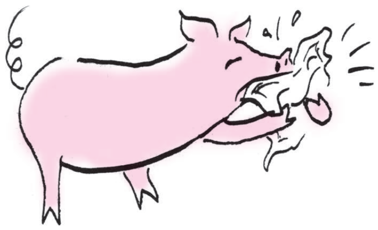
Wzór Bayesa a świńska grypa

*Piotr DWORNICZAK**



Jesień 2009 roku. Na Ukrainie na początku listopada choruje na grypę już 200 000 osób. W kilkunastu przypadkach potwierdzono obecność wirusa A/H1N1, zwanego potocznie wirusem świńskiej grypy. W Polsce narasta niepokój. Potwierdzono obecność tego wirusa u kilku osób. Czy będziemy mieli epidemię? Rząd uspokaja: liczba zachorowań na grypę nie rośnie. Ile z tych zachorowań dotyczy świńskiej grypy, dokładnie nie wiadomo, wykonuje się testy na obecność wirusa. Z danych Państwowego Zakładu Higieny wynika, że do 28 października 2009 roku odnotowano w Polsce 171 przypadków grypy typu A/H1N1. Główny Inspektorat Sanitarny w specjalnym komunikacie nakazuje natychmiastowe zgłaszanie się do lekarza wszystkim powracającym z zagranicy, którzy mają gorączkę, katar, ból gardła czy bóle mięśniowo-stawowe(1). Niepokoić zaczynają się także ci, którzy mają jakieś z tych objawów, mimo że za granicą nie byli. Proszą lekarzy o zlecenia na badanie krwi pod kątem obecności wirusa. Po wywiadzie lekarskim najczęściej okazuje się, że nie ma podstaw do takiego badania. Ale niepokój zostaje. Niektórzy decydują się zrobić badanie na swój koszt.

(1) *Gazeta Wyborcza*, Wrocław 2.11.2009.

Test i prawdopodobieństwo

Metodą, która może być stosowana w celu wykrycia obecności wirusa, jest metoda Real-time PCR (ang. *Polymerase Chain Reaction*)(2). Obecność materiału genetycznego wirusów jest uznawana za oznakę zakażenia. Wielu ekspertów zgodnie uznaje, że PCR to najlepsza metoda umożliwiająca monitoring chorych i mająca ogromną wartość diagnostyczną dla lekarza prowadzącego.

(2) Zob. np. `http://pl.wikipedia.org/wiki/Real_time_PCR` , zob. też hasła PCR, RT-PCR, nested PCR.

Niezależnie od tego, jak dobra jest metoda, mogą wystąpić przypadki błędu wynikającego, na przykład, z jakości materiału badawczego lub błędu operatora. Dla poprawnego działania testu PCR ważna jest koncentracja drobnoustrojów, liczona najczęściej w CFU/ml (ang. *Colony Forming Units*). Oczywiście, im koncentracja jest większa, tym większe jest prawdopodobieństwo, że przeprowadzony test da poprawny wynik. W książce wprowadzającej w tę technikę autorzy podają(3), że badania nad pierwotniakiem *Cryptosporidium parvum* pozwoliły uzyskać prawdopodobieństwo poprawnego wyniku rzędu 95% dla logarytmu dziesiętnego(4) CFU/ml wynoszącego około 4,5. Przy koncentracji dziesięciokrotnie mniejszej prawdopodobieństwo wynosi około 75%, natomiast przy 100-krotnie mniejszej tylko około 30%. Trudno jest przy jednokrotnym stosowaniu metody uzyskać prawdopodobieństwo rzędu 99%, niezależnie od koncentracji. Oczywiście, przy wykonaniu większej liczby niezależnych testów prawdopodobieństwo wykrycia istniejącego w próbce pierwotniaka wzrasta. Pewne badania(5) dotyczące komórek nowotworowych pokazują zgodność testu ze stanem faktycznym poniżej 95%. Dla wirusów ocena CFU lub podobna nie jest przeprowadzana, nie ma bowiem praktycznej możliwości stwierdzenia, jakie jest ich stężenie, gdyż mikroskopy nie zauważają tak małych organizmów, a wirusy nie tworzą kolonii. Gdyby bardzo optymistycznie przyjąć zgodność testu równą 99,9%, oznaczałoby to, że jest bardzo wysoka. Znaczy to też, że średnio test myliłby się raz na 1000 razy lub, innymi słowy, że wyznaczone częstościowo prawdopodobieństwo, iż test da wynik zgodny z rzeczywistością, wynosi 0,999.

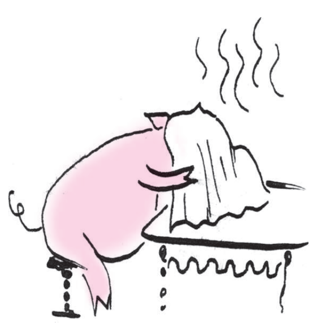
(3) Carl W. Dieffenbach, Gabriela S. Dveksler, *PCR Primer: A Laboratory Manual*, New York 2003, s. 154.

(4) Wielkości CFU/ml są rzędu 10^2–10^8, dlatego też przy określaniu dokładności testu podaje się zazwyczaj logarytm liczby CFU/ml.

(5) S. E. Ten Heuvel, H. J. Hoekstra, A. J. Suurmeijer, *Diagnostic Accuracy of FISH and RT-PCR in* 50 *Routinely Processed Synovial Sarcomas*, Appl. Immunohistochem. Mol. Morphol., 2008, May, 16 (3), 246–250.

Wzór Bayesa

Thomas Bayes, osiemnastowieczny matematyk angielski, zapewne nie przypuszczał, że podany przez niego wzór, nazywany dziś wzorem Bayesa, stosowany będzie powszechnie w statystyce i przyczyni się choćby do rozwoju współczesnego rynku finansowego (wycena instrumentów pochodnych). Wzór ten związany jest z pojęciem prawdopodobieństwa całkowitego, czyli prawdopodobieństwa wystąpienia zdarzenia, które może zajść po uprzednim zajściu pewnych warunków. Formalnie o prawdopodobieństwie całkowitym mówi poniższe twierdzenie.



Jeżeli A jest dowolnym zdarzeniem, $B_1, B_2, \ldots, B_n$ zaś są zdarzeniami

(a) *wykluczającymi się parami, tzn. $B_i \cap B_j = \emptyset$ dla $i \neq j$,*
(b) *o dodatnich prawdopodobieństwach zajścia, tzn. $P(B_i) > 0$ dla $i = 1, 2, \ldots, n$,*
(c) *takimi, że ich suma jest zdarzeniem pewnym, tzn. $B_1 \cup B_2 \cup \ldots \cup B_n = \Omega$,*

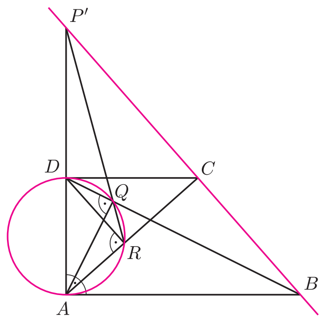
to prawdopodobieństwo (nazywane całkowitym) zajścia zdarzenia A określone jest równością

$$(*) \qquad P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \ldots + P(B_n)P(A|B_n).$$

*Katedra Matematyki Stosowanej, Uniwersytet Ekonomiczny w Poznaniu

Rozwiązanie zadania M 1295.
Oznaczmy przez P' punkt przecięcia prostych AD i QR. Należy dowieść, że $P = P'$.

Rozpatrzmy okrąg o średnicy AD.



Okrąg ten przechodzi przez punkty Q i R. Korzystając zatem z twierdzenia Pascala dla „sześciokąta” $AAQRDD$, wnioskujemy, że punkty B, C oraz P' leżą na jednej prostej. Stąd $P = P'$.

Twierdzenie to w skrótowej formie można wypowiedzieć następująco: jeżeli skutek A może nastąpić po zaistnieniu jednej z jedynie możliwych, wykluczających się przyczyn $B_1, B_2, \ldots, B_n$, to prawdopodobieństwo zajścia skutku A wyraża się wzorem $(*)$.

Z powyższego twierdzenia wynika kolejne, nazywane twierdzeniem Bayesa.

Jeżeli A jest zdarzeniem o dodatnim prawdopodobieństwie, a zdarzenia $B_1, B_2, \ldots, B_n$ spełniają warunki (a), (b) i (c) *poprzedniego twierdzenia, to prawdopodobieństwo warunkowe $P(B_i|A)$ zdarzenia B_i pod warunkiem A określone jest równością*

$$(**)\qquad P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \ldots + P(B_n)P(A|B_n)}.$$

Można to wypowiedzieć następująco: jeżeli skutek A wystąpił po zaistnieniu jednej z jedynie możliwych, wykluczających się przyczyn $B_1, B_2, \ldots, B_n$, to prawdopodobieństwo tego, że przyczyną zajścia skutku A była przyczyna B_i, wyraża się wzorem $(**)$.

Twierdzenie Bayesa określa zatem prawdopodobieństwo przyczyny, gdy wiemy, że nastąpił pewien skutek możliwy do spowodowania przez nią.

Myślenie racjonalne

W połowie XVI wieku w pracach Cardano pojawiło się prawdopodobieństwo zdefiniowane jako proporcja – czyli iloraz – liczby zdarzeń sprzyjających zajściu badanego zdarzenia do liczby wszystkich możliwych zdarzeń elementarnych (tych, które mogą wystąpić w danym doświadczeniu). Pojęcie to rozwinięte zostało w XVII wieku w pracach Bernoulliego, Fermata i Pascala, a później było badane przez Laplace’a. Wtedy też dla oszacowania możliwości wygranych w grach losowych zaczęto wyznaczać prawdopodobieństwa i wielkości nazywane dziś wartościami oczekiwanymi zmiennych losowych. Do czasu pojawienia się w połowie XIX wieku teorii użyteczności Gossena szacowanie możliwości było traktowane jako najbardziej prawidłowy, pod względem naukowym, sposób racjonalnego rozwiązywania problemu wyboru w warunkach ryzyka. Tak pojmowana racjonalność przyjmuje również, że rozpatrując możliwość wystąpienia przyczyny, gdy zaistniał pewien skutek, powinniśmy wykorzystywać wzór Bayesa.

Badania przesiewowe, myślenie codzienne i rachunki

Grypa jest niebezpieczna, a skoro istnieją testy wykrywające wirusa, to czy nie warto zrobić powszechnych badań na jego obecność?

Badanie przesiewowe to w medycynie rodzaj powszechnego badania, które wykonuje się w populacji lub tylko w tzw. grupach wysokiego ryzyka. Celem badania jest wykrycie choroby, ustalenie liczby lub frakcji chorych i umożliwienie zastosowania odpowiedniej terapii. W przypadku pozytywnego wyniku badania przesiewowego choroba musi być potwierdzona innymi metodami diagnostycznymi. Oczywiście, badania są obarczone błędem. Zarzuca się im ponadto „uspokojenie” pacjentów z ujemnym wynikiem testu, mimo że tak naprawdę są chorzy. Testy używane w badaniach powinny wykazywać się wysoką dokładnością[(6)].

[(6)] Przez dokładność lub poprawność testu (ang. *accuracy*) rozumie się zgodność wyniku testu z rzeczywistością.

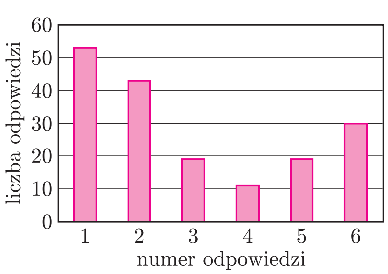
Gdyby wykonać badanie przesiewowe, to jak bardzo należy się przejmować w konkretnym przypadku jego wynikiem?

W listopadzie 2009 r. przeprowadziłem badanie ankietowe 175 studentów uczelni ekonomicznej. Zadałem im pytanie jak poniżej[(7)].

[(7)] Proponuję, aby Czytelnik spróbował odpowiedzieć (bez rachunków!) na to pytanie.

W Polsce coraz więcej mówi się o tzw. świńskiej grypie. Robione są testy potwierdzające obecność wirusa A/H1N1. *Testy te są bardzo dokładne. Pomyłka zachodzi raz na* 1000 *razy. Załóżmy, że w Polsce, która ma* 38 *mln mieszkańców, jest* 200 *osób chorych na świńską grypę* (*dotychczas łączna liczba takich przypadków wynosiła* 171). *Wybrano losowo jedną osobę. Test wykrył obecność wirusa* A/H1N1. *Jak oceniasz prawdopodobieństwo, że ta osoba jest rzeczywiście chora?*
Czy jest to raczej mało prawdopodobne
(1° *poniżej* 0,01, 2° *między* 0,01 *a* 0,1, 3° *między* 0,1 *a* 0,5),
czy raczej bardzo prawdopodobne
(4° *między* 0,5 *a* 0,7, 5° *między* 0,7 *a* 0,9, 6° *powyżej* 0,9)?
Proszę o zaznaczenie jednej odpowiedzi.

Liczba twierdzących odpowiedzi dla poszczególnych punktów była następująca:
1° 53, 2° 43, 3° 19, 4° 11, 5° 19, 6° 30,
co dla zobrazowania wygodnie jest przedstawić na wykresie.



Widać tu preferowanie odpowiedzi skrajnych ($1^\circ + 6^\circ$) – łącznie ponad 47%. Odpowiedź pierwszą wybrało 30% ankietowanych i zbliżoną odpowiedź drugą prawie 25%, co razem daje prawie 55% ankietowanych. Oznacza to, że wśród badanych mniej więcej co druga osoba uznała, że opisane zdarzenie jest mało prawdopodobne. Stosunkowo duży procent odpowiedzi 6° może świadczyć o tym, że ankietowani w odpowiedzi kierowali się jedynie wysoką wiarygodnością testu, lekceważąc informację statystyczną, mówiącą o tym, że niezmiernie małe jest prawdopodobieństwo spotkania osoby chorej. Czy taki rozkład odpowiedzi należy uznać za dziwny? Raczej nie. Wyniki eksperymentu zgadzają się z badaniami[8] Tversky'ego oraz Tversky'ego i Kahnemana, na podstawie których badacze doszli do wniosku, że ludzie zazwyczaj nie kierują się w wyborach regułami wynikającymi z probabilistycznej natury zjawiska. O ile bez obliczeń jest w oczywisty sposób dość trudno podać w miarę dokładną liczbową ocenę prawdopodobieństwa, o tyle stwierdzenie, czy zdarzenie jest raczej mało prawdopodobne, czy raczej bardzo prawdopodobne, powinno być dokonane w przeważającej większości w sposób właściwy. W przeprowadzonym badaniu jest tak w stosunku 115:60. Oznacza to, że prawie co trzecia osoba diametralnie myliła się w ocenie – zaraz zobaczymy, jak bardzo[9].

[8] Zob. np. D. Kahneman, A. Tversky, *Subjective probability: A judgment of representativeness*, Cognitive Psychology 3 (1973), ss. 430–454; A. Tversky, *Features of similarity*, Psychological Review 84 (1977), ss. 327–352; cytowane w bardzo dobrej książce T. Tyszki (red.), *Psychologia ekonomiczna*, GWP, Gdańsk 2004.

[9] Zauważmy, że osoby ankietowane są studentami kierunków ekonomicznych. Uważam, że gdyby poddać eksperymentowi losowo wybraną grupę osób z Polski, wyniki świadczyłyby o jeszcze gorszym szacowaniu prawdopodobieństw.

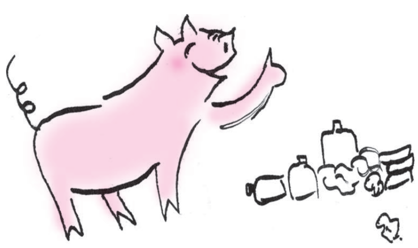
Spróbujmy zatem pomyśleć o badaniach przesiewowych w kontekście rachunku prawdopodobieństwa, przechodząc do „racjonalnego” myślenia i wzoru Bayesa.

Załóżmy, że Polska ma 38 000 000 mieszkańców, z których 200 jest chorych na świńską grypę. Gdybyśmy wybrali losowo jedną osobę, to prawdopodobieństwo tego, że jest ona chora, wynosi 200/38000000, czyli w przybliżeniu 0,0000053. Oczywiście, prawdopodobieństwo tego, że nie jest chora, wynosi $1 - 0{,}0000053$, czyli 0,9999947. Przyjmijmy, że test wykazał obecność wirusa. Zastanówmy się, jakie jest prawdopodobieństwo tego, że pacjent, u którego stwierdzono obecność wirusa, jest rzeczywiście chory.

Skutkiem w naszym doświadczeniu jest stwierdzenie, że pacjent jest chory. Pytamy o prawdopodobieństwo, że istotnie tak jest. Oznaczając przez A zdarzenie, że u pacjenta stwierdzono obecność wirusa A/H1N1, natomiast przez B_1 i B_2 odpowiednio zdarzenia, że pacjent jest chory, oraz że chory nie jest, zgodnie ze wzorem $(**)$ otrzymujemy

$$P(B_1|A) = \frac{0{,}999 \cdot 0{,}0000053}{0{,}999 \cdot 0{,}0000053 + 0{,}001 \cdot 0{,}9999947} \approx 0{,}005.$$

Powyższy wynik oznacza, że przy losowo wybranej osobie prawdopodobieństwo zdarzenia, że ma ona istotnie świńską grypę, mimo iż test na obecność wirusa dał rezultat pozytywny, jest małe. Dzieje się tak mimo wielkiej dokładności testu! Przypomnijmy, że w badaniu ankietowym odpowiedź tę (tzn. około 0,005) wybrało 30% ankietowanych, a zbliżoną odpowiedź drugą prawie 25% pytanych. Oznacza to, że wśród badanych mniej więcej co druga osoba podała wynik zbliżony do obliczonego ze wzoru Bayesa.



A już poza ankietą – jak byłoby, gdyby wynik testu był negatywny? Oznaczając przez A' zdarzenie przeciwne do A (tzn. zdarzenie, że wynik testu jest negatywny), można obliczyć prawdopodobieństwo zdarzenia, że osoba jest chora, mimo że wynik testu jest negatywny. Otrzymujemy wtedy

$$P(B_1|A') = \frac{0{,}001 \cdot 0{,}0000053}{0{,}001 \cdot 0{,}0000053 + 0{,}999 \cdot 0{,}9999947} \approx 0{,}000000005.$$

Oznacza to, że jeśli wynik testu jest negatywny, to osoba testowana prawie na pewno rzeczywiście nie ma świńskiej grypy.

Skąd wobec tego takie odpowiedzi ankietowanych[10]? Dlaczego nie myślimy kategoriami prawdopodobieństwa? Ulegamy panice? Może trochę. Dlaczego tak jest? Być może dlatego, że lepiej spanikować i uciekać nawet przed mało prawdopodobnym nieszczęściem, niż skazać się na prawie niemożliwą, ale bardzo niekorzystną sytuację. Co ma robić społeczeństwo? Opinia publiczna wymusza czasem na politykach pewne decyzje. Bywa, że decyzji tych nie można zaliczyć do racjonalnych w sensie probabilistycznym. Dotychczas nie notuje się w mediach propozycji badań przesiewowych populacji. I bardzo dobrze, gdyż z rachunków wynika, że obarczone są wielkim błędem. Mimo dużej dokładności testu jego wynik, jakikolwiek by nie był, niesie ze sobą mało informacji o przyczynie. A koszty są niemałe.

[10] A Twoja odpowiedź, Czytelniku?

Co robić, aby uniknąć groźnej choroby? Stosować się do takich samych zaleceń jak w przypadku innych chorób zakaźnych – unikać kontaktu z chorymi, stosować podstawowe zasady higieny i cieszyć się, że wirusy grypy dość szybko giną poza organizmem nosiciela lub chorego[11], a my mamy przed nimi naturalną obronę.

[11] Od kilku godzin do około 2 dni w zależności od temperatury, wilgotności i podłoża.