RP dla bioinformatyki, pisemne zadania domowe I

Termin: 14.11.2018, godz. 10.15.

1. Anatol rzuca monetą. Siłą rzeczy wyniki układają się krótsze lub dłuższe serie: najpierw wypada pewna liczba (0, 1, 2, ...) kolejnych orłów, potem pewna liczba (1, 2, ...) kolejnych reszek, itd. Przykładowo, jeśli Anatol wyrzucił kolejno

$$RROROR RROO,$$

to serii było łącznie 6 (dwie reszki, orzeł, reszka, orzeł, trzy reszki, dwa orły).

Po wykonaniu pewnej liczby rzutów Anatol poinformował nas, że w sumie wyrzucił k razy orła i l razy reszkę.

a) Uzasadnić, że prawdopodobieństwo, że było dokładnie $2n$ serii jest równe

$$p_{2n} = \frac{2\binom{k-1}{n-1}\binom{l-1}{n-1}}{\binom{k+l}{k}}.$$

b) Jakim wzorem dane jest p_{2n+1} – prawdopodobieństwo, że serii było $2n+1$?

c) Anatol, Błażej oraz Cyryl rzucają po 30 razy monetą. Jakie jest prawdopodobieństwo, że wszyscy trzej otrzymają po dokładnie 20 reszek i 10 orłów?

d) Anatol, Błażej i Cyryl twierdzą, że wszyscy otrzymali 20 reszek i 10 orłów, a na dowód tego prezentują nam zapisane przez siebie wyniki:
   A: RRORRORRORRORRORRORRORRORRO,
   B: RRRROORRRRRROORRROOORRRRORRROO,
   C: ORRRRROOORRRRORRORROOROORRRRRO.
   W świetle rozważań z podpunktów a) i b), który z tych trzech ciągów wygląda najbardziej, a który najmniej podejrzanie?

2. W jeziorze jest N ryb (ich liczba jest ustalona, ale nieznana). Z jeziora wyłowiono m ryb, oznakowano je i wypuszczono z powrotem do wody. Po pewnym czasie wyłowiono n ryb, a wśród nich było k oznakowanych. Za rozsądną ocenę liczby ryb w jeziorze można uznać liczbę ryb, dla której zrealizowało się zdarzenie o największym prawdopodobieństwie. Jaka to liczba? Co wychodzi dla $m = 200$, $n = 100$, $k = 8$?

Pytania kontrolne.[1] Czy nasze postulowane oszacowanie N jest na pewno niemniejsze niż $\max\{m, n\}$? Co jeśli okazało się, że $k = m$? A co jeśli $k = n$? A jeśli $k = 0$?

3. Rozważamy schemat klasyczny $(\Omega, 2^\Omega, \mathbb{P})$, $\mathbb{P}(\{\omega\}) = \frac{1}{\#\Omega}$ dla $\omega \in \Omega$, przy czym $\#\Omega$ jest liczbą pierwszą. Wykazać, że w takiej przestrzeni nie ma nietrywialnych zdarzeń niezależnych (tj. jeśli zdarzenia A i B są niezależne, to co najmniej jedno z nich jest równe $\emptyset$ lub Ω).

4. Brat Tuck, Mały John i Robin Hood zabawiają się strzelaniem z łuku. Brat Tuck trafia w środek tarczy z prawdopodobieństwem $0,4$, Mały John – $0,7$, zaś Robin Hood – $0,9$.

a) Jakie jest prawdopodobieństwo, że losowo wybrany strzelec trafi w środek tarczy?

b) Losowo wybrany z nich strzelił i trafił w środek tarczy. Jakie jest prawdopodobieństwo, że to Robin Hood strzelał?

c) Losowo wybrany z nich strzelił trzykrotnie i za każdym razem trafił w środek tarczy. Jakie jest prawdopodobieństwo, że to Robin Hood strzelał?

d) Trzykrotnie losowano strzelca i za każdym razem wylosowany strzelec trafił w środek tarczy. Jakie jest prawdopodobieństwo, że za każdym razem strzelał Robin Hood?

e) Wszyscy naraz oddają strzał. Jakie jest prawdopodobieństwo tego, że co najmniej dwie strzały utkwią w środku tarczy?

---

[1] Nie trzeba na nie pisemnie odpowiadać, ale – jak zawsze! – warto się zastanowić, czy otrzymany wynik jest choć trochę sensowny i co się dzieje w rozmaitych „skrajnych” przypadkach.

f) Wszyscy naraz oddają strzał; wiemy, że co najmniej jedna strzała utkwiła w środku tarczy. Jakie jest prawdopodobieństwo, że dokładnie dwie strzały utkwiły w środku tarczy?

g) Wszyscy naraz oddali strzał, ale tylko jedna strzała utkwiła w środku tarczy. Jakie jest prawdopodobieństwo, że to brat Tuck trafił?

5. Rzucamy sześcienną kostką, na której liczby oczek na ściankach są określone przez kolejne cyfry naszego numeru indeksu (proszę podstawić odpowiednie spersonalizowane dane). Niech X będzie liczbą wyrzuconych oczek.

a) Opisać rozkład zmiennej losowej X – ile wynosi $\mathbb{P}(X = n)$ dla $n \in \{0, 1, \ldots, 9\}$?

b) Naszkicować dystrybuantę zmiennej losowej X.

c) Znaleźć wartość oczekiwaną zmiennej losowej X oraz jej wariancję.

d) Wyznaczyć wszystkie mediany zmiennej losowej X, tzn. wszystkie liczby rzeczywiste M, takie że $\mathbb{P}(X \leqslant M) \geqslant 1/2$ oraz $\mathbb{P}(X \geqslant M) \geqslant 1/2$.

e) Wyznaczyć wszystkie kwantyle rzędu $1/3$ zmiennej losowej X, tzn. wszystkie liczby rzeczywiste ρ, takie że $\mathbb{P}(X \leqslant \rho) \geqslant 1/3$ oraz $\mathbb{P}(X \geqslant \rho) \geqslant 2/3$.

6. Gramy w lotto (obstawiamy 6 liczb spośród 49). Przyjmijmy następujący, uproszczony model: wypłata za trafienie szóstki to 2 miliony złotych, piątki – 5000 złotych, czwórki – 200 złotych, trójki – 24 złote. Załóżmy też na razie, że całkowicie pomijamy innych graczy (nie ma kumulacji, nagrody nie są dzielone, etc.).

a) Jakiego zysku powinniśmy się średnio spodziewać w jednej grze (uwzględniając cenę kuponu – 3 zł)? A jeśli wypłata za szóstkę byłaby równa 30 milionów?

b) Żeby zagwarantować sobie wygranie dwóch milionów postanawiamy wypełnić kupony ze wszystkimi możliwymi kombinacjami. Ile pieniędzy zyskamy (uwzględniając cenę wszystkich kuponów i nagrody za trafione piątki, czwórki i trójki)? Ile czasu zajęłoby nam wypełnienie wszystkich kuponów (zakładając optymistycznie, że jesteśmy w stanie wypełnić jeden kupon w sekundę i będziemy w stanie pracować bez przerwy)?

Z oczywistych powodów porzucamy poprzednią strategię i postanawiamy „puszczać” jeden kupon przy każdym losowaniu. Załóżmy, że losowania odbywają się 3 razy w tygodniu, przez 52 tygodnie roku.

c) Ponieważ trafienie szóstki ma dodatnie prawdopodobieństwo, to w nieskończonym ciągu gier p.n. kiedyś trafimy szóstkę. Ile razy musimy średnio zagrać zanim to nastąpi? Ile zajęłoby nam to czasu?

d) Ile razy powinniśmy zagrać, jeśli chcemy, żeby nasza szansa na na trafienie co najmniej jeden szóstki była większa niż jeden promil? Ile zajęłoby nam to czasu?

Nasza gra jest tak świetna, że razem z nami gra w nią wielu innych graczy (co od teraz będziemy uwzględniać). Z tego powodu doszło do kumulacji i wypłata za trafienie szóstki wzrosła do 30 milionów złotych. Jednak w prawdziwym życiu, jeśli kilka osób trafi szóstkę, to nagrodę główną dzieli się pomiędzy nie.

e) Załóżmy, że w naszą grę gra n osób (każda obstawia liczby niezależnie od innych). Jakie jest prawdopodobieństwo tego, że w danej rundzie dokładnie k osób trafiło szóstkę, jeśli wiemy, że co najmniej jedna osoba trafiła w tej rundzie szóstkę? Jaką wartość otrzymamy dla $n = 10^6$ i $k = 2$; $n = 10^7$ i $k \in \{2, 3\}$; $n = 10^8$ i $k \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$? (W Polsce mieszka mniej niż 10^7 osób, ale możemy sobie wyobrazić, że w przypływie euforii każdy grający wypełnia więcej niż jeden kupon).