Igła Buffona: dowód z księgi

Metody probabilistyczne informatyki
Instytut Informatyki Analitycznej, Uniwersytet Jagielloński
Jędrzej Hodor, Piotr Micek

12 grudnia 2023

Rozwiązujemy problem krótkiej igły Buffona. Zatem igłę długości ℓ, rzucamy na podłogę z poziomymi liniami w odstępach d i szukamy prawdopodobieństwa tego, że igła przecina którąś linię. Rozważamy problem dla $\ell < d$. Polecamy najpierw przeczytać rozwiązanie problemu[1], w którym definiujemy model przestrzeni probabilistycznej oparty na dwu niezależnych zmiennych: X o rozkładzie jednostajnym na $[0, \frac{d}{2}]$ i A o rozkładzie jednostajnym na $[0, \frac{\pi}{2}]$. Przy takiej formalizacji doświadczenia szukane prawdopodobieństwo to $\frac{2d}{\pi\ell}$. Teraz przestudiujemy zupełnie inny dowód—ciekawszy, magiczny, z Księgi [1].

1 Szkic

Najpierw przejdziemy przez dowód szybko przymykając oko na niektóre szczegóły. Po pierwsze, skoro $\ell < d$, to możemy przeciąć igłą tylko jedną linię. Dlatego,

$$\mathbb{P}(\text{igła przecina którąś linię}) = \mathbb{E}(\text{liczba linii przeciętych przez igłę}).$$

Po drugie, zauważmy, że liczba przecięć igły z liniami jest równa sumie liczby przecięć połówek igły z liniami. Co więcej, jeśli popatrzymy na połówkę igły i zapomnimy o drugiej połówce to zachowuje się ona tak jak dwa razy krótsza igła. Więc z liniowości wartości oczekiwanej

$$\mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } \ell) = 2 \cdot \mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } \ell/2).$$

Stąd nietrudno wyprowadzić, że

$$\mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } \ell) = \ell \cdot \mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } 1).$$

Co więcej, nie ma znaczenia, czy rzucamy igłą, czy też łamaną, która ma taką samą długość jak igła (można ją podzielić na kawałki, a potem posklejać do igły, w obu krokach korzystając z liniowości wartości oczekiwanej). Czyli,

$$\mathbb{E}(\text{liczba linii przeciętych przez łamaną o dł. } \ell) = \ell \cdot \mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } 1).$$

Teraz rozważmy okrąg o średnicy d. Po pierwsze, rzucając taki okrąg, on zawsze trafi w dokładnie dwie linie, ponadto obwód okręgu to $d\pi$. Okrąg nie jest łamaną, ale może być aproksymowany przez łamane dowolnie dokładnie. Zatem

$$2 = \mathbb{E}(\text{liczba linii przeciętych przez okrąg o dł. } d\pi) = d\pi \cdot \mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } 1).$$

Zatem finalnie,

$$\mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } \ell) = \ell \cdot \mathbb{E}(\text{liczba linii przeciętych przez igłę dł. } 1) = \ell \cdot \frac{2}{d\pi} = \frac{2\ell}{d\pi}.$$

[1] https://probabil.tcs.uj.edu.pl/docs/igla.pdf

2 Szczegóły

Rozwiązanie przedstawione w ten sposób ma dużo uroku. Jednak tak naprawdę wiele rzeczy jest zamiecionych pod dywan. Dlatego, w drugiej części tej notki spróbujemy zajrzeć pod ten dywan i zobaczyć co się tam kryje.

Najpierw powinniśmy zrozumieć w jakim modelu pracujemy. Rzucanie igły na podłogę jest mało matematyczne. Mamy tutaj dwie opcje. Pierwsza to opisanie modelu analitycznie tak jak w rozwiązaniu ze slajdów (zmienne losowe X i A). Jednak wtedy szybko znajdziemy się w sytuacji, gdzie prościej będzie podążyć za rozwiązaniem obliczeniowym. Dlatego spróbujemy inaczej i wypiszemy kilka abstrakcyjnych warunków, które chcemy, aby nasz model spełniał. Wszystkie warunki będą intuicyjnie jasne. To, czy "jednostajne rzucanie igły na podłogę" spełnia te warunki to pytanie bardziej natury filozoficznej. Z drugiej strony można udowodnić, że model analityczny (ze zmiennymi X i A) spełnia te warunki.

Ustalamy zatem model eksperymentu, gdzie zadana jest pewna przestrzeń probabilistyczna (nie specyfikujemy jaka explicite) i następujące zmienne losowe przyjmujące wartości nieujemne całkowite:

- dla każdej krzywej L na płaszczyźnie zmienna Y_L i
- dla każdej krzywej L na płaszczyźnie i dla każdego segmentu L' krzywej L zmienna $Z_L^{L'}$.

Następnie, wypiszemy warunki, które wyżej wymienione zmienne powinny spełniać. Zmienne można interpretować w następujący sposób: Y_L — liczba przecięć L z krawędziami desek po rzucie na podłogę obiektu o kształcie L; $Z_L^{L'}$ — liczba przecięć L' jako segmentu L z krawędziami desek po rzucie na podłogę obiektu o kształcie L. Wymagamy od tych zmiennych aby:

1. (sumowanie) $Y_L = \sum_{i=1}^k Z_L^{L_i}$ o ile L jest łamaną składającą się z segmentów $L_1, \dots, L_k$, gdzie każde dwa segmenty mają co najwyżej jeden wspólny punkt;
2. (jednostajność) rozkład $Z_L^{L'}$ jest taki sam jak rozkład $Y_{L'}$ (czyli rzucenie segmentu w ramach całości tak samo się zachowuje jak rzucenie jedynie tego segmentu);
3. (monotoniczność) jeżeli L_1, L_2 będą brzegami wypukłych figur F_1, F_2 na płaszczyźnie opowiednio takimi, że $F_1 \subseteq F_2$, to mamy $\mathbb{E}(Y_{L_1}) \leqslant \mathbb{E}(Y_{L_2})$;
4. (wypełnianie) jeżeli C jest okręgiem o średnicy d to $Y_C = 2$ z prawdopodobieństwem 1.

Niech $\ell > 0$. Oznaczmy przez Y_ℓ zmienną Y_L, gdzie L jest odcinkiem o długości ℓ. Niech $\ell_1, \ell_2 > 0$ będą takie, że $\ell = \ell_1 + \ell_2$. Niech L_1, L_2 będą podsegmentami L sumującymi się do L o długościach ℓ_1, ℓ_2 odpowiednio. Z 1, mamy $Y_\ell = Y_L = Z_L^{L_1} + Z_L^{L_2}$. Obkładając równość wartością oczekiwaną i korzystając z 2 i liniowości wartości oczekiwanej mamy $\mathbb{E}(Y_\ell) = \mathbb{E}(Z_L^{L_1}) + \mathbb{E}(Z_L^{L_2}) = \mathbb{E}(Y_{\ell_1}) + \mathbb{E}(Y_{\ell_2})$. Niech $f(\ell) := \mathbb{E}(Y_\ell)$. Wtedy powyższe można napisać jako: $f(\ell) = f(\ell_1) + f(\ell_2)$. Dodatkowo zauważmy, że funkcja f jest monotoniczna (wynika to z równości wyżej i faktu, że z definicji $f(\ell) \geqslant 0$). Daje to klasyczne równanie funkcyjne, którego rozwiązaniem jest funkcja spełniająca $f(\ell) = \ell f(1)$ dla każdego ℓ.[2]

W szczególności, jeżeli W jest wielokątem o obwodzie ℓ (suma skończenie wielu łamanych) to $\mathbb{E}(Y_W) = \ell f(1)$. Niech C będzie okręgiem o średnicy d. Pamiętajmy, że wtedy z 4, $\mathbb{E}(Y_C) = 2$. Okrąg można aproksymować przez n-kąt foremny W_n wpisany w C i n-kąt foremny W'_n opisany na C. Jeśli oznaczymy obwód przez $|\cdot|$, to mamy $|W_n| \leqslant |C| \leqslant |W'_n|$ i dodatkowo $\lim_{n\to\infty} |W_n| = \lim_{n\to\infty} |W'_n| = |C|$. Z 3, dla każdego n mamy

$$\mathbb{E}(Y_{W_n}) \leqslant \mathbb{E}(Y_C) \leqslant \mathbb{E}(Y_{W'_n}).$$

Korzystając z tego co wiemy o przypadku łamanych i przypadku okręgu,

$$|W_n| f(1) \leqslant 2 \leqslant |W'_n| f(1).$$

Zatem przechodząc do granicy,

$$2 = |C| f(1) = \pi d f(1).$$

Przypomnijmy sobie, że celem było obliczenie prawdopodobieństwa, że igła długości ℓ przetnie krawędź którejś deski.

$$\mathbb{P}(Y_\ell = 1) = \mathbb{E}(Y_\ell) = f(\ell) = \ell f(1) = \frac{2\ell}{d\pi}.$$

[2] https://math.stackexchange.com/questions/2671397/suppose-fx-is-linear-i-e-fx-y-fx-fy-and-monotone-on-infty

Literatura

[1] Martin Aigner and Günter M. Ziegler. *Proofs from THE BOOK*. Springer Publishing Company, Incorporated, 6th edition, 2018. pages 189–192, Buffon's needle problem.