deltoid 115

Nieprawdopodobne!

Joanna JASZUŃSKA

Zagadnienia związane z prawdopodobieństwem i statystyką bywają zaskakujące i nieintuicyjne. Zdarza się też często, że okazują się one znacznie łatwiejsze niż się na pierwszy rzut oka wydaje.

1. W czasie II Wojny Światowej matematyk Abraham Wald analizował zniszczenia ostrzelanych samolotów wojskowych powracających do amerykańskich baz. Zasugerował on wzmocnienie tych ich części, które według statystyk uszkadzane były *najrzadziej.* Dlaczego?

2. W trakcie I Wojny Światowej żołnierzy, uprzednio noszących na głowie jedynie zwykłe czapki, wyposażono w stalowe hełmy. Wedle statystyk z czasem *wzrosła* liczba rannych w głowę. Dlaczego?

3. Klasa chce iść do kina na jeden z filmów: A, B lub C. Każdy uczeń ma ustalone preferencje. W głosowaniu porównującym filmy A i B większość wybrała A. W głosowaniu pomiędzy B i C większość poparła B. Niestety, kino wycofało film B z repertuaru, pozostały do wyboru tylko A i C. Czy większość woli A od C?

4. Dane są liczby $2018, 31, 12345, 506$. Jaki procent ich sumy stanowi ich średnia?

5. Pracownicy pewnej firmy protestowali, że większość z nich zarabia poniżej średniej (w tej firmie). Szef to zmienił, *obniżając* asystentce pensję. Jak to możliwe?

6. Na loterii jest 150 losów: 5 losów *wygrywasz milion*, 50 losów *wygrywasz kolejny los* i 95 losów *nic nie wygrywasz*. Kupuję jeden los (i jeśli wygrywam kolejne, to je biorę). Z jakim prawdopodobieństwem wygram milion?

7. Na pewien kraj wieki temu czarodziejka rzuciła urok: *każda rodzina może mieć dowolnie wiele córek, ale jeśli urodzi im się syn, nie będą już mieć więcej potomstwa.* Mimo to w tym kraju wciąż jest mniej więcej tyle samo kobiet co mężczyzn. Jak to wyjaśnić?

8. Roztargniona sekretarka ma n listów, n odpowiadających im zaadresowanych kopert, ale wkłada listy do kopert losowo. Dla $k = 0, 1, 2, \ldots, n$ przez P_k oznaczmy prawdopodobieństwo tego, że *dokładnie* k listów trafi do właściwych kopert. Wykaż, że dla $n > 2018$ zachodzi nierówność $P_0 \cdot P_1 \cdot P_2 \cdot \ldots \cdot P_n < (2{,}7)^{-(n^n)!}$.

9. Losujemy liczbę całkowitą ze zbioru $\{1, 2, \ldots, 10^{100}\}$. Z jakim prawdopodobieństwem nie ma ona w swym zapisie dziesiętnym żadnej cyfry 7?

10. Skuteczność pewnej szczepionki testowano dwukrotnie. W próbie I szczepionka okazała się skuteczna dla 75% kobiet i 60% mężczyzn, w próbie II – dla 50% kobiet i 25% mężczyzn. Czy wynika stąd, że szczepionka ta jest skuteczniejsza dla kobiet?

Rozwiązania

R1. Wald analizował zniszczenia samolotów, które *wróciły*, a więc których uszkodzenia nie uniemożliwiły im dalszego lotu. Najrzadziej obserwowane uszkodzenia zinterpretował jako najniebezpieczniejsze – to one powodowały, że dany samolot nie wracał. □

R2. Dzięki nowym hełmom żołnierze ranni w głowę częściej przeżywali. Wcześniej widnieliby w statystykach jako *polegli*, nie zaś jako ranni. □

Powyższe dwa zadania to przykłady tzw. *błędu przeżywalności.*

R3. Niekoniecznie. Niech na przykład 10 z 30 osób ma preferencje $[A, B, C]$ (najbardziej chcą obejrzeć A, najmniej C), 10 osób uważa, że $[B, C, A]$, a 10, że $[C, A, B]$. Wtedy A z B wygrywa $20 : 10$, podobnie B z C, ale tak samo C z A! □

Jest to tzw. *paradoks Condorceta.*

R4. Średnia czterech liczb stanowi 25% ich sumy. □

R5. Początkowo szef zarabiał 11 tys., a każdy z pozostałych pracowników po 10 tys. Średnia płaca była więc ciut powyżej 10 tys. i wszyscy prócz szefa zarabiali poniżej średniej.

Po protestach szef nadal zarabia 11 tys., asystentka 8 tys., a pozostali pracownicy po 10 tys. Średnia jest teraz odrobinę poniżej 10 tys. i wszyscy prócz asystentki zarabiają powyżej średniej. □

R6. Rozważmy inną loterię, na której jest 100 losów: 5 losów *wygrywasz milion* i 95 losów *nic nie wygrywasz.* Prócz tego do pudełka z losami wpadło 50 liści – w razie wylosowania liścia należy go wyrzucić i losować raz jeszcze. Prawdopodobieństwo wygranej jest równe $5/100$ (nie zależy od losowania liści, drapania się po głowie itp.). Ponieważ ta loteria nie różni się de facto od opisanej w zadaniu, więc odpowiedź jest ta sama. □

R7. Każde rodzące się dziecko jest na 50% chłopcem, a na 50% dziewczynką – nie zależy to od liczby i płci posiadanego przez nie starszego rodzeństwa. □

R8. Jeśli $n-1$ listów trafi do właściwych kopert, to n-ty też musi trafić, więc $P_{n-1} = 0$. □

R9. Każdą z liczb od 1 do 10^{100} można zapisać, używając dokładnie 100 cyfr, początkowe z nich mogą być zerami (liczbę 10^{100} kodują same zera). Liczb bez cyfry 7 jest 9^{100}, bo na każdym ze 100 miejsc zapisu dziesiętnego można wybrać dowolną spośród 9 cyfr różnych od 7. Wobec tego szukane prawdopodobieństwo równe jest $9^{100}/10^{100} = (0{,}9)^{100} \approx 0{,}00002656$. □

R10. Niekoniecznie. Jeśli np. w próbie I uczestniczyły 4 kobiety i 100 mężczyzn, a w próbie II testowano 100 kobiet i 4 mężczyzn, to łącznie szczepionka okazała się skuteczna dla $3 + 50$ spośród 104 kobiet oraz dla $60 + 1$ z takiej samej liczby mężczyzn. □

…ale może przebadano np. tylko młodzieńców i sędziwe damy? Problem ten znany jest jako *paradoks Simpsona.*