} \, . \end{aligned}$$

Podobnie wygląda sytuacja z estymatorem prostym B_{S}, ponieważ

$$\begin{aligned} \lim_{t\to\infty} \frac{B(t)}{A(t)} &= \lim_{t\to\infty} \frac{B(t)}{t} \frac{t}{A(t)} \\ &= \lim_{t\to\infty} \frac{B(t)}{\lambda t} \, . \end{aligned}$$

Natomiast dla estymatora niebezpośredniego aby pokazać zgodnośc pokażemy tzw wzór Little'a [1]:

$$\hat{l} = \lambda(1 - p_{\mathrm{block}})/\mu \, .$$

Jeśli natomiast by założyć, ze system jest w warunkach stacjonarności, to można pokazać, że estymatory prosty i niebezpośredni są zgodne. W pracy Srikanta i Whitta [10] opisano następujący eksperyment z $\lambda = 380$, $\mu = 1$ i $c = 400$. Dla tego systemu wysymulowano, startując z pustego systemu, przy użyciu estymatora niebezpośredniego, z $t_{\max} = 5400$, mamy $p_{\mathrm{block}} = 0.00017$. Tutaj TAVC wynosi około 0.00066. W tym zadaniu można policzyć numerycznie asymptotyczne obciążenie co przedstawione jest w tabl. 2.1 z $\alpha = \rho/c$. Zauważmy, że w przypadku $N(0) = 0$, TAVC jest tego samego

α	$N(0) = 0$	$N(0) = c$
0.7	1.00	-0.42
0.8	1.00	-0.24
0.9	0.99	-0.086
1.0	0.85	-0.0123
1.1	0.66	-0.0019
1.2	0.52	-0.0005

Tablica 2.1: Asymptotyczne obciążenie dla $\hat{B}_{\mathrm{I}}$

rzędu co asymptotyczne obciążenia.

[1] Sprawdzić tożsamość.

W następnym przykładzie zajmiemy się efektem obiążenia.

Przykład 2.7 Pytanie o to kak długi należy dobrać horyzont symulacji $T_{\max}$ aby osiągnąć zadaną precyzję zależy często od obciążenia systemu który symulujemy. W systemie M/M/1 mozna policzyć TAVC $\bar{\sigma}^2$, która to wielkość poprze fundamentalny związek rzutuje na $T_{\max}$. Mianowicie (patrz Whitt [13]) mamy

$$\bar{\sigma}^2 = \frac{2\rho(1+\rho)}{(1-\rho)^4} \, .$$

Oczywiście musimy założyć $\rho < 1$. Widzimy, że jeśli $\rho \uparrow 1$, to $\bar{\sigma}^2$ bardzo szybko rośnie. Stąd, korzystając z fundamentalnego związku, mamy że tak samo szybko rośnie horyzont symulacji aby utrzymać zadaną dokładność.

3 Uwagi bibliograficzne

Cały cykl prac na temat planowania symulacji w teorii kolejek napisał Whitt z współpracownikami; [11, 12, 13, 10]. Tym tematem zajmował się również Grassmann [14]

Rozdział VII

Dodatek

1 Zmienne losowe i ich charakterystyki

1.1 Rozkłady zmiennych losowych

W teorii prawdopodobieństwa wprowadza się zmienną losową X jako funkcję zdefiniowaną na przestrzeni zdarzeń elementarnych Ω. W wileu przypadka, ani Ω and X ni są znane. Ale co jest ważne, znany jest rozkład zmiennej losowej X, tj. przepis jak liczyć $F(B)$, że X przyjmie wartość z zbioru B, gdzie B jest podzbiorem (borelowskim) prostej $\mathbb{R}$.

Ze względu na matematyczną wygodę rozważa się rodzinę $\mathcal{F}$ podzbiorów Ω, zwaną rodziną zdarzeń, oraz prawdopodobieństwo $\mathbb{P}$ na $\mathcal{F}$ przypisujące zdarzeniu A z $\mathcal{F}$ liczbę $\mathbb{P}(A)$ z odcinka $[0,1]$.

Rozkład można opisać prze podanie *dystrybuanty*, tj funkcji $F : \mathbb{R} \to [0,1]$ takiej, że $F(x) = \mathbb{P}(X \leq x)$. Prze $\overline{F}(x) = 1 - F(x)$ będziemy oznaczać *ogon* of F.

W praktyce mamy dwa typy zmiennych losowych: dyskretne i absolutnie ciągłe. Mówimy, że X jest *dyskretną* jeśli jeśli istnieje przeliczalny zbiór liczb $E = \{x_0, x_1, \ldots\}$ z $\mathbb{R}$ taki, że $\mathbb{P}(X \in E) = 1$. W tym przypadku definiujemy *funkcję prawdopodobieństwa* $p : E \to [0,1]$ przez $p(x_k) = \mathbb{P}(X = x_k)$; para (E, p) podaje pełny przepis jak liczyć prawdopopodobieństwa $\mathbb{P}(X \in B)$ a więc zadaje rozkład. Najważniejszą podklasą są zmienne losowe przyjmujące wartości w E które jes podzbiorem kraty $\{hk,\ k \in \mathbb{Z}_+\}$ dla pewnego $h > 0$. Jeśli nie będzie powiedziane inaczej to zakładamy, że $h = 1$, tzn $E \subset \mathbb{Z}_+$, i.e. $x_k = k$. Wtedy piszemy prosto $p(x_k) = p_k$ i mówimy, że rozkład jest *kratowy*.

Z drugiej strony mówimy, że X jest *absolutnie ciągłą* jeśli istnieje funkcja $f : \mathbb{R} \to \mathbb{R}_+$ taka, że $\int f(x)\,dx = 1$ i $\mathbb{P}(X \in B) = \int_B f(x)\,dx$ dla każdego $B \in \mathcal{B}(\mathbb{R})$. Mówimy wtedy, że $f(x)$ jest *gęstością* zmiennej X.

Rozkład dyskretnej zmiennej losowej nazywamy rozkładem dyskretnym. Jeśli X jest kratowy, to jego rozkład też nazywamy kratowy. Analogoicznie rozkład zmiennej losowej absolutnie ciągłej nazywamy rozkładem absolutnie ciągłym. Czasami dystrybuanta nie jest ani dyskretna ani absolutnie ciągła. Przykładem jest *mieszanka* $F = \theta F_1 + (1-\theta)F_2$, gdzie F_1 is dyskretna z prawdopodobieństwem θ i F_2 jest absolutnie ciągła z prawdopodobieństwem $1-\theta$.

Aby zaznaczyć, że dystrybuanta, gęstość, funkcja prawdopodbieństwa, jest dla zmiennej losowej X, piszemy czasem $F_X, p_X, f_X, \ldots$.

Dla dwóch zmiennych losowych X i Y o tym samym rozkładzie piszemy $X \stackrel{\mathrm{d}}{=} Y$. To samo stosujemy do wektoów.

Mówoimy, że wektor losowy $(X_1, \ldots, X_n)$ jest *absolutnie ciągły* jeśli istnieje nieujemna i całkowalna funkcja $f : \mathbb{R}^n \to \mathbb{R}_+$ z

$$\int_{\mathbf{R}} \ldots \int_{\mathbf{R}} f(x_1, \ldots, x_n)\,dx_1 \ldots dx_n = 1$$

taka, że $\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B) = \int_{B_1} \ldots \int_{B_n} f(x_1, \ldots, x_n)\,dx_n \ldots dx_1$ for all Borel sets $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$. Funkcja $f(x)$ nazywa się *gęstością* $\boldsymbol{X} = (X_1, \ldots, X_n)^T$. Przytoczymy teraz pożyteczny rezultat dotyczący wzoru na gęstość przekształconego wektora losowego. Mianowicie niech będą dane funkcje $g_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$, które spełniają następujące warunki:

- (i) przekształcenie $\boldsymbol{g} = (g_1, \ldots, g_n)^T$ jest wzajemnie jednoznaczne na zbiorze A,
- (ii) przekształcenie $\boldsymbol{g}$ ma na A ciągłe pierwsze pochodne cząstkowe ze względu na wszystkie zmienne,
- jakobian $J_{\boldsymbol{g}}(\boldsymbol{x}) = \frac{\partial(g_1,\ldots,g_n)}{\partial(x_1,\ldots,x_n)}$ jest różny od zera na A.

Oznaczmy przez $\boldsymbol{h} = (h_1, \ldots, h_n)^T$ przekształcenie odwrotne do $\boldsymbol{g}$, tzm $\boldsymbol{x} = \boldsymbol{h}(\boldsymbol{y}) = (h_1(\boldsymbol{y}), \ldots, h_n(\boldsymbol{y}))$ dla każdego $\boldsymbol{y} \in \boldsymbol{g}(A)$. Rozważmy wektor $\boldsymbol{Y} = (\boldsymbol{X})$. Wektor ten ma gęstość $f_{\boldsymbol{Y}}$ zadaną wzorem:

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = f_{\boldsymbol{X}}(h_1(y), \ldots, h_n(\boldsymbol{y}))|J_h(\boldsymbol{y})|.$$

Jeśli zmienne losowe $(X_1, \ldots, X_{n-1})$ $(X_1, \ldots, X_n)$ są absolutnie ciągłe z odpowiednio z gęstościami $f_{X_1,\ldots,X_{n-1}}$ i $f_{X_1,\ldots,X_n}$ i jeśli $f_{X_1,\ldots,X_{n-1}}(x_1, \ldots, x_{n-1}) > 0$, to

$$f_{X_n|X_1,\ldots,X_{n-1}}(x_n \mid x_1, \ldots, x_{n-1}) = \frac{f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)}{f_{X_1,\ldots,X_{n-1}}(x_1, \ldots, x_{n-1})}$$

nazywa się i *warunkową gęstością* X_n pod warunkiem $X_1 = x_1, \ldots, X_{n-1} = x_{n-1}$.

Dla ustalonego n i wszystkich $k = 1, 2, \ldots, n$ i $\omega \in \Omega$, niech

$X_{(k)}(\omega)$ oznacza k-th najmniejszą wartośyć $X_1(\omega), \ldots, X_n(\omega)$. Składowe wektora losowego $(X_{(1)}, \ldots, X_{(n)})$ nazywają się *statystyka porządkową* $(X_1, \ldots, X_n)$.

1.2 Podstawowe charakerystyki

Niech X będzie zmienną losową. i $g : \mathbb{R} \to \mathbb{R}$ funkcją. Definiujemy $\mathbb{E}\, g(X)$ przez

$$\mathbb{E}\, g(X) = \begin{cases} \sum_k g(x_k)p(x_k) & \text{jeśli } X \text{ jest dyskretny} \\ \int_{-\infty}^{\infty} g(x)f(x)\,dx & \text{jeśli } X \text{ jest absolutnie ciągłe} \end{cases}$$

pod warunkiem odpowiednio, że $\sum_k |g(x_k)|p(x_k) < \infty$ i $\int_{-\infty}^{\infty} |g(x)|f(x)\,dx < \infty$. Jeśli g jest nieujemna, to używamy symbolu $\mathbb{E}\, g(X)$ również gdy równa się nieskończoność. Jeśli rozkład jest mieszanką dwóch typów to definiujemy $\mathbb{E}\, g(X)$ by

$$\mathbb{E}\, g(X) = \theta \sum_k g(x_k)p(x_k) + (1-\theta) \int_{-\infty}^{\infty} g(x)f(x)\,dx\,. \tag{1.1}$$

Czasami wygodne jest zapisanie $\mathbb{E}\, g(X)$ w terminach *całki Stieltjesa* t.j.

$$\mathbb{E}\, g(X) = \int_{-\infty}^{\infty} g(x)dF(x)\,,$$

która jest dobrze określona przy pewnych warunkach na g i F. W szczególności dla $g(x) = x$ wartość $\mu = \mathbb{E}\, X$ jest nazywana *średnią* , lub *wartością oczekiwaną* lub *pierwszy moment* X. Dla $g(x) = x^n$, $\mu^{(n)} = \mathbb{E}\,(X^n)$ jest nazywana n-tym *momentem*. *Wariancją* X jest $\sigma^2 = \mathbb{E}\,(X-\mu)^2$ i $\sigma = \sqrt{\sigma^2}$ jest *odchyleniem standardowym* lub *dyspersją* . Czasami piszemy $\operatorname{Var} X$ zamiast σ^2. *współczynnik zmienności* jest zdefiniowany przez $\operatorname{cv}_X = \sigma/\mu$, i *współczynnik asymetrii* przez $\mathbb{E}\,(X-\mu)^3\sigma^{-3}$. Dla dwóch zmiennych losowych X, Y dfiniujemy we *kowariancję* $\operatorname{Cov}(X,Y)$ przez $\operatorname{Cov}(X,Y) = \mathbb{E}\,((X-\mathbb{E}\,X)(Y-\mathbb{E}\,Y))$ jeśli tylko $\mathbb{E}\,X^2, \mathbb{E}\,Y^2 < \infty$. zauważmy, że równoważnie mamy $\operatorname{Cov}(X,Y) = \mathbb{E}\,XY - \mathbb{E}\,X\,\mathbb{E}\,Y$. Jeśli $\operatorname{Cov}(X,Y) > 0$, to mówimy, że X, Y are *dodatnio skorelowane*, jeśli $\operatorname{Cov}(X,Y) < 0$ to *ujemnie skorelowane* i *nieskorelowane* jeśli $\operatorname{Cov}(X,Y) = 0$. Przez *współczynnik korelacji* pomiędzy X, Y rozumiemy

$$\operatorname{corr}(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}.$$

Medianą zmienej losowej X jest dowolna liczba is $\zeta_{1/2}$ taka, że

$$\mathbb{P}(X \le \zeta_{1/2}) \ge \tfrac{1}{2}\,, \qquad \mathbb{P}(X \ge \zeta_{1/2}) \ge \tfrac{1}{2}\,.$$

Przez indykator $\mathbf{1}(A) : \Omega \to \mathbb{R}$ rozumiemy zmienną losową

$$\mathbf{1}(A, \omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

A więc jeśli $A \in \mathcal{F}$ to $\mathbf{1}(A)$ jest zmienną losową i $\mathbb{P}(A) = \mathbb{E}\,\mathbf{1}(A)$. Używa się nastepującej konwencji, że $\mathbb{E}\,[X; A] = \mathbb{E}\,(X\mathbf{1}(A))$ dla zmiennej losowej X i zdarzenia $A \in \mathcal{F}$.

Jeśli chcemy zwrócić uwagę, że średnia, n-ty moment, itd. są zdefiniowane dla zmiennej X lub dystrybunaty F to piszemy $\mu_X, \mu_X^{(n)}, \sigma_X^2, \ldots$ or $\mu_F, \mu_F^{(n)}, \sigma_F^2, \ldots$.

1.3 Niezależność i warunkowanie

Mówimy, że zmienne losowe $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ są *niezależne* jeśli

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \prod_{k=1}^{n} \mathbb{P}(X_k \in B_k)$$

for all $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ lub, równoważnie,

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \prod_{k=1}^{n} \mathbb{P}(X_k \leq x_k)$$

dla wszystkich $x_1, \ldots, x_n \in \mathbb{R}$. Nieskończony ciąg $X_1, X_2, \ldots$ jest ciągiem niezależnych zmiennych losowych, jeśli każdy skończony podciąg ma tą własność. Mówimy, że dwa ciągi $X_1, X_2, \ldots$ i $Y_1, Y_2, \ldots$ są niezależne jeśli

$$\mathbb{P}\Big(\bigcap_{i=1}^{n}\{X_i \leq x_i\} \cap \bigcap_{i=1}^{m}\{Y_i \leq y_i\}\Big) = \mathbb{P}\Big(\bigcap_{i=1}^{n}\{X_i \leq x_i\}\Big)\mathbb{P}\Big(\bigcap_{i=1}^{m}\{Y_i \leq y_i\}\Big)$$

dla wszystkicj $n, m \in \mathbb{N}$ i $x_1, \ldots, x_n$, $y_1, \ldots, y_m \in \mathbb{R}$. Dla ciągu $X_1, X_2, \ldots$ niezależnych i jednakowo rozłożonych zmiennych losowych wygodne jest wprowadzenie generycznej zmiennej losowej X, która ma taki sam rozkład. Mówimy, że zmienna losowa jest niezależna od zdarzenie, jeśli jest niezależna od indykatora $\mathbf{1}(A)$ tego zdarzenia.

Niech $A \in \mathcal{F}$. *Warunkowym prawdopodobieństwem* zadarzenia A' pod warunkiem A jest

$$\mathbb{P}(A' \mid A) = \begin{cases} \mathbb{P}(A' \cap A)/\mathbb{P}(A) & \text{if } \mathbb{P}(A) > 0, \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

Często użyteczne jest wzór na *prawdopodobieństwo całkowite*. Niech $A, A_1, A_2, \ldots \in \mathcal{F}$ będzie ciągiem zdarzeń taki, że $A_i \cap A_j = \emptyset$ dla $i \neq j$ i $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = 1$. Wtedy,

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A \mid A_i)\, \mathbb{P}(A_i)\,. \tag{1.2}$$

Warunkową wartością oczekiwaną $\mathbb{E}\,(X \mid A)$ zmiennej losowej X pod warunkiem zdarzenia A jest warunkową wartością oczekiwaną względem rozkładu warunkowego z dystrybunatą $F_{X|A}$, where $F_{X \mid A}(x) = \mathbb{P}(\{X \leq x\} \cap A)/\mathbb{P}(A)$.

1.4 Transformaty

Niech $I = \{s \in \mathbb{R} : \mathbb{E}\, e^{sX} < \infty\}$. Zauważmy, że I musi być odcinkiem (właściwym) lub prostą półprostą lub nawet punktem $\{0\}$. *Funkcją tworzącą momenty* $\hat{m} : I \to \mathbb{R}$ zmiennej losowej X jest zdefiniowana przez $\hat{m}(s) = \mathbb{E}\, e^{sX}$. Zauważmy różnice pomiędzy funkcją tworzącą momenty a *trasformatą Laplace-Stieltjesa* $\hat{l}(s) = \mathbb{E}\, e^{-sX} = \int_{-\infty}^{\infty} e^{-sx}\, dF(x)$ zmiennej losowej X lub dystrybuanty F zmiennej X.

Dla zmiennych kratowych z funkcją prawdopodobie/nstwa $\{p_k,\ k \in \mathbb{N}\}$ wprowadza się *funkcję tworzącą* $\hat{g} : [-1, 1] \to \mathbb{R}$ zdefiniowaną przez $\hat{g}(s) = \sum_{k=0}^{\infty} p_k s^k$.

Niech teraz X będzie dowolną zmienną losową z dystrybuantą F. *Funkcją charakterystyczną* $\hat{\varphi} : \mathbb{R} \to \mathbb{C}$ zmiennej X jest zdefiniowna przez

$$\hat{\varphi}(s) = \mathbb{E}\, e^{isX}. \tag{1.3}$$

W szczególności jesłi F jest absolutnie ciągła z gęstością f, to

$$\hat{\varphi}(s) = \int_{-\infty}^{\infty} e^{isx} f(x)\, dx = \int_{-\infty}^{\infty} \cos(sx) f(x)\, dx + i \int_{-\infty}^{\infty} \sin(sx) f(x)\, dx\,.$$

Jeśli X jest kratową zmienną losową z funkcją prawdopodobie/nstwa $\{p_k\}$, to

$$\hat{\varphi}(s) = \sum_{k=0}^{\infty} e^{isk} p_k\,. \tag{1.4}$$

Jeśli chcemy podkreśli, że transformata jest zmiennej X z dydtrybuantą F, to piszemy $\hat{m}_X, \ldots, \hat{\varphi}_X$ lub $\hat{m}_F, \ldots, \hat{\varphi}_F$.

Zauważmy, też formalne związki $\hat{g}(e^s) = \hat{l}(-s) = \hat{m}(s)$ lub $\hat{\varphi}(s) = \hat{g}(e^{is})$. Jednakże delikatnym problemem jest zdecydowanie, gdzie dana transformata jest określona. Na przykład jeśli X jest nieujemna, to $\hat{m}(s)$ jest zawsze dobrze zdefiniowana przynajmiej dla $s \le 0$ podczes gdy $\hat{l}(s)$ jest dobrze zdefiniowana przynajmniej dla $s \ge 0$. Są też przykłady pokazujące, że $\hat{m}(s)$ może bć ∞ dla wszystkich $s > 0$, podczas gdy w innych przypadkach, że $\hat{m}(s)$ jest skończone $(-\infty, a)$ dla pewnych $a > 0$.

Ogólnie jest wzajemna jednoznaczność pomiędzy rozkładem a odpowiadającą jemu transformatą, jednakże w każdym wypadku sformułowanie twierdzenia wymaga odpowiedniego wysłowienia. Jeśli n-ty moment zmiennej losowej $|X|$ jest sko/nczony, to n-ta pochodna $\hat{m}^{(n)}(0), \hat{l}^{(n)}(0)$ istnieje jeśli tylko $\hat{m}(s), \hat{l}(s)$ są dobrze zdefiniowane w pewnym otoczeniu $s = 0$. Wtedy mamy

$$\mathbb{E}\, X^n = \hat{m}^{(n)}(0) = (-1)^n \hat{l}^{(n)}(0) = (-i)^n \hat{\varphi}^{(n)}(0)\,. \tag{1.5}$$

Jeślo $\hat{m}(s)$ i $\hat{l}(s)$ są odpowienio dobrze zdefiniowane na $(-\infty, 0]$ i $[0, \infty)$, to trzeba pochodne w (1.5) zastąpić pochodnymi jednostronnymi, t.j.

$$\mathbb{E}\, X^n = \hat{m}^{(n)}(0-) = (-1)^n \hat{l}^{(n)}(0+)\,, \tag{1.6}$$

Jeśli X przyjmuje wartośc w podzbiorze $\mathbb{Z}_+$ i jeśli $\mathbb{E}\, X^n < \infty$, to

$$\mathbb{E}\, \left(X(X-1)\dots(X-n+1)\right) = \hat{g}^{(n)}(1-)\,. \tag{1.7}$$

Inną ważną własnością $\hat{m}(s)$ (i również dla $\hat{l}(s)$ i $\hat{\varphi}(s)$) jest własność , że jeśli $X_1, \dots, X_n$ są niezależne to

$$\hat{m}_{X_1+\dots+X_n}(s) = \prod_{k=1}^{n} \hat{m}_{X_k}(s)\,. \tag{1.8}$$

Podobnie, jeśli $X_1, \dots, X_n$ są niezależny zmienne kratowe z $\mathbb{Z}_+$ to

$$\hat{g}_{X_1+\dots+X_n}(s) = \prod_{k=1}^{n} \hat{g}_{X_k}(s)\,. \tag{1.9}$$

Ogólnie dla dowolnych funkcji $g_1, \dots, g_n : \mathbb{R} \to \mathbb{R}$ mamy

$$\mathbb{E}\left(\prod_{k=1}^{n} g_k(X_k)\right) = \prod_{k=1}^{n} \mathbb{E}\, g_k(X_k) \tag{1.10}$$

jeśli tylko $X_1, \dots, X_n$ są niezleżne.

Niech $X, X_1, X_2, \dots$ będą zmiennymi losowymi. Mówimy, że ciąg $\{X_n\}$ *zbiega słabo* lub *zbiega według rozkładu*) do X jeśli $F_{X_n}(x) \to F_X(x)$ dla wszystkich punktów ciągłości F_X. Piszemy wtedy $X_n \xrightarrow{\mathrm{d}} X$.

2 Rodziny rozkładów

Podamy teraz kilka klas rozkładów dyskretnych i absolutnie ciągłych. Wszystkie są scharakteryzowane przez skończony zbiór parametrów. Podstawowe charakterystyki podane są w tablicy rozkładów w podrozdziale 3.

W zbiorze rozkładów absolutnie ciągłych wyróżniamy dwie podklasy: rozkłady z lekkimi ogonami i ciężkimi ogonami. Rozróżnienie to jest bardzo ważne w teorii Monte Carlo.

Mówimy, że rozkład z dystrybuntą $F(x)$ ma wykładniczo ograniczony ogon jeśli dla pewnego $a, b > 0$ mamy $\overline{F}(x) \leq a\mathrm{e}^{-bx}$ dla wszystkich $x > 0$. Wtedy mówimy też że rozkład ma *lekki ogon*. Zauważmy, że wtedy $\hat{m}_F(s) < \infty$ dla $0 < s < s_0$, gdzie $s_0 > 0$. W przeciwnym razie mówimy, że rozkład ma *ciężki ogon* a więc wtedy $\hat{m}_F(s) = \infty$ dla wszystkich $s > 0$.

Uwaga Dla dystrybunaty z ciązkim ogonem F mamy

$$\lim_{x\to\infty} \mathrm{e}^{sx}\overline{F}(x) = \infty \tag{2.1}$$

dla wszystkich $s > 0$.

2.1 Rozkłady dyskretne

Wśród dyskretnych rozkładów mamy:

- *Rozkład zdegenerowany* δ_a skoncentrowany w $a \in \mathbb{R}$ z $p(x) = 1$ jeśli $x = a$ i $p(x) = 0$ w przeciwnym razie.
- *Rozkład Bernoulliego* $\mathrm{Ber}(p)$ z $p_k = p^k(1-p)^{1-k}$ dla $k = 0, 1$; $0 < p < 1$: rozkład przyjmujący tylko wartości 0 i 1.
- *Rozkład dwumianowy* $\mathrm{B}(n, p)$ z $p_k = \binom{n}{k}p^k(1-p)^{n-k}$ dla $k = 0, 1 \ldots, n$; $n \in \mathbb{Z}_+, 0 < p < 1$: rozkład sumy n niezależnych i jednakowo rozóżonych zmiennych losowych o rozkładzie $\mathrm{Ber}(p)$. A więc , $\mathrm{B}(n_1, p) * \mathrm{B}(n_2, p) = \mathrm{B}(n_1 + n_2, p)$.
- *Rozkład Poissona* $\mathrm{Poi}(\lambda)$ z $p_k = \mathrm{e}^{-\lambda}\lambda^k/k!$ dla $k = 0, 1 \ldots$; $0 < \lambda < \infty$: jedna z podstawowych cegiełek na których zbudowany jest rachunek prawdopodobieństwa. Historycznie pojawił się jako graniczny dla rozkładów dwumianowych $\mathrm{B}(n, p)$ gdy $np \to \lambda$ dla $n \to \infty$. Ważna jest własność zamkniętości ze względu na splot, tzn. suma niezależnych zmiennych Poissonowskich ma znowu rozkład Poissona. Formalnie piszemy, że $\mathrm{Poi}(\lambda_1) * \mathrm{Poi}(\lambda_2) = \mathrm{Poi}(\lambda_1 + \lambda_2)$.

- *Rozkład geometryczny* $\mathrm{Geo}(p)$ z $p_k = (1-p)p^k$ dla $k = 0, 1, \ldots;\ 0 < p < 1$: rozkład mający dyskretny brak pamięci. To jest $\mathbb{P}(X \geq i + j \mid X \geq j) = \mathbb{P}(X \geq i)$ dla wszystkih $i, j \in \mathbb{N}$ wtedy i tylko wtedy gdy X jest geometrycznie rozłożony.

- *Rozkład obciąty geometryczny* $\mathrm{TGeo}(p)$ z $p_k = (1-p)p^{k-1}$ dla $k = 1, 2, \ldots;\ 0 < p < 1$: jeśli $X \sim \mathrm{Geo}(p)$, to $X|X > 0 \sim \mathrm{TGeo}(p)$, jest to rozkład do liczby prób Bernoulliego do pierszego sukcesy, jeśli prawdopobieństwo sukcesu wynosi $1 - p$.

- *Rozkład ujemny dwumianowy* lub inaczej *rozkład Pascala* $\mathrm{NB}(\alpha, p)$ z $p_k = \Gamma(\alpha + k)/(\Gamma(\alpha)\Gamma(k+1))(1-p)^\alpha p^k$ dla $k = 0, 1, \ldots;\ \alpha > 0, 0 < p < 1$. Powyżej piszemy

$$\binom{x}{0} = 1, \qquad \binom{x}{k} = \frac{x(x-1)\ldots(x-k+1)}{k!},$$

  dla $x \in \mathbb{R}$, $k = 1, 2, \ldots$. Mamy

$$p_k = \binom{\alpha + k - 1}{k}(1-p)^\alpha p^k = \binom{-\alpha}{k}(1-p)^\alpha(-p)^k.$$

  Co więcej, dla $\alpha = 1, 2, \ldots$, $\mathrm{NB}(\alpha, p)$ jest rozkładem sumy α niezależnych i jednakowo rozłożonych $\mathrm{Geo}(p)$ zmiennych losowych. To oznacza na przykład, że podklasa ujemnych rozkładów dwumianowych $\{\mathrm{NB}(\alpha, p), \alpha = 1, 2, \ldots\}$ jest zamknięta ze względu na splot t.j. $\mathrm{NB}(\alpha_1, p) * \mathrm{NB}(\alpha_2, p) = \mathrm{NB}(\alpha_1 + \alpha_2, p)$ if $\alpha_1, \alpha_2 = 1, 2, \ldots$. co więcej poprzedni wzór jest spełniny dla $\alpha_1, \alpha_2 > 0$ and $0 < p < 1$.

- *Rozkład logarytmiczny* $\mathrm{Log}(p)$ z $p_k = p^k/(-k\log(1-p))$ for $k = 1, 2, \ldots;\ 0 < p < 1$: pojawia się jako granica obciętych rozkładów dwumianowych.

- *Rozkład (dyskretny) jednostajny* $\mathrm{UD}(n)$ z $p_k = n^{-1}$ dla $k = 1, \ldots, n$; $n = 1, 2, \ldots$.

2.2 Rozkłady absolutnie ciągłe

Wśród rozkładów absolutnie ciągłych mamy:

- *Rozkład normalny* $\mathrm{N}(\mu, \sigma^2)$ z $f(x) = (2\pi\sigma^2)^{-1/2}\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}$ dla wszystkich $x \in \mathbb{R}$; $\mu \in \mathbb{R}, \sigma^2 > 0$: obok rozkładu Poissona druga cegiełka

na której opiera się rachunek prawdopodobieństwa. Pojawia się jako granica sum niezalażnych asymptotycznie zaniedbywalnych zmiennych losowych. Rzokłady normalne są zamnięte na branie splotów, tj. $\mathrm{N}(\mu_1, \sigma_1^2) * \mathrm{N}(\mu_2, \sigma_2^2) = \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

- *Rozkład wykładniczy* $\mathrm{Exp}(\lambda)$ z $f(x) = \lambda \mathrm{e}^{-\lambda x}$ dla $x > 0;\ \lambda > 0$: podstawowy rozkład w teorii procesów Markowa ze względu na własność braku pamięci, t.j. $\mathbb{P}(X > t + s \mid X > s) = \mathbb{P}(X > t)$ dla wszystkich $s, t \geq 0$ jeśli X jest rozkładem wykładniczym.

- *Rozkład Erlanga* $\mathrm{Erl}(n, \lambda)$ z $f(x) = \lambda^n x^{n-1} \mathrm{e}^{-\lambda x}/(n-1)!;\ n = 1, 2, \ldots$: rozkład sumy n niezależnych i jednakowo rozłożonych $\mathrm{Exp}(\lambda)$ zmiennych losowych. Tak więc, $\mathrm{Erl}(n_1, \lambda) * \mathrm{Erl}(n_2, \lambda) = \mathrm{Erl}(n_1 + n_2, \lambda)$.

- *Rozkład chi kwadrat* $\chi^2(n)$ z $f(x) = (2^{n/2}\Gamma(n/2))^{-1} x^{(n/2)-1} \mathrm{e}^{-x/2}$ if $x > 0$ and $f(x) = 0$ if $x \leq 0$; $n = 1, 2, \ldots$: Rozkład sumy n niezależnych i jednakowo rozłożonych zmiennych losowych, które są kwadratami $\mathrm{N}(0, 1)$ zmiennych losowych To oznacza, że $\chi^2(n_1) * \chi^2(n_2) = \chi^2(n_1 + n_2)$.

- *Rozkład gamma* $\Gamma(a, \lambda)$ z $f(x) = \lambda^a x^{a-1} \mathrm{e}^{-\lambda x}/\Gamma(a)$ for $x \geq 0$; z paametrem kształtu $a > 0$ i parametreem skali $\lambda > 0$: jeśli $a = n \in \mathbb{N}$, to $\Gamma(n, \lambda) = \mathrm{Erl}(n, \lambda)$. Jeśli $\lambda = 1/2$ i $a = n/2$, to $\Gamma(n/2, 1/2) = \chi^2(n)$.

- *Rozkład jednostajny* $\mathrm{U}(a, b)$ z $f(x) = (b - a)^{-1}$ dla $a < x < b;\ -\infty < a < b < \infty$.

- *Rozkład beta distribution* $\mathrm{Beta}(a, b, \eta)$ z

$$f(x) = \frac{x^{a-1}(\eta - x)^{b-1}}{B(a, b)\eta^{a+b-1}}$$

  da $0 < x < \eta$ gdzie $B(a, b)$ oznacza funkcję beta(partz **[??som.spe.fun??]**); $a, b, \eta > 0$; jeśli $a = b = 1$, to $\mathrm{Beta}(1, 1, \eta) = \mathrm{U}(0, \eta)$.

- *Rozkład odwrotny gaussowski* (*inverse Gaussian distribution*) $\mathrm{IG}(\mu, \lambda)$ z

$$f(x) = \left(\lambda/(2\pi x^3)\right)^{1/2} \exp\left(-\lambda(x - \mu)^2/(2\mu^2 x)\right)$$

  dla wszystkich $x > 0;\ \mu \in \mathbb{R}, \lambda > 0$.

- *Rozkład statystyki ekstremalnej* $\mathrm{EV}(\gamma)$ z $F(x) = \exp(-(1+\gamma x)_+^{-1/\gamma})$ dla wszystkich $\gamma \in \mathbb{R}$ gdzie wielkość $-(1+\gamma x)_+^{-1/\gamma}$ jest zdefiniowana jako e^{-x} kiedy $\gamma = 0$. W tym ostatnim przypadku rozkład jest znany jako *Rozkład Gumbela*. Dla $\gamma > 0$ otrzymujemy *rozkład Frécheta* ; rokład ten jest skoncentrowany na półprostej $(-1/\gamma, +\infty)$. W końcu, dla $\gamma < 0$ otrzymujemy *rozkład ekstremalny Weibulla*, skoncentrowany na półprostej $(-\infty, -1/\gamma)$.

2.3 Rozkłady z ciężkimi ogonami

Przykładami absolutnie ciągłych rozkładów, które mają ciężki ogon są podane niżej.

- *Rozkład logarytmiczno normalny* $\mathrm{LN}(a,b)$ z $f(x) = (xb\sqrt{2\pi})^{-1}\exp\left(-(\log x - a)^2/(2b^2)\right)$ dla $x > 0$; $a \in \mathbb{R}, b > 0$; jeśli X jest o rozkłdzie $\mathrm{N}(a,b)$ to e^X ma rozkład $\mathrm{LN}(a,b)$.

- *Rozkład Weibulla* $\mathrm{W}(r,c)$ z $f(x) = rcx^{r-1}\exp(-cx^r)$ dla $x > 0$ gdzie parametr kształtu $r > 0$ i parametr skali $c > 0$; $\overline{F}(x) = \exp(-cx^r)$; jeśli $r \geq 1$, to $\mathrm{W}(r,c)$ ma lekki ogon, w przeciwnym razie $\mathrm{W}(r,c)$ ma ciężki.

- *Rozkład Pareto* $\mathrm{Par}(\alpha)$ z $f(x) = \alpha(1/(1+x))^{\alpha+1}$ for $x > 0$; gdzie eksponent $\alpha > 0$; $\overline{F}(x) = (1/(1+x))^{\alpha}$.

- *Rozkład loggamma* $\mathrm{L}\Gamma(a,\lambda)$ z $f(x) = \lambda^a/\Gamma(a)(\log x)^{a-1}x^{-\lambda-1}$ dla $x > 1$; $\lambda, a > 0$; jeśli X jest $\Gamma(a,\lambda)$, to e^X ma rozkład $\mathrm{L}\Gamma(a,\lambda)$.

2.4 Rozkłady wielowymiarowe

Przykładami absolutnie ciągłych wielowymiarowych rozkładów są podane niżej.

- *Rozkład wielowymiarowy normalny* $\mathrm{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$ ma wektor losowy $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ z wartością oczekiwaną $\mathbb{E}\,\boldsymbol{X} = \boldsymbol{m}$ oraz macierza kowariancji $\boldsymbol{\Sigma}$, jeśli dowolna kombinacja liniowa $\sum_{j=1}^n a_j X_j$ ma jednowymiarowy rozkład normalny. Niech $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T$. Jeśli $\boldsymbol{Z}$ ma rozkład $\mathrm{N}((0,\ldots,0), \boldsymbol{I})$, gdzie $\boldsymbol{I}$ jest macierzą jednostkową, to $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{Z}$ ma rozkład normalny $\mathrm{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$

3 Tablice rozkładów

Następujące tablice podają najważniejsze rozkłady. Dla każdego rozkładu podane są: Średnia, wariancja, współczynnik zmienności zdefiniowany przez $\sqrt{\operatorname{Var} X}/\mathbb{E}\, X$ (w.zm.). Podane są również: funkcja tworząca $\mathbb{E}\, s^N$ (f.t.) w przypadku dyskretnego rozkładu, oraz funkcja tworząca momenty $\mathbb{E}\, \exp(sX)$ (f.t.m.) w przypadku absolutnie ciągłych rozkładów, jeśli tylko itnieją jawne wzory. W przeciwnym razie w tabeli zaznaczamy $*$. Więcej o tych rozkładach można znaleźć w w podrozdziale VII.

rozkład	parametry	funkcja prawdopodobieństwa $\{p_k\}$
$\mathrm{B}(n,p)$	$n=1,2,\ldots$ $0\le p\le 1, q=1-p$	$\binom{n}{k}p^k q^{n-k}$ $k=0,1,\ldots,n$
$\mathrm{Ber}(p)$	$0\le p\le 1, q=1-p$	$p^k q^{1-k}; k=0,1$
$\mathrm{Poi}(\lambda)$	$\lambda>0$	$\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}; k=0,1,\ldots$
$\mathrm{NB}(r,p)$	$r>0, 0<p<1$ $q=1-p$	$\frac{\Gamma(r+k)}{\Gamma(r)k!}q^r p^k$ $k=0,1,\ldots$
$\mathrm{Geo}(p)$	$0<p\le 1, q=1-p$	$qp^k; k=0,1,\ldots$
$\mathrm{Log}(p)$	$0<p<1$ $r=\frac{-1}{\log(1-p)}$	$r\frac{p^k}{k}; k=1,2,\ldots$
$\mathrm{UD}(n)$	$n=1,2,\ldots$	$\frac{1}{n}; k=1,\ldots,n$

Tablica 3.1: Tablica rozkładów dyskretnych

średnia	wariancja	w.zm.	f.t.
np	npq	$\left(\frac{q}{np}\right)^{\frac{1}{2}}$	$(ps+q)^n$
p	pq	$\left(\frac{q}{p}\right)^{\frac{1}{2}}$	$ps+q$
λ	λ	$\left(\frac{1}{\lambda}\right)^{\frac{1}{2}}$	$\exp[\lambda(s-1)]$
$\frac{rp}{q}$	$\frac{rp}{q^2}$	$(rp)^{-\frac{1}{2}}$	$q^r(1-ps)^{-r}$
$\frac{p}{q}$	$\frac{p}{q^2}$	$p^{-\frac{1}{2}}$	$q(1-ps)^{-1}$
$\frac{rp}{1-p}$	$\frac{rp(1-rp)}{(1-p)^2}$	$\left(\frac{1-rp}{rp}\right)^{\frac{1}{2}}$	$\frac{\log(1-ps)}{\log(1-p)}$
$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\left(\frac{n-1}{3(n+1)}\right)^{\frac{1}{2}}$	$\frac{1}{n}\sum_{k=1}^{n} s^k$

Tablica 3.1: Tablica rozkładów dyskretnych (cd.)

rozkład	parametry	f. gęstości $f(x)$
$\mathrm{U}(a,b)$	$-\infty < a < b < \infty$	$\frac{1}{b-a}; x \in (a,b)$
$\Gamma(a,\lambda)$	$a, \lambda > 0$	$\frac{\lambda^a x^{a-1} \mathrm{e}^{-\lambda x}}{\Gamma(a)}; x \geq 0$
$\mathrm{Erl}(n,\lambda)$	$n = 1, 2 \ldots, \lambda > 0$	$\frac{\lambda^n x^{n-1} \mathrm{e}^{-\lambda x}}{(n-1)!}; x \geq 0$
$\mathrm{Exp}(\lambda)$	$\lambda > 0$	$\lambda \mathrm{e}^{-\lambda x}; x \geq 0$
$\mathrm{N}(\mu,\sigma^2)$	$-\infty < \mu < \infty,$ $\sigma > 0$	$\frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma}; x \in \mathbb{R}$
$\chi^2(n)$	$n = 1, 2, \ldots$	$\left(2^{n/2}\Gamma(n/2)\right)^{-1} x^{\frac{n}{2}-1} \mathrm{e}^{-\frac{x}{2}}; x \geq 0$
$\mathrm{LN}(a,b)$	$-\infty < a < \infty, b > 0$ $\omega = \exp(b^2)$	$\frac{\exp\left(-\frac{(\log x - a)^2)}{2b^2}\right)}{xb\sqrt{2\pi}}; x \geq 0$

Tablica 3.2: Tablica rozkładów absolutnie ciągłych

średnia	wariancja	w.zm.	f.t.m.
$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{b-a}{\sqrt{3}(b+a)}$	$\frac{\mathrm{e}^{bs}-\mathrm{e}^{as}}{s(b-a)}$
$\frac{a}{\lambda}$	$\frac{a}{\lambda^2}$	$a^{-\frac{1}{2}}$	$\left(\frac{\lambda}{\lambda-s}\right)^a; s<\lambda$
$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$n^{-\frac{1}{2}}$	$\left(\frac{\lambda}{\lambda-s}\right)^n; s<\lambda$
$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	1	$\frac{\lambda}{\lambda-s}; s<\lambda$
μ	σ^2	$\frac{\sigma}{\mu}$	$\mathrm{e}^{\mu s+\frac{1}{2}\sigma^2 s^2}$
n	$2n$	$\sqrt{\frac{2}{n}}$	$(1-2s)^{-n/2}; s<1/2$
$\exp\left(a+\frac{b^2}{2}\right)$	$\mathrm{e}^{2a}\omega(\omega-1)$	$(\omega-1)^{\frac{1}{2}}$	$*$

Tablica 3.2: Tablica rozkładów absolutnie ciągłych (c.d.)

rozkład	parametry	f. gęstości $f(x)$
$\mathrm{Par}(\alpha)$	$\alpha > 0$	$\alpha \left(\frac{1}{1+x}\right)^{\alpha+1}; x > 0$
$\mathrm{W}(r,c)$	$r, c > 0$ $\omega_n = \Gamma\left(\frac{r+n}{r}\right)$	$rcx^{r-1}\mathrm{e}^{-cx^r}; x \geq 0$
$\mathrm{IG}(\mu,\lambda)$	$\mu > 0, \lambda > 0$	$\left[\frac{\lambda}{2\pi x^3}\right]^{\frac{1}{2}} \exp\left(\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right); x > 0$
$\mathrm{PME}(\alpha)$	$\alpha > 1$	$\int_{\frac{\alpha-1}{\alpha}}^{\infty} \alpha \left(\frac{\alpha-1}{\alpha}\right)^{\alpha} y^{-(\alpha+2)}\mathrm{e}^{-x/y}dy$ $x > 0$

Tablica 3.2: Tablica rozkładów absolutnie ciągłych (c.d.)

średnia	wariancja	w.z.	f.t.m.
$\frac{1}{\alpha-1}$ $\alpha>1$	$\frac{\alpha}{(\alpha-1)^2(\alpha-2)}$ $\alpha>2$	$[\frac{\alpha}{\alpha-2}]^{-\frac{1}{2}}$ $\alpha>2$	$*$
$\omega_1 c^{-1/r}$	$(\omega_1-\omega_2^2)c^{-2/r}$	$\left[\frac{\omega_2}{\omega_1^2}-1\right]^{\frac{1}{2}}$	$*$
μ	$\frac{\mu^3}{\lambda}$	$\left(\frac{\mu}{\lambda}\right)^{\frac{1}{2}}$	$\frac{\exp\left(\frac{\lambda}{\mu}\right)}{\exp\left(\sqrt{\frac{\lambda^2}{\mu^2}-2\lambda s}\right)}$
1	$1+\frac{2}{\alpha(\alpha-2)}$	$\sqrt{1+\frac{2}{\alpha(\alpha-2)}}$	$\frac{\int_0^{\frac{\alpha}{\alpha-1}}\frac{x^\alpha}{x-s}dx}{\frac{\alpha^{\alpha-1}}{(\alpha-1)^\alpha}}$ $s<0$

Tablica 3.2: Tablica rozkładów absolutnie ciągłych (cd.)

4 m-pliki

Podamy tutaj przykładowe m-pliki.

4.1 Generowanie zmiennych losowych o zadanym rozkładach

Dyskretna zmienna losowa

```
function [sample] = simdiscr(npoints, pdf, val)
% SIMDISCR random numbers from a discrete random
% variable X which attains values x1,...,xn with probabilities
% p1,...,pn
% [sample] = simdiscr(n_points, pdf [, val])
%
% Inputs: npoints - sample size
%         pdf - vector of probabilities (p1, ...,  pn). They should
%            sum up to 1.
%         val - optional, vector of values (x1, ... xn). Default is
%            assumed (1, ..., n).
%
% Outputs: sample - vector of random numbers
%
% See also SIMBINOM, SIMEXP, SIMGEOM, SIMPARETO

% Authors: R.Gaigalas, I.Kaj
% v1.2 04-Oct-02

  % check arguments
  if (sum(pdf) ~= 1)
    error('Probabilities does not sum up to 1');
  end

  n = length(pdf);

  % if the last argument omitted
  % assign (1, ..., n) to the value vector
  if (nargin==2)
    val = [1:n];
```

```
  end

  cumprob = [0 cumsum(pdf)]; % the jump points of cdf

  runi = rand(1, npoints); % random uniform sample

  sample = zeros(1, npoints);
  for j=1:n
    % if the value of U(0,1) falls into the interval (p_j, p_j+1)
    % assign the value xj to X
    ind = find((runi>cumprob(j)) & (runi<=cumprob(j+1)));
    sample(ind) = val(j);
  end
```

Rozkład Pareto

```
function [sample] = simparetonrm(M, N, alpha, gamma)

% SIMPARETONRM Generate a matrix of random numbers from the
% normalized Pareto distribution with
%
% pdf      f(x) = alpha*gamma/(1+gamma*x)^(1+alpha),  x>0,
% cdf      F(x) = 1-(1+gamma*x)^(-alpha).
%
% The additional parameter gamma allows to control the expected
% value. For alpha>1 the expected value exists and is equal to
% 1/(gamma*(alpha-1)).
%
% [sample] = simparetonrm(M, N, alpha, gamma)
%
% Inputs:
%        M - number of rows in the output matrix
%        N - number of columns in the output matrix
%        alpha - tail parameter of the distribution
%        gamma - parameter of the distribution
%
% Outputs:
%        sample - MxN matrix of random numbers
```

```
%
% See also SIMBINOM, SIMDISCR, SIMGEOM, SIMEXP, SIMPARETO

% Authors: R.Gaigalas, I.Kaj
% v2.0 17-Oct-05

  sample = ((1-rand(M, N)).^(-1/alpha)-1)./gamma;
```

5 Wybrane instrukcje MATLAB'a

5.1 hist

5.2 rand

```
Uniformly distributed pseudorandom numbers

Syntax

Y = rand
Y = rand(n)
Y = rand(m,n)
Y = rand([m n])
Y = rand(m,n,p,...)
Y = rand([m n p...])
Y = rand(size(A))
rand(method,s)
s = rand(method)



Description


Y = rand returns a pseudorandom, scalar value drawn from a uniform
    distribution on the unit interval.

Y = rand(n) returns an n-by-n matrix of values derived as described
    above.

Y = rand(m,n) or Y = rand([m n]) returns an m-by-n matrix of the same.

Y = rand(m,n,p,...) or Y = rand([m n p...]) generates an m-by-n-by-p-by-...
    array of the same.

Note    The size inputs m, n, p, ... should be nonnegative integers.
        Negative integers are treated as 0.

Y = rand(size(A)) returns an array that is the same size as A.
```

```
rand(method,s) causes rand to use the generator determined by method,
    and initializes the state of that generator using the value of s.

The value of s is dependent upon which method is selected. If method is
set to 'state' or 'twister', then s must be either
a scalar integer value from 0 to 2^32-1 or the output of rand(method).
If method is set to 'seed', then s must be either
a scalar integer value from 0 to 2^31-2 or the output of rand(method).
To set the generator to its default initial state,
set s equal to zero.

The rand and randn generators each maintain their own internal state
information. Initializing the state of one has no effect
on the other.

Input argument method can be any of the strings shown in the table below:

methodDescription'state'Use a modified version of Marsaglia's subtract
with borrow algorithm (the default in MATLAB versions 5 and later).
This method can generate all the double-precision values in the closed
interval [2^(-53), 1-2^(-53)]. It theoretically can
generate over 2^1492 values before repeating itself.'seed'Use a multiplicati
congruential algorithm (the default in MATLAB version 4).
This method generates double-precision values in the closed interval
[1/(2^31-1), 1-1/(2^31-1)], with a period of 2^31-2.
'twister'Use the Mersenne Twister algorithm by Nishimura and Matsumoto.
This method generates double-precision values in
the closed interval [2^(-53), 1-2^(-53)], with a period of (2^19937-1)/2.

For a full description of the Mersenne twister algorithm, see

http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html

s = rand(method) returns in s the current internal state of the generator
selected by method. It does not change the generator being used.

Remarks
```

```
The sequence of numbers produced by rand is determined by
the internal state of the generator. Setting the generator
to the same fixed state enables you to repeat computations.
Setting the generator to different states leads to unique
computations. It does not, however, improve statistical
properties.

Because MATLAB resets the rand state at startup, rand
generates the same sequence of numbers in each session
unless you change the value of the state input.

Examples


Example 1


Make a random choice between two equally probable alternatives:
 if rand < .5
    'heads'
 else
    'tails'
 end



Example 2


Generate a 3-by-4 pseudorandom matrix:
R = rand(3,4)
R =
    0.2190    0.6793    0.5194    0.0535
    0.0470    0.9347    0.8310    0.5297
    0.6789    0.3835    0.0346    0.6711
```

Example 3

```
Set rand to its default initial state:
rand('state', 0);


Initialize rand to a different state each time:
rand('state', sum(100*clock));


Save the current state, generate 100 values, reset the state,
and repeat the sequence:
s = rand('state');
u1 = rand(100);
rand('state',s);
u2 = rand(100);      % Contains exactly the same values as u1.
```

Example 4

```
Generate uniform integers on the set 1:n:
n = 75;
f = ceil(n.*rand(100,1));

f(1:10)
ans =
    72
    34
    25
     1
    13
```

```
    63
    49
    25
     2
    27
```

Example 5

Generate a uniform distribution of random numbers on a specified interval [a,b]. To do this, multiply the output of rand by (b-a), then add a.
For example, to generate a 5-by-5 array of uniformly distributed random numbers on the interval [10,50],

```
a = 10; b = 50;
x = a + (b-a) * rand(5)
x =

   18.1106   10.6110   26.7460   43.5247   30.1125
   17.9489   39.8714   43.8489   10.7856   38.3789
   34.1517   27.8039   31.0061   37.2511   27.1557
   20.8875   47.2726   18.1059   25.1792   22.1847
   17.9526   28.6398   36.8855   43.2718   17.5861
```

Example 6

Use the Mersenne twister generator, with the default initial state used by Nishimura and Matsumoto:

```
rand('twister',5489);
```

References

[1] Moler, C.B., "Numerical Computing with MATLAB," SIAM, (2004), 336 pp.
Available online at http://www.mathworks.com/moler.

[2] G. Marsaglia and A. Zaman "A New Class of Random Number Generators,
" Annals of Applied Probability, (1991), 3:462-480.

[3] Matsumoto, M. and Nishimura, T. "Mersenne Twister: A 623-Dimensionally
Equidistributed Uniform Pseudorandom Number Generator,"
ACM Transactions on Modeling and Computer Simulation, (1998), 8(1):3-30

[4] Park, S.K. and Miller, K.W. "Random Number Generators: Good Ones Are
Hard to Find," Communications of the ACM, (1988), 31(10):1192-1201.

5.3 randn

Normally distributed random numbers

Syntax

```
Y = randn
Y = randn(n)
Y = randn(m,n)
Y = randn([m n])
Y = randn(m,n,p,...)
Y = randn([m n p...])
Y = randn(size(A))
randn(method,s)
s = randn(method)
```

Description

Y = randn returns a pseudorandom, scalar value drawn from

```
    a normal distribution with mean 0 and standard deviation 1.

Y = randn(n) returns an n-by-n matrix of values derived as
    described above.

Y = randn(m,n) or Y = randn([m n]) returns an m-by-n
    matrix of the same.

Y = randn(m,n,p,...) or Y = randn([m n p...]) generates
    an m-by-n-by-p-by-... array of the same.

Note    The size inputs m, n, p, ... should be nonnegative
integers. Negative integers are treated as 0.

Y = randn(size(A)) returns an array that is the same size as A.

randn(method,s) causes randn to use the generator determined
by method, and initializes the state of that generator using
the value of s.

The value of s is dependent upon which method is selected.
If method is set to 'state', then s must be either a scalar
integer value from 0 to 2^32-1 or the output of rand(method).
If method is set to 'seed', then s must be either a scalar
integer value from 0 to 2^31-2 or the output of rand(method).
To set the generator to its default initial state, set s
equal to zero.

The randn and rand generators each maintain their own internal
state information. Initializing the state of one has no effect
on the other.

Input argument method can be either of the strings shown in
the table below:

methodDescription'state'Use Marsaglia's ziggurat algorithm
(the default in MATLAB versions 5 and later). The period
is approximately 2^64.'seed'Use the polar algorithm (the default
```

```
in MATLAB version 4). The period is approximately (2^31-1)*(pi/8).

s = randn(method) returns in s the current internal state
of the generator selected by method. It does not change
the generator being used.

Examples


Example 1


R = randn(3,4) might produce
 R =
    1.1650    0.3516    0.0591    0.8717
    0.6268   -0.6965    1.7971   -1.4462
    0.0751    1.6961    0.2641   -0.7012



For a histogram of the randn distribution, see hist.

Example 2


Set randn to its default initial state:
randn('state', 0);



Initialize randn to a different state each time:
randn('state', sum(100*clock));



Save the current state, generate 100 values, reset the state,
and repeat the sequence:
s = randn('state');
```

```
u1 = randn(100);
randn('state',s);
u2 = randn(100);      % Contains exactly the same values as u1.
```

Example 3

```
Generate a random distribution with a specific mean and variance .
To do this, multiply the output of randn by the standard
deviation , and then add the desired mean. For example,
to generate a 5-by-5 array of random numbers with a mean
of .6 that are distributed with a variance of 0.1,
x = .6 + sqrt(0.1) * randn(5)
x =

    0.8713    0.4735    0.8114    0.0927    0.7672
    0.9966    0.8182    0.9766    0.6814    0.6694
    0.0960    0.8579    0.2197    0.2659    0.3085
    0.1443    0.8251    0.5937    1.0475   -0.0864
    0.7806    1.0080    0.5504    0.3454    0.5813
```

References

[1] Moler, C.B., "Numerical Computing with MATLAB," SIAM, (2004), 336 pp. Available online at http://www.mathworks.com/moler.

[2] Marsaglia, G. and Tsang, W.W., "The Ziggurat Method for Generating Random Variables," Journal of Statistical Software, (2000), 5(8). Available online at http://www.jstatsoft.org/v05/i08/.

[3] Marsaglia, G. and Tsang, W.W., "A Fast, Easily Implemented Method for Sampling from Decreasing or Symmetric Unimodal Density Functions," SIAM Journal of Scientific and Statistical

Computing, (1984), 5(2):349-359.

[4] Knuth, D.E., "Seminumerical Algorithms," Volume 2 of The Art of Computer Programming, 3rd edition Addison-Wesley (1998).

5.4 randperm

5.5 stairs

Bibliografia

[1] Asmussen S. and Glynn (2007) *Stochastic Simulations; Algorithms and Analysis*. Springer, New York.

[2] Feller W. (1987) *Wstęp do racunku prawdopodobieństwa*. PWN, Warszawa.

[3] Fishman G. (1966) *Monte Carlo*. Springer, New York.

[4] Fishman G. (2001) *Discrete-Event Simulation*. Springer, New York.

[5] Glasserman P. (2004) *Monte Carlo Methods in Financial Engineering*. Springer, New York.

[6] Madras N. (2002) *Lectures on Monte Carlo Methods*. AMS, Providence.

[7] Resing J. (2008) Simulation. Manuscript, Eindhoven University of Technology.

[8] Ripley B. (1987) *Stochastic Simulation*. Wiley, Chichester.

[9] Ross S. (1991) *A Course in Simulation*. Macmillan, New York.

[10] Srikant R. and Whitt W. (2008) Simulation run length planning for stochastic loss models. *Proceedings of the 1995 Winter Simulation Conference* **22**: 415–429.

[11] Whitt W. (1989) Planing queueing simulation. *Management Science* **35**: 1341–1366.

[12] Whitt W. (1991) The efficiency of one long run versus independent replications in steady-state simulations. *Management Science* **37**: 645–666.

[13] Whitt W. (1992) Asymptotic formulas for markov processes with applications to simulation. *Operations Research* **40**: 279–291.

[14] W.K. G. (2008) Warm-up periods in simulation can be detrimental. *PEIS* **22**: 415–429.

[15] Zieliński R. (1970) *Metody Monte Carlo*. WNT, Warszawa.

Skorowidz nazw