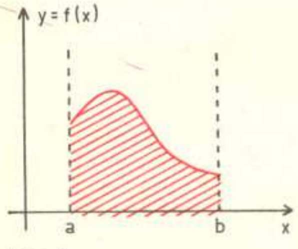
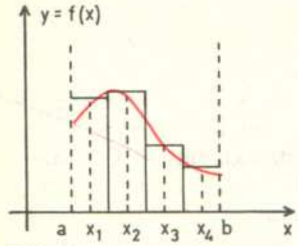
Metody Monte Carlo (V)

Dr Ryszard ZIELIŃSKI

OBLICZANIE CAŁEK (c.d.)



Rys. 1



Rys. 2

Podobnie jak w poprzednim odcinku, będziemy zajmowali się metodami Monte Carlo obliczania całki

(1) $$I = \int_a^b f(x)dx.$$

Prezentację bardziej zaawansowanych metod musimy jednak poprzedzić pewnymi uwagami i przypomnieć pewne fakty ze szkolnego kursu rachunku prawdopodobieństwa.

Dla uproszczenia dalszych rozważań będziemy — podobnie jak przy prezentacji poprzednio omawianej metody „orzeł-reszka" — zakładali, że f jest funkcją nieujemną i traktowali całkę (1) jako pole powierzchni ograniczonej osią odciętych, prostymi $x = a$, $x = b$ i wykresem funkcji f (rys. 1). Podzielmy odcinek (a,b) na m równych części i niech x będzie odciętą punktu będącego środkiem j-tego odcinka (rys. 2). Mamy wtedy, jak łatwo sprawdzić

(2) $$x_j = a + \left(j - \frac{1}{2}\right) \cdot \frac{b-a}{m}, \quad j = 1,2, \dots, m.$$

Zauważmy, że jeżeli m jest duże, to pole pod krzywą $y = f(x)$ nie różni się dużo od pola figury utworzonej przez wszystkie „słupki" przedstawione na rys. 2. Całka I będzie wtedy w przybliżeniu równa:

(3) $$I \approx (b-a) \sum_{j=1}^{m} y_j \cdot \frac{1}{m}.$$

Żeby to przybliżenie było „dobre", funkcja f musi być dostatecznie regularna (wystarczy np., żeby była ciągła), a liczba m odpowiednio duża, ale nie będziemy dyskutowali tego szczegółowo.

A teraz sięgnijmy do szkolnego kursu rachunku prawdopodobieństwa. Niech $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$ będzie zbiorem zdarzeń elementarnych i niech $p_j = P(\{\omega_j\})$. Na zbiorze Ω określimy zmienną losową Y i niech y_j oznacza wartość tej zmiennej losowej w punkcie ω_j, tzn. $y_j = Y(\omega_j)$. Wartość oczekiwana EY tej zmiennej losowej wyraża się znanym wzorem

(4) $$EY = \sum_{j=1}^{m} y_j p_j,$$

a jeżeli wszystkie prawdopodobieństwa p_j są równe $\left(\text{a więc jeżeli } p_j = \frac{1}{m}\right)$, otrzymujemy

(5) $$EY = \sum_{j=1}^{m} y_j \cdot \frac{1}{m},$$

co jest identyczne z sumą występującą po prawej stronie wzoru (3). Mamy więc następujący przybliżony wzór dla całki I:

(6) $$I \approx (b-a) \cdot EY,$$

gdzie Y jest odpowiednią zmienną losową.

Przeprowadzimy następujący eksperyment. Weźmy m jednakowych kartek papieru, na kartkach tych wypiszmy kolejno wartości $y_1, y_2, \dots, y_m$ (niektóre z liczb y_j mogą oczywiście być jednakowe) wrzućmy wszystkie kartki do kapelusza, wymieszajmy je starannie i wyciągnijmy na chybił-trafił jedną z nich. Będzie nas interesowała liczba napisana na wylosowanej kartce. Jest to oczywiście jedna z wartości zmiennej losowej, która z prawdopodobieństwem $\frac{1}{m}$ przyjmuje każdą z wartości $y_1, y_2, \dots, y_m$. Po wyciągnięciu kartki z kapelusza i zapisaniu sobie liczby z tej kartki, wkładamy ją z powrotem do kapelusza i powtarzamy nasze losowanie, powiedzmy, n razy. Niech $Y^{(k)}$ oznacza wynik k-tego losowania. Oczywiście dla każdego k mamy

$$P\{Y^{(k)} = y_j\} = \frac{1}{m}, \ j = 1, 2, \dots, m.$$

Wyniki kolejnych losowań są niezależne (losowanie ze zwracaniem).

Z prawa wielkich liczb (patrz poprzednie odcinki) wiemy, że z prawdopodobieństwem równym jedności

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} Y^{(k)} = EY,$$

a centralne twierdzenie graniczne pozwala nam oszacować różnicę między $\frac{1}{n} \sum_{k=1}^{n} k = Y^{(k)}$

Rozwiązanie zadania M 77.

Niech x będzie wysokością wieży. Mamy wówczas: $\operatorname{tg}\alpha = \frac{x}{a}$, $\operatorname{tg}\beta = \frac{x}{b}$, $\operatorname{tg}(90^\circ - (\alpha+\beta)) = \frac{x}{c}$.

Ale $\frac{x}{c} = \operatorname{tg}(90^\circ - (\alpha+\beta)) = \operatorname{ctg}(\alpha+\beta) = \frac{1}{\operatorname{tg}(\alpha+\beta)} = \frac{1 - \operatorname{tg}\alpha \operatorname{tg}\beta}{\operatorname{tg}\alpha + \operatorname{tg}\beta} = \frac{1 - \frac{x^2}{ab}}{\frac{x}{a} + \frac{x}{b}}$.

Rozwiązując otrzymane równanie otrzymujemy kolejno $\frac{x}{c}\left(\frac{x}{a} + \frac{x}{b}\right) = 1 - \frac{x^2}{ab}$,

$x^2\left(\frac{1}{ac} + \frac{1}{bc} + \frac{1}{ab}\right) = 1$, $x^2 \frac{a+b+c}{abc} = 1$,

$$x = \sqrt{\frac{abc}{a+b+c}}$$

(x jako odległość jest liczbą dodatnią, więc pomijamy rozwiązanie ujemne).

Rozwiązanie zadania M 78.

Zachodzi równość $(a-1)!\cdot a = a!$. Jeżeli przyjmiemy, że $a = y!$, to otrzymamy nieskończenie wiele rozwiązań danego równania $x = y!-1$, $z = y!$, y — dowolne. Znamy jeszcze jedno rozwiązanie tego równania (nie objęte podanymi wzorami), a mianowicie $6!\,7! = 10!$; nie wiemy, czy istnieją inne rozwiązania tego równania.

oraz EY: błąd oszacowania wartości oczekiwanej za pomocą takiej średniej wynosi $2S_n/\sqrt{n}$, gdzie (patrz poprzednie odcinki):

$$S_n = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(Y^{(k)})^2 - \left(\frac{1}{n}\sum_{k=1}^{n}Y^{(k)}\right)^2}, \tag{7}$$

W ten sposób szacowanie wartości oczekiwanej zmiennej losowej Y, a więc szacowanie całki (1), sprowadza się do wykonania odpowiedniego eksperymentu i kilku prostych rachunków. Zastąpienie całki I sumą i korzystanie z przybliżenia (3) było nam potrzebne tylko po to, abyśmy mogli wyjaśnić mechanizm odpowiedniego eksperymentu posługując się pojęciami zmiennej losowej i wartości oczekiwanej zmiennej losowej, znanymi ze szkolnego kursu rachunku prawdopodobieństwa. Możemy uniknąć tego przybliżenia zastępując losowanie kartek z kapelusza losowaniem punktów z przedziału (a,b), np. tak jak to już robiliśmy w poprzednim odcinku. Prowadzi to do następującego algorytmu szacowania całki (1):

1) wybrać na chybił-trafił punkt z przedziału (a,b);
2) obliczyć wartość funkcji f w wylosowanym punkcie;
3) powtórzyć obie czynności n razy i obliczyć średnią arytmetyczną liczb otrzymanych w punkcie 2;
4) obliczyć oszacowanie całki mnożąc wynik z punktu 3 przez liczbę $(b-a)$.

Jeżeli oznaczymy przez x_k, $k = 1,2, \ldots, m$, współrzędną punktu z przedziału (a,b), otrzymanego w k-tym losowaniu, to dla oszacowania $\hat{I}$ całki I otrzymujemy wzór:

$$\hat{I} = \frac{b-a}{n}\sum_{k=1}^{n} f(x_k), \tag{8}$$

a błąd tego oszacowania (por. wzór (7)) jest równy

$$\frac{2(b-a)}{\sqrt{n}}\sqrt{\frac{1}{n}\sum_{k=1}^{n}f^2(x_k) - \left(\frac{1}{n}\sum_{k=1}^{n}f(x_k)\right)^2} \tag{9}$$

Dla ilustracji przytoczymy dwa wyniki liczbowe. Obliczano całkę $\int_0^1 \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx$ (której wartość, równa 0,3413, może być dokładnie obliczona innymi metodami). Obliczenia wykonano raz metodą orzeł-reszka i raz metodą przed chwilą opisaną, w każdym przypadku losując 25 punktów w przedziale całkowania (0,1). W przypadku metody orzeł-reszka otrzymano wynik 0,40 z błędem 0,1960; w przypadku przedstawionej wyżej metody otrzymano wynik 0,3442 z błędem 0,0192.

Obliczamy całkę $I = \int_0^1 f(x)dx$ dla $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

metodą „orzeł-reszka"

Numer losowania j	Wylosowany punkt x_j	Wylosowany punkt y_j	$f(x_j)$	Czy $y_j < f(x_j)$?
1	0,10	0,09	0,397	tak
2	0,73	0,25	0,306	tak
3	0,33	0,76	0,378	nie
4	0,52	0,01	0,348	tak
5	0,35	0,86	0,375	nie
6	0,34	0,67	0,377	nie
7	0,35	0,48	0,375	nie
8	0,76	0,80	0,299	nie
9	0,95	0,90	0,254	nie
10	0,91	0,17	0,264	tak
11	0,39	0,29	0,370	tak
12	0,27	0,49	0,385	nie
13	0,45	0,37	0,361	nie
14	0,54	0,20	0,345	tak
15	0,48	0,05	0,356	tak
16	0,64	0,89	0,325	nie
17	0,47	0,42	0,357	nie
18	0,96	0,24	0,252	tak
19	0,80	0,52	0,290	nie
20	0,40	0,37	0,368	nie
21	0,20	0,63	0,391	nie
22	0,61	0,04	0,331	tak
23	0,02	0,00	0,399	tak
24	0,82	0,29	0,285	nie
25	0,16	0,65	0,394	nie

Liczba sukcesów („tak"): 10

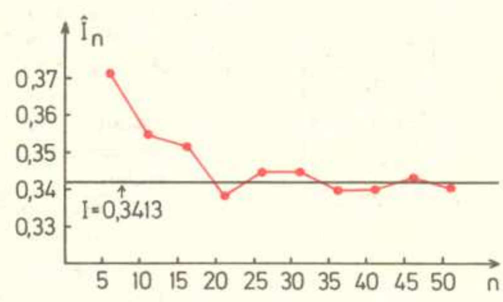
Oszacowanie całki $\hat{I} = 10/25 = 0,4$

metodą podstawową

Numer losowania j	Wylosowany punkt x_j	$f(x_j)$
1	0,10	0,3970
2	0,09	0,3973
3	0,73	0,3056
4	0,25	0,3867
5	0,33	0,3778
6	0,76	0,2989
7	0,52	0,3485
8	0,01	0,3989
9	0,35	0,3752
10	0,86	0,2756
11	0,34	0,3765
12	0,67	0,3187
13	0,35	0,3752
14	0,48	0,3555
15	0,76	0,2989
16	0,80	0,2897
17	0,95	0,2541
18	0,90	0,2661
19	0,91	0,2637
20	0,17	0,3932
21	0,39	0,3697
22	0,29	0,3825
23	0,27	0,3847
24	0,49	0,3538
25	0,45	0,3605
	Suma	8,6043

Oszacowanie całki $\hat{I} = 8,6043/25 = 0,3442$

Oznaczamy przez I_n kolejne oszacowania całki uzyskane metodą podstawową po n losowaniach.
Biorąc za podstawę oszacowań wyniki z tabelki (i dalsze, nie podane w tej tabelce) otrzymamy wykres:



Przypomnijmy jeszcze raz (por. »Delta« 11/1975), co oznacza zdanie „otrzymano wynik 0,3442 z błędem 0,0192".
Zdanie to można sformułować również w następujący sposób: „stwierdzono, że poszukiwana wartość całki jest liczbą z przedziału (0,3250, 0,3634)". Gdybyśmy jeszcze raz powtórzyli nasze rachunki, otrzymalibyśmy z pewnością inne oszacowanie dla całki i inne oszacowanie dla błędu. Zdanie „poszukiwana wartość całki jest liczbą z przedziału (oszacowanie całki − błąd, oszacowanie całki + błąd)" może być zdaniem prawdziwym lub zdaniem fałszywym, ale prawdopodobieństwo tego, że będzie to zdanie prawdziwe, jest wysokie i wynosi około 0,95. Wartość 0,95 tego prawdopodobieństwa jest związana ze współczynnikiem 2 występującym we wzorze (9) na wielkość błędu.
Podana wyżej metoda (nazywa się ją czasami „metodą podstawową") jest zawsze dokładniejsza od metody orzeł-reszka w tym sensie, że ma mniejszy błąd (chociaż może się zdarzyć, że wynik otrzymany metodą orzeł-reszka będzie bliższy prawdy niż wynik otrzymany metodą podstawową). Ilustrują to przytoczone wyżej przykłady liczbowe. Znane są jeszcze dokładniejsze metody Monte Carlo. Najprostsze z nich to metoda losowania warstwowego i metoda średniej ważonej. Pierwsza polega na tym, że przedział całkowania (a w przypadku funkcji wielu zmiennych — obszar całkowania) rozbija się na sumę rozłącznych przedziałów („warstw") i całkę przedstawia się jako sumę odpowiednich całek liczonych na poszczególnych warstwach.
Np. dla liczb $c_1, c_2, \ldots, c_s$ takich, że $a \leqslant c_1 \leqslant c_2 \leqslant \ldots \leqslant c_s \leqslant b$ mamy

$$\int_a^b f(x)dx = \int_a^{c_1} f(x)dx + \int_{c_1}^{c_2} f(x)dx + \ldots + \int_{c_s}^{b} f(x)dx$$

Każdą z całek takiej sumy szacuje się metodą Monte Carlo, a poszukiwaną wartość całki I szacuje się jako sumę tych oszacowań. Okazuje się, że prowadzi to do znacznej redukcji błędu. Na przykład dla podawanej już całki $\frac{1}{\sqrt{2\pi}}\int_0^1 e^{-x^2/2}\,dx$ przy rozbiciu jej na sumę pięciu całek otrzymano — losując, jak poprzednio, ogółem 25 punktów — wynik 0,3406 z błędem 0,0034.
Metoda średniej ważonej polega, mówiąc z grubsza, na tym, że nie każdy punkt z przedziału całkowania ma jednakowe szanse na wylosowanie. Okazuje się, że w niektórych przypadkach można tak zorganizować losowanie punktów w obszarze całkowania, żeby błąd oszacowania całki był dowolnie mały.
Prezentacja wszystkich metod Monte Carlo obliczania całek wymagałaby znacznie bogatszego aparatu teorii prawdopodobieństwa i statystyki matematycznej, ograniczymy się więc do tego, co już wyżej powiedzieliśmy, a zainteresowanego Czytelnika odsyłamy do obszernej książki „*Metody Monte Carlo*" (Wydawnictwa Naukowo-Techniczne, Warszawa 1970).

Zadania

Redaguje mgr Andrzej MĄKOWSKI

M 76. Wyznaczyć wszystkie wielomiany f spełniające dla każdego wielomianu g i każdej liczby rzeczywistej x równość

$$f(g(x)) = g(f(x)).$$

Rozwiązanie na str. 17

M 77. Z punktów odległych o a, b, c od wieży widać ją pod kątami α, β, $90° - (\alpha+\beta)$. Wyznaczyć wysokość wieży nie korzystając z tablic trygonometrycznych.

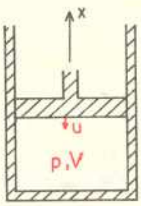
Rozwiązanie na str. 12

M 78. Udowodnić, że równanie $x!y! = z!$ ma nieskończenie wiele rozwiązań w liczbach naturalnych x, y, z większych od 1.

Rozwiązanie na str. 13

Redaguje dr Andrzej ZIEMIŃSKI

F 26. Gaz doskonały znajduje się w naczyniu w kształcie walca, którego jedną z podstaw stanowi ruchomy tłok.



Ścianki naczynia i tłok nie przewodzą ciepła.
Tłok przesuwamy równomiernie z prędkością u, znacznie mniejszą od średniej prędkości cząsteczek gazu.
Znajdźcie, jak zmienia się ciśnienie P gazu w naczyniu, w zależności od objętości V. Rozważcie przypadek gazu jedno- i dwu-atomowego (odpowiednio o trzech i pięciu stopniach swobody ruchu).
Przed przystąpieniem do rozwiązywania powyższego zadania przypomnijmy sobie wyprowadzanie równania stanu gazu z kinetyczno-molekularnej teorii materii.
(Fizyka dla klasy I str. 165).

Rozwiązanie na str. 17