Paradoks jednomyślności

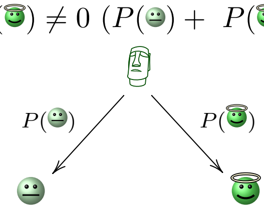
Marcin Makowski
Wydział Fizyki, Uniwersytet w Białymstoku

4 marca 2020

Zgodnie ze starożytnym żydowskim prawem, jeśli oskarżony w wyniku procesu przed Sanhedrynem został jednogłośnie uznany za winnego zarzucanych mu czynów to odstępuje się od wymierzenia kary! Taka zasada dla wielu czytelników jest zapewne szokująca. Jednak ówcześni twórcy prawa uznawali, że jednomyślność sędziów często wskazuje na niedopatrzenie przy ocenie dowodów winy. To mogło wypaczyć wynik procesu. Intuicyjnie rozumiano, że jeśli coś wydaje się zbyt idealne by mogło być prawdziwe to najprawdopodobniej popełniono błąd, nawet jeśli jego natury nie dało się jednoznacznie ustalić. Podobny motyw często występuje w literaturze lub filmie gdy nabieramy wątpliwości wraz z licznie pojawiającymi się dowodami, zbyt jednoznacznie wskazującymi na winę jednego z bohaterów (może ktoś celowo wplątuje go w popełnione przestępstwo?).

Spróbujmy skonstruować prosty probabilistyczny model, w którym da się zaobserwować paradoksalny efekt jednomyślności. To pozwoli nam lepiej zrozumieć jakie są jego przyczyny. Posłużymy się pomysłem zaczerpniętym z pracy: *L. J. Gunn, F. Chapeau-Blondeau, M. McDonnell, B. Davis, A. Allison and D. Abbott, "Too good to be true: when overwhelming evidence fails to convince", Proceedings of the Royal Society A vol. 472, 20150748 (2016).*

Rozważmy oskarżonego 🗿 o popełnienie pewnego przestępstwa. Możemy go scharakteryzować jedną z dwóch wzajemnie wykluczających się cech: 😐 – winny, 😇 – niewinny. Każda z nich występuje odpowiednio z prawdopodobieństwem $P(😐) \neq 0$, $P(😇) \neq 0$ ($P(😐) + P(😇) = 1$).



Załóżmy, że o winie bądź niewinności oskarżonego decyduje 23 sędziów. Właśnie tylu członków liczył każdy z 5 okręgowych Sanhedrynów w starożytnej Judei. Chcemy obliczyć warunkowe prawdopodobieństwo $P(\text{😐}|k)$ że oskarżony jest winny gdy k sędziów ($k = 1, \ldots, 23$) uzgodniło taki werdykt. Przyjmijmy następujące oznaczenia:

- $P(k|\text{😐})$ – prawdopodobieństwo warunkowe, że k sędziów uznało oskarżonego za winnego jeśli był on winny,
- $P(k|\text{😇})$ – prawdopodobieństwo warunkowe, że k sędziów uznało oskarżonego za winnego jeśli był on niewinny,
- $P(k)$ – prawdopodobieństwo uznania przez k sędziów winy oskarżonego.

Wykorzystując twierdzenie o prawdopodobieństwie całkowitym, definicję prawdopodobieństwa warunkowego i twierdzenie Bayesa otrzymujemy:

$$P(k) = P(k|\text{😐}) \cdot P(\text{😐}) + P(k|\text{😇}) \cdot P(\text{😇}),$$

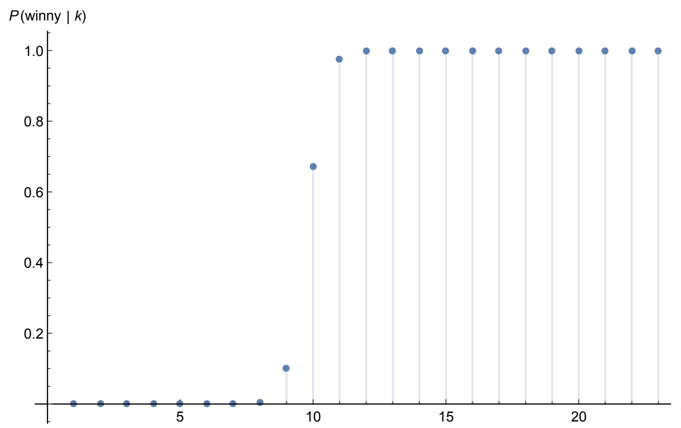
$$P(\text{😐}|k) := \frac{P(\text{😐}\&k)}{P(k)} = P(k|\text{😐}) \cdot \frac{P(\text{😐})}{P(k)}.$$

Symbol $\text{😐}\&k$ oznacza jednoczesne zajście dwóch zdarzeń: oskarżony jest winny i k sędziów uznało jego winę. Zakładając, że każdy sędzia podejmuje decyzję niezależnie i z jednakowym prawdopodobieństwem rozstrzyga o winie oskarżonego możemy przyjąć rozkład dwumianowy jako model prawdopodobieństwa. Mamy

$$P(k|\text{😐}) = \binom{n}{k} r^k \cdot (1-r)^{n-k},\ P(k|\text{😇}) = \binom{n}{k} s^k \cdot (1-s)^{n-k},$$

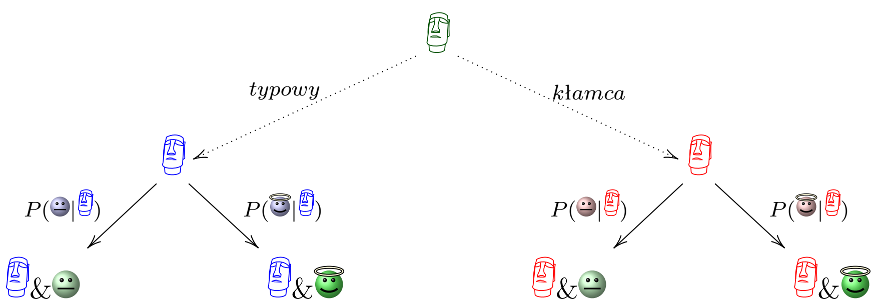
gdzie n oznacza ilość wszystkich sędziów a k to liczba tych, którzy uznali podsądnego za winnego, $s = P(1|\text{😇})$, $r = P(1|\text{😐})$. Rysunek 1 przedstawia jak wartość $P(\text{😐}|k) =: P(\text{winny}|k)$ (dla pewnych ustalonych wartości interesujących nas prawdopodobieństw) zależy od k – ilości sędziów, którzy uzgodnili werdykt uznając oskarżonego za winnego. Wyniki są zgodne z naszą intuicją. Gdy ponad połowa z 23 sędziów uznaje oskarżonego za winnego to prawdopodobieństwo, że jest winny jest bliskie 1.

Skomplikujmy nieco sytuacje. Co jeśli populacja potencjalnych oskarżonych nie jest tak homogeniczna? Być może sędziowie doskonale sobie radzą w typowych sytuacjach, gdy oskarżeni zachowują się zgodnie ze znanym im i dobrze zbadanym schematem. Wprowadźmy do naszego modelu rozróżnienie. Załóżmy, że populacja potencjalnych oskarżonych składa się z osób



Rysunek 1: $n = 23$, $P(1|\textcolor{green}{\text{😇}}) = 0,14$, $P(1|\textcolor{green}{\text{😐}}) = 0,75$, $P(\textcolor{green}{\text{😐}}) = P(\textcolor{green}{\text{😇}})$.

typowych i pewnej domieszki nietypowych, których nazwiemy *kłamcami*. Każdy z nich może być winny albo niewinny a sędziowie nie wiedzą z kim mają do czynienia – Rysunek 2.



Rysunek 2: Rozbicie populacji oskarżonych na dwa rozłączne podzbiory: typowych i kłamców. Prawdopodobieństwo wystąpienia każdego z nich to odpowiednio $P(\textcolor{blue}{\text{🗿}})$ oraz $P(\textcolor{red}{\text{🗿}})$.

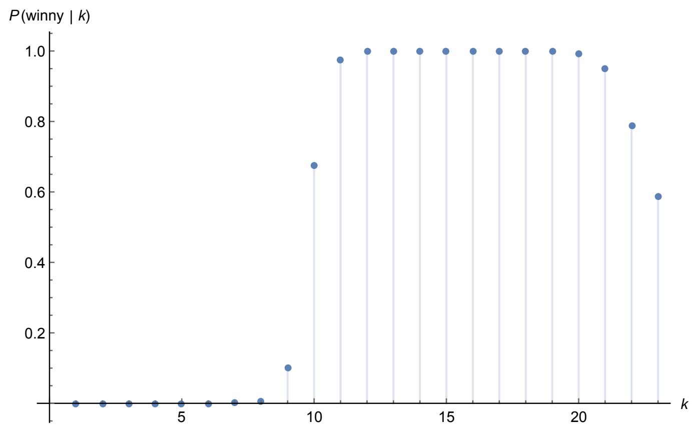
W tym przypadku prawdopodobieństwo, że oskarżony jest winny gdy k sędziów uznało go za winnego przedstawia poniższa formuła:

$$P(\textcolor{green}{\text{😐}}|k) = \frac{P(k|\textcolor{blue}{\text{😐}}) \cdot P(\textcolor{blue}{\text{😐}}|\textcolor{blue}{\text{🗿}}) \cdot P(\textcolor{blue}{\text{🗿}}) + P(k|\textcolor{red}{\text{😐}}) \cdot P(\textcolor{red}{\text{😐}}|\textcolor{red}{\text{🗿}}) \cdot P(\textcolor{red}{\text{🗿}})}{P(k)},$$

gdzie

$$\begin{aligned} P(k) =& P(k|\textcolor{blue}{\text{😐}}) \cdot P(\textcolor{blue}{\text{😐}}|\textcolor{blue}{\text{🗿}}) \cdot P(\textcolor{blue}{\text{🗿}}) + P(k|\textcolor{red}{\text{😐}}) \cdot P(\textcolor{red}{\text{😐}}|\textcolor{red}{\text{🗿}}) \cdot P(\textcolor{red}{\text{🗿}}) + \\ & P(k|\textcolor{blue}{\text{😇}}) \cdot P(\textcolor{blue}{\text{😇}}|\textcolor{blue}{\text{🗿}}) \cdot P(\textcolor{blue}{\text{🗿}}) + P(k|\textcolor{red}{\text{😇}}) \cdot P(\textcolor{red}{\text{😇}}|\textcolor{red}{\text{🗿}}) \cdot P(\textcolor{red}{\text{🗿}}) . \end{aligned}$$

Na rysunku 3 możemy zaobserwować jak interesujące nas prawdopodobieństwo zależy od wartości k. Okazuje się, że gdy ilość sędziów uznających winę oskarżonego jest bliska 23 to prawdopodobieństwo $P(\text{😐}|k) =: P(winny|k)$ spada! Dla jednomyślnej decyzji ($k = 23$) wynosi ok 0,58. To dość zaskakujący efekt.



Rysunek 3: $n = 23$, $P(\text{🗿}) = 10^{-2}$, $P(1|\text{😇}) = 0,14$, $P(1|\text{😐}) = 0,75$, $P(1|\text{😐}) = P(1|\text{😇}) = 0,95$, $P(\text{😐}|\text{🗿}) = P(\text{😇}|\text{🗿}) = P(\text{😐}|\text{🗿}) = P(\text{😇}|\text{🗿})$.

Prawdopodobieństwo tego, że oskarżony jest winny gdy sędziowie byli jednomyślni nie odbiega zbytnio od prawdopodobieństwa wyrzucenia orła w rzucie symetryczną monetą. Przyzwyczajeni jesteśmy raczej do sytuacji gdy decyzje podejmowane są wolą większości a im więcej osób się na nią zgodziło tym lepiej. Nasz przykład pokazuje, że im bardziej jednomyślny staje się rezultat procesu to jego wynik jest mniej wiarygodny. Ma to związek z niepewnością jaką wprowadziliśmy do modelu dodając do analizy nietypowego oskarżonego – *kłamcę*. Niepewność skutkuje nieregularnościami wyników. Wyobraźmy sobie grę w Monopol. Gdyby naszemu przeciwnikowi w każdym rzucie kośćmi wypadałyby dwie szóstki to uznalibyśmy, że coś jest nie tak. Oczywiście nie mądrze byłoby sugerować, że każda jednomyślna decyzja powinna budzić wątpliwości. Na przykład gdy poprosimy kilkanaście osób aby z grupy trzech zwierząt: kota, psa i papugi wskazały kota to najpewniej wszyscy wskażą to samo zwierzę. Ten wybór jest jednomyślny bo nie miały na niego wpływu czynniki losowe. Jednak tak klarowne sytuacje to rzadkość. Zwykle nasze decyzje obarczone są pewnym ryzykiem wynikającym z braku pełnej informacji, fizycznego zmęczenia, doświadczeń życiowych, wywieranej presji i wielu innych czynników. Żyjemy w tak złożonym środowisku i sami jesteśmy na tyle skomplikowani, że gdy nasze wybory staną się jednomyślne to warto się zatrzymać i zastanowić czy to nie jest błąd.