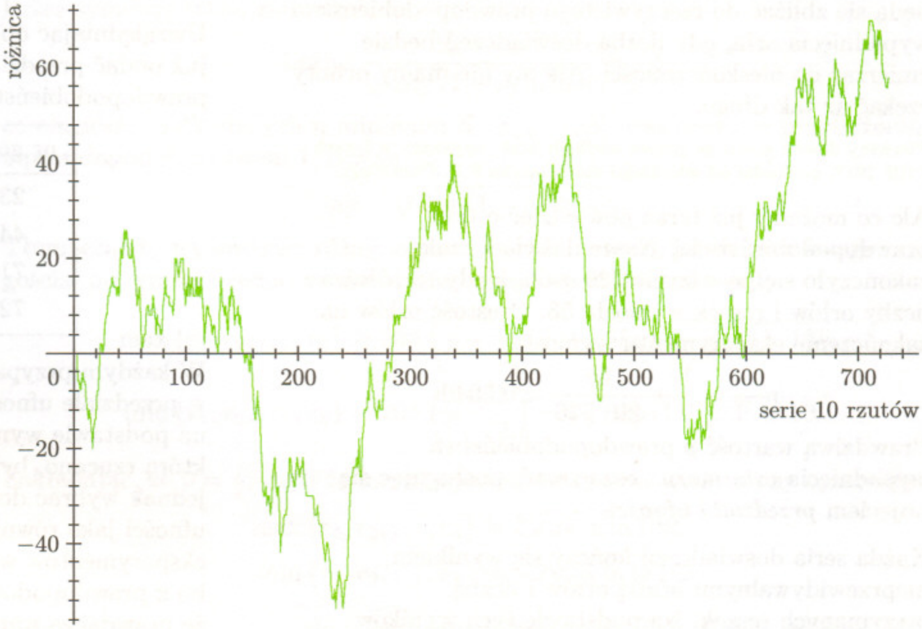
# Co zrobić z ORRO?

*Andrzej DĄBROWSKI*

Jednym z doświadczeń wykonywanych w ubiegłym roku przez gości Festiwalu Nauki w Warszawie było rzucanie monetą. Każdy z uczestników doświadczenia rzucał monetą 10 razy. Do komputera wpisywano różnicę między uzyskaną liczbą orłów i reszek. Na ekranie komputera, po wizycie $k$-tego gościa, rysowany był wykres krzywej $r_k$, będącej różnicą liczby orłów i reszek, uzyskanych dotąd w eksperymencie.

Opis doświadczenia z rzucaniem monetą znaleźć można w prawie każdym podręczniku rachunku prawdopodobieństwa. Rzadko jednak spotkamy zapis wyników takiego doświadczenia, bo zapewne autorzy podręczników stwierdzili, że oprócz „smrodku dydaktycznego" żadnych ciekawych wniosków nie można wyciągnąć z konkretnych danych. Bo cóż ciekawego jest w nudnym ciągu wyników ORRO…?




Patrząc na wykres sumy różnic między liczbą orłów i reszek, uzyskany w doświadczeniu z Festiwalu, zauważyć można kilka niepokojących faktów. Gdyby moneta, którą rzucano, była symetryczna, to różnice powinny oscylować w pobliżu 0. Tymczasem wykres przez większą część czasu przebywał ponad osią $x$, wskazując na przewagę liczby orłów nad liczbą reszek. Po okresie załamania koniunktury orłów (w okresie od momentu pojawienia się 150. gościa aż do gościa nr 237, gdzie przewaga liczby reszek nad liczbą orłów wyniosła 54) i po okresie burzliwej, ale dodatniej stabilizacji (między wizytą trzechsetnego a pięćsetnego gościa) zaczęła się hossa, kończąca się w momencie zakończenia doświadczenia na różnicy, wynoszącej 70 orłów. Szkoda, że już nikt nie zobaczy, jak długo trwałaby ta hossa i jakim wynikiem mogłaby się zakończyć. I szkoda, że nie są to notowania spółki giełdowej, w którą zainwestowaliśmy nasze oszczędności.

Czy taki przebieg doświadczenia może nam coś powiedzieć o prawdopodobieństwie wyrzucenia orła? Czy słusznie podejrzewamy, że moneta jest asymetryczna i jak wielka jest ta asymetria? Zazwyczaj do oszacowania prawdopodobieństwa zdarzenia $A$ bierze się jego częstość w serii $n$ doświadczeń, czyli ułamek postaci

$$p_n(A) = \frac{l_n(A)}{n},$$

gdzie $l_n(A)$ jest liczbą doświadczeń, w których wystąpiło zdarzenie $A$.

Oznaczając przez $l_{10k}(„O")$ liczbę orłów po wizycie $k$-tego gościa, a liczbę reszek przez $l_{10k}(„R")$, otrzymamy równość

$$r_k = l_{10k}(„O") - l_{10k}(„R") = 2l_{10k}(„O") - 10k.$$

Częstość orłów po wizycie $k$-tej osoby można więc zapisać wzorem

$$p_k = \frac{1}{2} + \frac{r_k}{20k}.$$

Obserwacje krzywej różnic pozwalają nam oszacować prawdopodobieństwo wyrzucenia orła w dowolnym momencie naszego doświadczenia. W momencie absolutnej bessy dla orłów częstość wypadnięcia orła była równa

$$p_{237} = \frac{1}{2} + \frac{-54}{20 \cdot 237} \approx 0{,}4889.$$

Kiedy po kolejnych wahaniach krzywa różnic wzniosła się na wysokość 46 po wizycie 440 osoby, częstość orłów była równa

$$p_{440} = \frac{1}{2} + \frac{46}{20 \cdot 440} \approx 0{,}5052.$$

Najwyższą wartość różnicy liczby orłów i reszek, równą 70, osiągnięto po wizycie 710 osoby. Oznacza to, że w tym momencie częstość orłów była równa

$$p_{710} = \frac{1}{2} + \frac{70}{20 \cdot 710} \approx 0{,}5049,$$

co mogłoby wskazywać na stabilizację wokół wartości 0,505. Nie wiadomo, jak dalej potoczą się losy naszej krzywej i związanej z nią częstości orłów.

Twierdzenie, zwane mocnym prawem wielkich liczb, mówi, że z prawdopodobieństwem 1 częstości orłów będą się zbliżać do rzeczywistego prawdopodobieństwa wypadnięcia orła, gdy liczba doświadczeń będzie zmierzać do nieskończoności. Ale my nie mamy ochoty czekać aż tak długo.

Pierwszy dowód mocnego prawa wielkich liczb, uzyskany w latach 1916–1917, pochodzi od włoskiego matematyka F.P. Cantellego.

Ale co możemy już teraz powiedzieć o tym prawdopodobieństwie? Nasze doświadczenie zakończyło się po wizycie 726 osób, kiedy to różnica liczby orłów i reszek wyniosła 58. Częstość orłów na zakończenie eksperymentu wyniosła

$$p_{726} = \frac{1}{2} + \frac{58}{20 \cdot 726} \approx 0{,}5040.$$

Prawdziwą wartość $p$ prawdopodobieństwa wypadnięcia orła można oszacować, posługując się pojęciem *przedziału ufności*.

Każda seria doświadczeń kończy się wynikiem nieprzewidywalnym: liczbą orłów i liczbą otrzymanych reszek. Na podstawie tych wyników możemy skonstruować przedział (też losowy) o tej własności, że prawdopodobieństwo, iż zawiera on liczbę $p$, jest co najmniej równe $\gamma$. Taki przedział nazywamy *przedziałem ufności na poziomie ufności* $\gamma$. W przypadku przedziału, zawierającego prawdopodobieństwo $p$ wyrzucenia orła, przedział taki ma postać

$$\left[p_n - d_n(\gamma), p_n + d_n(\gamma)\right].$$

Oznacza to, że jego środek jest równy częstości $p_n$, a błąd oszacowania $p$ wyraża się wzorem

$$d_n(\gamma) = z(\gamma)\sqrt{\frac{p_n(1-p_n)}{n}}.$$

Współczynnik $z(\gamma)$, wyrażający wpływ poziomu ufności $\gamma$ na błąd oszacowania parametru $p$, łatwo obliczyć, gdy wykonamy dużą liczbę rzutów $n$.

Wystarczy, że liczba pomiarów jest większa od 50.

W tym przypadku $z(\gamma)$ jest liczbą o tej własności, że zmienna losowa $Z$ o standardowym rozkładzie normalnym spełnia warunek $|Z| \le z(\gamma)$ z prawdopodobieństwem równym $\gamma$.

Rozkład normalny nazywamy standardowym, jeśli jego wartość oczekiwana jest równa 0, a odchylenie standardowe 1.

Zazwyczaj poziomy ufności wybiera się równe 0,95 lub 0,99, co jest równoważne przyjęciu wartości $z(\gamma) = 1{,}96$ lub odpowiednio $z(\gamma) = 2{,}58$.

Wróćmy do danych, zebranych w trakcie eksperymentu: oszacowanie prawdopodobieństwa wyrzucenia orła po wizycie 237. gościa za pomocą przedziału ufności na poziomie 0,95 miało błąd równy

$$1{,}96\sqrt{\frac{0{,}4889\,(1-0{,}4889)}{10 \cdot 237}} \approx 0{,}0201.$$

Dokładność przedziału ufności zwiększała się w miarę zbierania informacji

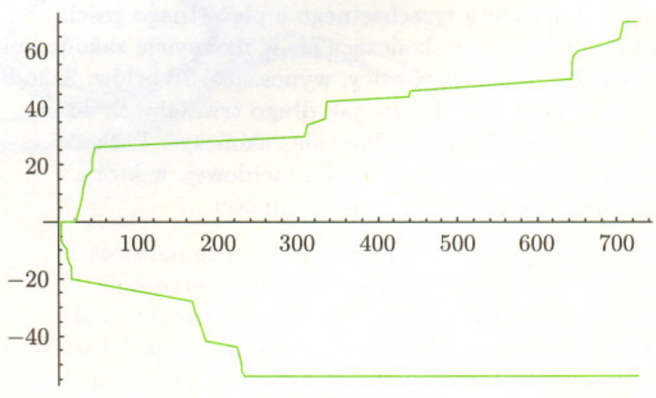
| nr gościa | 237 | 440 | 710 | 726 |
|---|---|---|---|---|
| błąd | 0,0201 | 0,0148 | 0,0116 | 0,0115 |

Uwzględniając obliczone wartości częstości, możemy już podać przedziały ufności na poziomie 0,95 dla prawdopodobieństwa wyrzucenia orła.

| nr gościa | przedział ufności |
|---|---|
| 237 | [0,4688 , 0,5090] |
| 440 | [0,4904 , 0,5200] |
| 710 | [0,4933 , 0,5163] |
| 726 | [0,4925 , 0,5155] |

W każdym przypadku wartość $p = \frac{1}{2}$ mieściła się w przedziale ufności, co oznacza, że moglibyśmy uznać na podstawie wyników tego doświadczenia, iż moneta, którą rzucano, była symetryczna. Sceptyk może jednak wybrać dowolnie inną wartość z przedziału ufności jako równie dobrą. Po zakończeniu wszystkich eksperymentów wiemy jedynie, i to nie na pewno, bo z prawdopodobieństwem co najmniej 0,95, że prawdziwa wartość prawdopodobieństwa wyrzucenia orła mieści się w przedziale [0,4925 , 0,5155].

Omawiając fluktuacje różnicy liczby orłów i reszek, zwróciliśmy uwagę na wartości maksymalne i minimalne tej różnicy. Maksimum obliczone po $k$ kolejnych wizytach gości jest losową, niemalejącą funkcją $k$. Podobnie, minimum jest losową, nierosnącą funkcją $k$. Jak wielkie wartości mogą osiągnąć te funkcje w miarę wzrostu liczby przeprowadzonych rzutów? Oto wykres kolejnych maksimów i minimów w zależności od numeru kolejnego uczestnika doświadczenia.



Jak widać, maksima rosną powoli, minima ustabilizowały się na poziomie wartości $-54$. Interesująca byłaby odpowiedź na pytanie, czy maksima mogłyby rosnąć do nieskończoności, a minima maleć do minus nieskończoności, gdyby obserwacje trwały bardzo długo.

Rosyjski matematyk A. Chinczyn udowodnił w 1924 roku twierdzenie, dotyczące tempa wzrostu liczby sukcesów $S_n$ w $n$ kolejnych niezależnych próbach, w których prawdopodobieństwo sukcesu jest stałe i wynosi $p$. Twierdzenie to nosi nazwę *prawa iterowanego logarytmu*. Stwierdza ono, że z prawdopodobieństwem 1 zachodzi równość

$$\limsup \frac{S_n - np}{\sqrt{2np(1-p)\ln\ln n}} = 1.$$

Równość ta oznacza, że dla dużych $n$ maksimum $S_1, S_2, \ldots, S_n$ jest równe w przybliżeniu, z prawdopodobieństwem 1, liczbie

$$np + \sqrt{2np(1-p)\ln\ln n}.$$

Przez symetrię łatwo zauważyć, że

$$\liminf \frac{S_n - np}{\sqrt{2np(1-p)\ln\ln n}} = -1,$$

co oznacza, że dla dużych $n$ minimum $S_1, S_2, \ldots, S_n$ jest równe w przybliżeniu, z prawdopodobieństwem 1, liczbie

$$np - \sqrt{2np(1-p)\ln\ln n}.$$

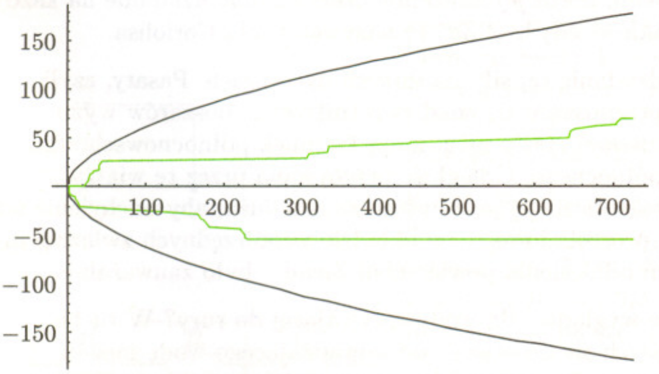
Tłumacząc to na wartości różnic $r_k$ liczby orłów i reszek po wizycie $k$-tego gościa, otrzymamy oszacowanie ważne z prawdopodobieństwem 1

$$\max\{r_1, r_2, \ldots, r_k\} \approx 20k\left(p - \frac{1}{2}\right) + 4\sqrt{5kp(1-p)\ln\ln 10k},$$
$$\min\{r_1, r_2, \ldots, r_k\} \approx 20k\left(p - \frac{1}{2}\right) - 4\sqrt{5kp(1-p)\ln\ln 10k}.$$

Zakładając, że $p = \frac{1}{2}$, czyli że moneta jest symetryczna, otrzymamy oszacowanie

$$\max\{r_1, r_2, \ldots, r_k\} \approx 2\sqrt{5k\ln\ln 10k},$$
$$\min\{r_1, r_2, \ldots, r_k\} \approx -2\sqrt{5k\ln\ln 10k}.$$

Ze wzorów tych wynika, że maksima i minima w miarę zwiększania liczby rzutów będą zmierzały do odpowiednich nieskończoności, choć dość wolno. Porównując na tym samym wykresie rzeczywiste wartości maksymalnych i minimalnych różnic z oszacowaniem z prawa iterowanego logarytmu, zauważymy, że nasze obawy, iż różnice liczby orłów i reszek są duże, były bezpodstawne. Po wizycie ostatniego gościa moglibyśmy oczekiwać nawet różnicy równej 175 (lub $-175$).




Zjawiska, związane z tak prostym doświadczeniem, jak rzut monetą, dostarczają nam, jak widać, materiału do wielu niebanalnych rozważań. Pozostało nam do rozstrzygnięcia jeszcze wiele ciekawych pytań.

Obserwując różnice między liczbą orłów i reszek, zauważamy, że czasami ta różnica wynosi 0, co oznacza, iż w tych momentach zrównała się liczba orłów i reszek. Stało się tak 12 razy: podczas wizyty gości numer 1, 24, 59, 127, 152, 153, 282, 388, 462, 471, 539, 594. Średni czas powrotu do zera wyniósł około 50 gości. Od ostatniego zera odwiedziły stoisko jeszcze 132 osoby i nie widać nadziei na wyrównanie orłów i reszek. W końcu doświadczenia byliśmy świadkami hossy orłów, różnica doszła do wartości 70. Czy kiedykolwiek powrócimy do 0? Jaki jest naprawdę średni czas powrotu?

Średni czas powrotu do różnicy równej $-6$ też wynosi około 50 gości. Czy tak będzie zawsze? Dla dowolnej wartości różnicy? Czy, obserwując w momencie $k$ jakąś różnicę, możemy po dostatecznie długim czasie dojść do dowolnej pomyślanej przez nas różnicy? Jeśli tak, to z jakim prawdopodobieństwem?

Mam nadzieję, że pytania, które tu postawiłem, zainteresowały niektórych Czytelników. Być może będzie okazja, aby do nich powrócić.