Jest to zapis odczytu wygłoszonego na XLII Szkole Matematyki Poglądowej, *Dowody i kontrprzykłady*, styczeń 2009.

Piękne i bestie

Grzegorz KOSIOROWSKI, Kraków

Matematyka. Nauka czystego intelektu. Bezduszna, nieludzka, nie zwracająca uwagi na kryteria estetyczne. Zawsze stawiająca precyzję ponad urodą. I za tym wzorem podążają matematycy – akolici surowej bogini. Wszyscy wiemy, że nie ma chyba nic bardziej fałszywego niż ten ciąg stereotypowych skojarzeń. Takie opinie jak „W swojej pracy starałem się łączyć prawdę i piękno, lecz gdy musiałem wybierać, wybierałem piękno" Hermanna Weyla lub „Nie ma na świecie miejsca dla brzydkiej matematyki" Godfreya Hardy'ego, choć w oczywisty sposób żartobliwe i przesadzone, są bliskie większości matematyków.

Jednak, choć doceniamy nietrywialne piękno prac matematycznych, nie jesteśmy zgodni, co do kryteriów matematycznego piękna i brzydoty. Zgodnie z tematyką ostatniej Szkoły Matematyki Poglądowej w Grzegorzewicach „Dowody i kontrprzykłady" artykuł ten skoncentruje się na kryteriach estetycznych, wedle których oceniamy matematyczne dowody tego samego twierdzenia (trudno porównać dowody różnych twierdzeń). W końcu każdy słyszał (lub sam wypowiadał) stwierdzenia typu: „Przepiękny dowód twierdzenia..." bądź „Tego dowodu przedstawiać nie będę. Jest długi, techniczny i paskudny". Co decyduje o przyjęciu przez nasze umysły niektórych rozumowań jako pięknych, a innych jako... bestie? Postaram się przeprowadzić w miarę rzetelną (jak na tak nieprecyzyjny temat) analizę, podpartą pewnymi przykładami.

Pierwszym przykładem będzie zadanie, pochodzące ze studenckich Międzynarodowych Zawodów Matematycznych w Ostrawie (z roku 2000). Należało udowodnić poniższe twierdzenie:

Twierdzenie 1. *Niech m, n będą liczbami naturalnymi, a $x \in [0, 1]$. Wtedy*
$$(1-(1-x)^m)^n + (1-x^n)^m \geq 1.$$

Samo zadanie specjalnie urokliwie nie wygląda. Ot, nierówność jak nierówność. Jednakże jeden z dowodów robi duże wrażenie.

Zaczyna się zaskakująco. Rozważamy macierz o złożoną z m wierszy i n kolumn o losowo wybieranych elementach: elementem o danych współrzędnych jest litera A z prawdopodobieństwem x, bądź litera B z prawdopodobieństwem $(1-x)$. Okazuje się, że dla udowodnienia nierówności z twierdzenia wystarczy przeliczyć prawdopodobieństwa kilku zdarzeń. Policzmy, jak można zinterpretować składniki sumy po lewej stronie:

- $(1-x)^m$ – jest to prawdopodobieństwo, że dana kolumna składa się wyłącznie z liter B. Zatem:
- $1-(1-x)^m$ – jest to prawdopodobieństwo, że w danej kolumnie występuje litera A. Zatem:
- $(1-(1-x)^m)^n$ – jest to prawdopodobieństwo, że w każdej kolumnie występuje litera A.

oraz:

- x^n – jest to prawdopodobieństwo, że dany wiersz składa się wyłącznie z liter A. Zatem:
- $1-x^n$ – jest to prawdopodobieństwo, że w danym wierszu występuje litera B. Zatem:
- $(1-x^n)^m$ – jest to prawdopodobieństwo, że w każdym wierszu występuje litera B.

Zauważmy, że jeśli oznaczymy następujące zdarzenia: X – w każdej kolumnie występuje litera A, Y – w każdym wierszu występuje litera B, to ich suma jest zdarzeniem pewnym. Gdyby bowiem zdarzenie X nie zachodziło, to istniałaby kolumna złożona z samych liter B, czyli w każdym wierszu litera B musiała by się znaleźć. Zgodnie z tymi rozważaniami mamy:
$$(1-(1-x)^m)^n + (1-x^n)^m = P(X) + P(Y) \geq P(X \cup Y) = 1,$$
co było do udowodnienia.

Autor ma bardzo osobisty stosunek do powyższego dowodu. Zwięzłość i jasność rozumowania oraz zaskakujący pomysł przewodni są przez niego obarczone dużą dozą odpowiedzialności za wybór studiów. Nic dziwnego, że chętnie dzielił się tym zachwytem ze znajomymi matematykami, którzy w zdecydowanej większości zgadzali się z tą oceną. Ale nie wszyscy, co jest mało zaskakujące w tak subiektywnej dziedzinie jak estetyka...

Pewien Znajomy Autora Ze Studiów (w skrócie odtąd zwany PZAZS) podał argument zaskakujący: „Ten dowód jest brzydki, bo używamy rachunku przawdopodobieństwa!". Co ciekawsze, nie był w swojej opinii zupełnie odosobniony. Co zatem PZAZS uważał za ładniejszy dowód? Oto szkic rozumowania, które po pewnym czasie przedstawił:

Dowód „bez prawdopodobieństwa" rozpoczyna się od udowodnienia następującej nierówności pomocniczej:

$$\forall m \in \mathbb{N}\ \forall a, b \in [0,1]\ \ : (a+b-ab)^m + (1-a^m)(1-b^m) \geq 1,$$

a właściwie równoważnej nierówności: $(a+b-ab)^m - a^m - b^m + (ab)^m \geq 0$.

Dowód prowadzony jest przez indukcję. Dla $m = 1$ równość jest oczywista. Założenie indukcyjne przyjmuje postać:

$$(a+b-ab)^{m-1} \geq a^{m-1} + b^{m-1} - (ab)^{m-1}.$$

Przemnażając go przez dodatnią liczbę $(a+b-ab)$ i odejmując stronami dostajemy:

$$\begin{aligned}(a+b-ab)^m - a^m - b^m + (ab)^m &\geq \\ &\geq a^{m-1}b - a^m b + ab^{m-1} - ab^m - a^m b^{m-1} - a^{m-1}b^m + 2a^m b^m = \\ &= (b^{m-1} - b^m)(a - a^m) + (a^{m-1} - a^m)(b - b^m) \geq 0.\end{aligned}$$

Zatem nierówność pomocnicza jest udowodniona. Tezę twierdzenia dowodzimy przez indukcję względem n. Oczywiście, dla $n = 1$ zachodzi równość. Założenie indukcyjne ma postać: $(1-(1-x)^m)^{n-1} \geq 1 - (1-x^{n-1})^m$. Mnożymy je obustronnie przez $1-(1-x)^m$, otrzymując:

$$(1-(1-x)^m)^n \geq (1-(1-x^{n-1})^m)(1-(1-x)^m).$$

Oznaczamy $a = 1 - x^{n-1}$, $b = 1 - x$ i, korzystając z założenia indukcyjnego oraz nierówności pomocniczej, przeliczamy:

$$\begin{aligned}(1-x^n)^m + (1-(1-x)^m)^n &\geq \\ &\geq (1 - x^{n-1}(1-(1-x)))^m + (1-(1-x^{n-1})^m)(1-(1-x)^m) = \\ &= (a + (1-a)b)^m + (1-a^m)(1-b^m) \geq 1.\end{aligned}$$

□

Drugi dowód jest jak najbardziej poprawny i dość zwięzły (choć jego precyzyjny zapis zawierający wszystkie przekształcenia powinien zająć znacznie więcej miejsca niż szkic przedstawiony powyżej). Wydaje się jednak, że na większości matematyków nie robi on takiego wrażenia jak jego „probabilistyczny" konkurent. Zanim spróbuję wyjaśnić przyczyny, przedstawię jeszcze jeden przykład: słynne zadanie hrabiego de Buffon o igle.

Zadanie 2. *Igłę o długości l upuszczamy losowo na arkusz poliniowany prostymi równoległymi, poprowadzonymi w równych odstępach $d > l$. Jakie jest prawdopodobieństwo, że po upadku igła będzie przecinać jedną z linii?*

Jak wiele z XVIII-wiecznych klasycznych problemów, ten też wywodzi się z gier hazardowych, a konkretnie z gry franc-carreau polegającej na rzucaniu okrągłą monetą na podłogę podzieloną na kwadraty i zakładaniu się, czy upadnie na linię. Odpowiedź zagadkę hrabiego jest dość niespodziewana, a wyliczyć ją można dosyć dzięki dość prostej całce.

Należy rozważyć kąt α, jaki igła tworzy z danymi liniami (dla ustalenia uwagi – poziomymi). Można założyć, że $0 < \alpha < \frac{\pi}{2}$, gdyż przypadek kątów ujemnych jest symetryczny. Zauważmy, że wysokość (czyli „pionowa składowa długości" igły wynosi $l \sin \alpha$. Wtedy łatwo stwierdzić, że prawdopodobieństwo, iż igła przecina którąś z linii wynosi $\frac{l \sin \alpha}{d}$. Zatem wystarczy obliczyć średnią tych prawdopodobieństw po danych kątach:

$$P = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{l \sin \alpha}{d} \, d\alpha = \frac{2}{\pi} \frac{l}{d} [-cos\alpha]_0^{\frac{\pi}{2}} = \frac{2}{\pi} \frac{l}{d}.$$

Jak widać, rozwiązanie to nie jest szczególnie skomplikowane. Jednak zadanie znalazło też swoje miejsce w książce „Dowody z Księgi" Martina Aignera i Guntera Zieglera, której tytuł i pomysł zostały zainspirowane znanym powiedzeniem Pála Erdősa o Księdze, w której Bóg gromadzi doskonałe dowody twierdzeń matematycznych. Tak więc książka ta miała przedstawiać wybór pięknych dowodów matematycznych. Między innymi przedstawia alternatywne, efektowne i elementarne (choć może nieco dłuższe) rozwiązanie zagadki Buffona. Jego autorem jest Joseph-Emile Barbier.

Rozważmy $E(l)$ – wartość oczekiwaną punktów przecięcia igły o długości l z danymi liniami. Dla $l < d$ jest $E(l)$ jest naturalnie równe szukanemu prawdopodobieństwu (jako, że igła nie może przeciąć więcej niż jednej linii). Jednocześnie, łatwo sprawdzić, że zależność E od l jest liniowa, zatem mamy dla pewnego rzeczywistego c

$$P = E(l) = cl.$$

Pozostaje wyznaczyć c. W tym celu zauważmy, że kształt „igły" nie ma znaczenia dla powyższego wzoru. Zależność wartości oczekiwanej liczby przecięć od długości igły pozostaje prawdziwa dla igieł w kształcie łamanych. I tu tkwi kluczowy punkt dowodu: rozważamy igłę w kształcie... okręgu o średnicy d. Okrąg taki przecina linie zawsze w dokładnie 2 punktach. Przybliżając okrąg wielokątami i korzystając ze wzoru na wartość oczekiwaną liczby przecięć dla łamanych otrzymamy $cd\pi = E(d\pi) = 2$ i w końcu

$$c = \frac{2}{d\pi}.$$

Pierwsza para przykładów sugeruje, że ważnym, o ile nie najważniejszym, kryterium oceny estetycznej dowodu, jest jego zwięzłość. W naturalny sposób preferujemy dowody krótsze, jako bliższe ideałowi. Przy czym warto zwrócić uwagę, że porównywać w ten sposób można tylko dowody zapisane na tym samym poziomie komplikacji, bez omijania poważnych fragmentów – na przykład drugi dowód w pierwszym przykładzie należy ocenić raczej jako szkic; jego formalny zapis byłby znacznie mniej zwięzły. Takie kryterium miałoby tę zaletę, że jest dość obiektywne (policzenie linijek nie jest trudne), niemniej zupełnie się nie sprawdza w przypadku „dowodu z Księgi" z zadania o igle. No i oczywiście nijak nie uwzględnia nietypowych, indywidualnych kryteriów typu PZAZS.

W celu scharakteryzowania najbardziej typowych kryteriów, autor przeprowadził serię dyskusji oraz ankietę pośród studentów matematyki Uniwersytetu Jagiellońskiego. Oczywiście, wyniki nie są miarodajne ze względu na małą i mało zróżnicowaną próbkę statystyczną, jednakże wydają się wskazywać najistotniejsze tendencje. Założeniem ankiety było, że porównujemy tylko dowody tego samego twierdzenia, które na dodatek zostały zapisane poprawnie, ściśle, jasno, bez luk do uzupełnienia dla czytelników i generalnie maksymalnie elegancko z technicznego punktu widzenia.

Studenci wskazali sześć kryteriów, permanentnie się powtarzających, natomiast każdy przypisywał im inne „wagi". By zilustrować różnorodność poglądów, wystarczy nadmienić, że każdy z wymienionych czynników był wskazywany przez co najmniej kilku ankietowanych jako najważniejszy, a przez inną grupę jako mało ważny. Z kolei każdy z poniższych elementów pojawiał się w zdecydowanej większości ankiet, jako wpływający na ocenę. Oto wskazane elementy oceny urody dowodu:

- Zwięzłość.
- Innowacyjność, zwana przez autora „prawem niespodzianki" – miłe wrażenie sprawia użycie w dowodzie nie znanej nam dotąd metody lub punktu widzenia

na pewne kwestie. Zwłaszcza, jeśli sposób użycia jest zaskakujący i pouczający. Doskonałymi przykładami są przedstawione wcześniej zagadnienia: w „probabilistycznym” rozwiązaniu zadania z Ostrawy „niespodzianka” występuje już na początku, podczas rysowania macierzy, która na pierwszy rzut oka nie ma wiele wspólnego z zadaniem. Ponadto, metoda tu użyta zawiera element dydaktyczny: zwraca uwagę na możliwość użycia rachunku prawdopodobieństwa w problemach nie odwołujących się do niego. Z kolei występująca w „rozwiązaniu z księgi” problemu Buffona igła w kształcie okręgu (konstrukt niewątpliwie zaskakujacy) może być eleganckim przykładem działania matematyki: użycia abstrakcyjnej struktury do rozwiązania konkretnego problemu.
- Przejrzystość głównej idei dowodu – estetykę dowodu zwiększa, gdy możemy docenić jego ideę przewodnią bez wgłębiania się w szczegóły, od pierwszego spojrzenia. Gorzej, jeśli dowód jest w postaci kilku stron obliczeń i, by zrozumieć, po co są wszystkie przejścia, konieczne jest przestudiowanie całości. Nawet jeśli fragmenty takiego dowodu są naprawdę pomysłowe, to trudniej uznać go za piękny. Chyba brak tej przejrzystości jest główną wadą dowodów alternatywnych w przykładach: długie przeliczenia w pierwszym przykładzie skutecznie zacierają ślady myśli przewodniej. W drugim przykładzie obliczenia, co prawda są krótkie, ale np. w dość sztuczny sposób objaśniają, skąd się w wyniku wzięła liczba π (co jest najciekawszym elementem wyniku Buffona).
- Elementarność – dowód, używający prostych środków wielu uważa za ładniejszy niż taki, który jest nawet krótszy, ale wymaga zastosowania potężnych i trudnych w dowodzie twierdzeń.
- Zgodność z tokiem rozumowania twórcy dowodu – dla studentów irytujące bywają dowody, które polegają mniej więcej na tym: wyciągamy z kapelusza jakąś funkcję, różniczkujemy ją 3 razy i łatwo zauważyć, że... Dowód jest prawidłowy, szybki, efektowny, zaskakujący, przejrzysty, ale tak naprawdę nie mamy pojęcia skąd autorowi dowodu przyszła do głowy ta funkcja – co niszczy poczucie użyteczności tej metody dowodowej i tworzy wrażenie, że nie zobaczyło się tego, co w dowodzie było naprawdę ważne.
- Jednolitość/różnorodność użytych metod – punkt dość zaskakujący, ale pojawiający się w wielu wypowiedziach i to właśnie w postaci „spolaryzowanej”. Istnieją osoby, które lubią, gdy znienacka w dowodzie jakiegoś twierdzenia z równań różniczkowych pojawia się lemat z teorii liczb. Inni wolą, gdy twierdzenia przypisane jednej dziedzinie są dowodzone środkami tej dziedzinie przynależnymi. Tak więc w zasadzie ten punkt możnaby zaliczyć do indywidualnych prefererencji, gdyby nie fakt, że pojawia się w sposób istotny w wypowiedziach tak dużej grupy studentów.

Zatem można stwierdzić, że nasza ocena piękna dowodu składa się z odpowiedniego wyważenia powyższych kryteriów oraz z dodatku upodobań indywidualnych. Dla niektórych dowód jest zdyskwalifikowany w konkursie piękności, gdy zawiera słowo „prawdopodobieństwo” (jak w przypadku PZAZS), względnie, gdy liczy się w nim całki, bądź gdy dowód jest niekonstruktywny (np. korzysta z pewnika wyboru). Z drugiej strony są osoby, które czują sentyment do pewnych metod i na przykład uwielbiają dowody korzystające z wybranego twierdzenia (przykłady cenionych ze względu na urodę stosowania twierdzeń z ankiety to twierdzenie Baire'a i lemat Kuratowskiego–Zorna).

Te wszystkie rozważania nie mają zapewne wielkiego związku z naturą matematyki. Ale wydaje mi się, że mają spory związek z naturą matematyków. Warto więc chyba się zastanowić, co tak naprawdę pociąga nas w naszej ulubionej dziedzinie wiedzy.