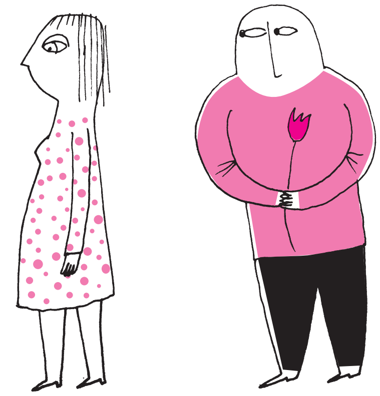
Matrymonialna matematyka

*Magdalena SIENIAWSKA**

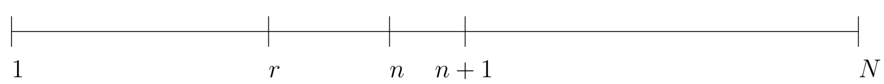
*Doktorantka, Centrum Astronomiczne im. Mikołaja Kopernika, PAN



Kiedy jest najlepszy moment na powiedzenie sakramentalnego „tak”? Czy obecny obiekt westchnień jest osobą, z którą spędzisz resztę życia? Ile związków musi się rozpaść, żeby stworzyć stabilną relację? Pewnie wiele osób zadaje (lub wkrótce zada) sobie takie pytania. Ciężko uwierzyć, że matematyka i algorytmika mogą nam pomóc również w romantycznych rozważaniach i sercowych dylematach. Wszystko sprowadza się przecież do określenia momentu, kiedy należy przerwać szukanie i podjąć decyzję, a ten problem jesteśmy w stanie wymodelować matematycznie.

Zagadnienie jest znane w kręgach matematycznych jako *problem sekretarki.* Nie wiadomo, kto i kiedy jako pierwszy sformułował tę zagadkę, ale swoją popularność zyskała w 1960 roku po publikacji lutowego numeru czasopisma *Scientific American.* Oryginalnie problem był postawiony w następujący sposób: szef biura pilnie potrzebuje sekretarki, w związku z czym ogłasza nabór na to stanowisko. Na ogłoszenie odpowiada N kandydatek, z którymi szef przeprowadza rozmowy kwalifikacyjne. Po każdym wywiadzie rekruter ma dwie opcje: przyjąć kandydatkę i tym samym zakończyć poszukiwania bądź odmówić rekrutowanej osobie i zaprosić następną kandydatkę. Zakładamy, że (np. z racji dużego ruchu na rynku pracy) nie ma możliwości zmiany raz podjętej decyzji. Kiedy jest najlepszy moment na zakończenie rekrutacji, jeśli celem jest wybranie najlepszej kandydatki spośród wszystkich N zgłoszonych?

Intuicyjnie wiemy, że decyzji nie należy podejmować zbyt szybko (wówczas jest duża szansa, że nigdy nie spotkamy najlepszej kandydatki), ani zwlekać zbyt długo (najlepszą kandydatkę możemy wówczas pochopnie odrzucić). Okazuje się, że najbardziej opłaca się szukać właściwej kandydatki dopiero po przesłuchaniu około 37% kandydatek. Skąd akurat taka liczba? Można zauważyć, że najkorzystniejsza strategia polega na podzieleniu rekrutacji na dwie fazy. Najpierw należy zorientować się w kwestii ogólnego poziomu kandydatek, tj. istnieje taka liczba osób r, których aplikacje należy z automatu odrzucić, a same rozmowy kwalifikacyjne traktować jako zbieranie informacji (oczywiście musi być $1 \leqslant r < N$). Po przeprowadzeniu r wywiadów (pula testowa), w drugiej fazie należy wybrać pierwszą kandydatkę, która będzie lepsza od najlepszej z puli testowej. Jakie r (w wersji asymptotycznej: jaki stosunek r/N, gdy $N \to \infty$) da nam największe prawdopodobieństwo wybrania najlepszej sekretarki z całej puli? Zauważmy, że w puli zasadniczej jest $N - r$ kandydatek. Wyżej postawiony problem, z matematycznego punktu widzenia, sprowadza się do znalezienia takiego r, dla którego pewien ciąg liczbowy osiąga maksimum. Rozważmy przypadek, gdy najlepsza kandydatka jest na pozycji $n + 1$. W oczywisty sposób, jeśli $n + 1 < r$, to nie jesteśmy w stanie jej znaleźć. Schemat poniżej ilustruje przypadek, gdy najbardziej kompetentna osoba jest w puli zasadniczej.



Aby najlepsza kandydatka została faktycznie wybrana, musi zajść następujące zdarzenie: najlepsza kandydatka w $[1, r]$ jest również najlepsza w $[1, n]$ (którego prawdopodobieństwo oczywiście wynosi r/n). Ten warunek gwarantuje wytrwanie procesu rekrutacji do rozmowy z kandydatką $n + 1$ (czyli najlepszą z całej grupy N osób). Teraz, aby określić prawdopodobieństwo całkowite sukcesu, sumujemy odpowiednie prawdopodobieństwa warunkowe, przy kolejnych założeniach, że najlepsza kandydatka zajmuje pozycję $n + 1 = r + 1, r + 2, r + 3, \ldots, N$ (przemnożone przez prawdopodobieństwa warunków, które w tym problemie stale wynoszą $1/N$):

$$P(\text{sukces}) = \frac{1}{N}\left[\frac{r}{r} + \frac{r}{r+1} + \frac{r}{r+2} + \cdots + \frac{r}{N-1}\right] \approx -\frac{r}{N} \cdot \ln\left(\frac{r}{N}\right),$$

gdzie ln jest oznaczeniem logarytmu naturalnego (sposób uzyskania wyniku na marginesie). Teraz wystarczy znaleźć maksimum wyrażenia na $P(\text{sukces})$,

$$P(\text{sukces}) = \frac{1}{N}\left[\frac{r}{r} + \frac{r}{r+1} + \cdots + \frac{r}{N-1}\right] = \frac{1}{N}\sum_{n=r}^{N-1}\frac{r}{n} \approx \frac{r}{N}\int_r^N \frac{1}{n}\,dn = -\frac{r}{N}\cdot\ln\left(\frac{r}{N}\right) = -x\cdot\ln(x)$$

Kolejne linie stanowią wyrażenia równoważne:

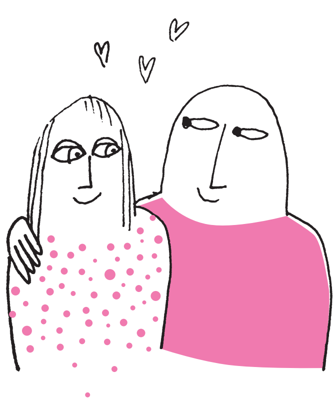
$$[-x \cdot \ln(x)]' = 0$$
$$-1 \cdot \ln(x) - x \cdot x^{-1} = 0$$
$$-\ln(x) - 1 = 0$$
$$\ln(x) = -1$$
$$x = e^{-1}$$

Uwaga: Aby upewnić się, że funkcja $-x \cdot \ln(x)$ na przedziale $[0,1]$ faktycznie osiąga maksimum w punkcie e^{-1}, należy jeszcze sprawdzić punkty brzegowe.

aby ustalić optymalny stosunek $\frac{r}{N}$. (Szukanie maksimum z użyciem pochodnej przyrównanej do zera – na marginesie.) Z naszych rozważań wynika, że wynosi on e^{-1}, czyli że pula testowa kandydatek powinna zawierać $1/e \approx 37\%$ wszystkich kandydatek. Wówczas szansa na znalezienie najlepszej osoby na stanowisko sekretarki wynosi $-e^{-1} \cdot \ln(e^{-1}) = e^{-1} \approx 37\%$.

W literaturze można spotkać się z różnymi modyfikacjami problemu sekretarki. Wynik będzie inny, jeśli możemy wrócić do początkowo odrzuconych osób albo zadowolimy się najlepszą lub prawie najlepszą kandydatką. Warto zauważyć, że tę matematyczną zagadkę można również rozważyć w sposób czasowy, nie ilościowy. Jeśli na przykład nabór na stanowisko sekretarki trwa dokładnie miesiąc, a nie wiemy, ile dostaniemy CV, to optymalną pulą testową będzie pierwsze 11 dni ($37\% \cdot 30$ dni), niezależnie od ilości rozmów kwalifikacyjnych. Następnie powinniśmy zatrudnić pierwszą osobę lepszą od osób z puli testowej.

Problem sekretarki dzięki swojej prostocie i jednoznacznemu rozwiązaniu znajduje zastosowanie w wielu dziedzinach życia, także codziennego. Matematyka podpowiada, że powinniśmy postępować zgodnie z regułą 37%, jeśli chodzi np. o wynajem mieszkania. Przeprowadzając się do nowego miasta (np. z powodu wyjazdu na studia), najczęściej nie mamy wiedzy, jaki jest tam rynek nieruchomości. Jeżeli szukamy mieszkania do wynajęcia, warto ustalić sobie limit, np. 30 potencjalnie interesujących lokali, z czego pierwsze 11 potraktować jako wstępne oględziny rynku nieruchomości. Podobnie rzecz ma się z randkami i ślubami, które także można sprowadzić do problemu sekretarki. Umawianie się za pośrednictwem portali randkowych niekiedy przypomina rozmowy rekrutacyjne. Z puli interesujących nas współużytkowników portalu staramy się wybrać „tę najlepszą” lub „tego najlepszego”. Schemat postępowania powinien być podobny: ustalić liczbę spotkań, na jakie jesteśmy gotowi, pierwsze 37% potraktować wyłącznie jako zbieranie informacji, a z puli pozostałych osób wybrać pierwszą lepszą niż kandydaci/-tki z puli testowej.



W przypadku bardziej poważnych decyzji, takich jak ślub i ustatkowanie się, także możemy zrobić użytek z problemu sekretarki. Zauważmy, że nie jesteśmy w stanie oszacować, ilu potencjalnych partnerów spotkamy w życiu. W takim przypadku lepiej jest się odnieść do limitu czasowego: z przyczyn biologicznych i prawnych najlepszym czasem na założenie rodziny jest wiek między 18 a 35 lat. Mamy więc w sumie $35 - 18 = 17$ lat na zbieranie informacji i podjęcie decyzji o ustatkowaniu się. Z powyższych rozważań wiemy już, że przez pierwsze $37\% \cdot 17$ lat ≈ 6 lat powinniśmy poświęcić wyłącznie gromadzeniu danych. W wieku ok. $18 + 6 = 24$ lat przychodzi moment, gdy musimy być gotowi na podjęcie ostatecznej decyzji.

Problem sekretarki był wielokrotnie testowany na ochotnikach podczas badań psychologicznych i behawiorystycznych. Naukowcy chcieli sprawdzić, czy ludzie intuicyjnie skłaniają się ku zasadzie 37%. Okazuje się, że nie. Z badań wynika, że rodzaj ludzki jest raczej niecierpliwy i przystępuje do podejmowania decyzji już po ok. 31% puli. Na zbyt szybkim podejmowaniu decyzji można stracić, np. gdy na wycieczce szukamy kawiarni. Przedwczesne zdecydowanie się na filiżankę cappuccino może sprawić, że bezsensownie przepłacimy, podczas gdy tańsza kawiarnia jest tuż za rogiem. O ile zapłacenie zbyt wiele za kawę nie jest tragedią, o tyle w przypadku sprzedaży domu czy zamążpójścia zbyt szybko podjęta decyzja może być brzemienna w skutkach. Warto o tym pamiętać, kiedy następnym razem będziemy podejmować decyzję, i być może powstrzymać się przed zbyt szybkim i pochopnym przerwaniem poszukiwań.

Autorka zaznacza, że powyższe rozważania są jedynie ciekawostką matematyczną, nieuwzględniającą wszystkich możliwych scenariuszy i parametrów związanych z prawdziwym życiem. Ostatecznie, jak zauważył Mark Twain: „Są trzy rodzaje kłamstw: kłamstwa, bezczelne kłamstwa i statystyki”.

Więcej o informatycznych algorytmach i ich odniesieniu do codziennego życia można przeczytać w książce *Algorytmy. Kiedy mniej myśleć* autorstwa Toma Griffithsa i Briana Christiana.