

Twierdzenie Bayesa

Esej ten będzie poświęcony pewnemu aspektowi zastosowania probabilistyki do zagadnień wyboru religii, mianowicie, do robiącemu (niezasłużenie) furorę, zwłaszcza za Oceanem, zastosowania tzw. wnioskowania Bayesowskiego w celu wyboru, która opcja religijna jest „bardziej prawdopodobna”. W szczególności dotyczy się to zagadnień typu czy Jezus istniał/ czy był Bogiem, czy zmartwychwstał itp.

Żeby wykazać jałowość takiego podejścia, musimy najpierw zacząć od podstawowych pojęć, mianowicie, co to tak naprawdę jest prawdopodobieństwo, czym jest tzw. twierdzenie Bayesa, jak się je stosuje, itp.

Przed wszystkim należy sobie zdać sprawę, czym w ogóle jest prawdopodobieństwo. Istnieje kilka definicji prawdopodobieństwa, najczęściej jest jednak przyjmowana definicja Kołmogorowa, jako najpełniejsza. Mówi ona tak:

- Istnieje pewien niepusty zbiór, nazwijmy go sobie

Czytaj więcej na:

<http://lewandowski.apologetyka.info/ateizm/ateistyczny-pseudoargument-z-osobistego-niedowierzenia.968.htm>

Czytaj więcej na:

<http://lewandowski.apologetyka.info/ateizm/ateistyczny-pseudoargument-z-osobistego-niedowierzenia.968.htm>

, zwany przestrzenią probabilistyczną

- Podzbiory tej przestrzeni A, nazywamy zdarzeniami losowymi
- Funkcja $P(A)$ odwzorowująca zdarzenie A na przedział liczb od 0 do 1 nazywamy prawdopodobieństwem. Ma ono wymienione poniżej własności.
- $P(\Omega)=1$ -prawdopodobieństwo przypisane całej przestrzeni probabilistycznej (a więc zdarzeniu pewnemu) wynosi 1
- Jeśli zdarzenia A i B są zbiorami rozłącznymi (bez wspólnej części) to $P(A+B)=P(A)+P(B)$, prawdopodobieństwo sumy zdarzeń rozłącznych jest równe sumie ich prawdopodobieństw.

Czyli po prostu pewna **abstrakcyjna definicja** za pomocą matematycznych pojęć mająca możliwie najogólniej opisać, co **intuicyjnie rozumiemy** przez prawdopodobieństwo. Definicja, którą ktoś kiedyś wymyślił, i która opisuje wiele, ale **nie da się za jej pomocą opisać wszystkiego**. Nie jest więc to żadna fundamentalna właściwość bytu.

Ale za jej pomocą da się opisać szereg twierdzeń dotyczących prawdopodobieństwa. Jednym z takich twierdzeń jest **twierdzenie Bayesa**.

Oznaczmy $P(A|B)=P(AB)/P(B)$ jako **prawdopodobieństwo warunkowe zdarzenia A pod warunkiem, że zaszło zdarzenie B**. Przekształcając formułę mamy $P(AB)=P(A|B)P(B)$ -prawdopodobieństwo że zaszło jednocześnie zdarzenie A i B wynosi prawdopodobieństwo A pod warunkiem że B, razy prawdopodobieństwo zdarzenia B. Zapisując zdarzenia A i B w odwrotnej kolejności mamy trywialną zależność:

$$P(A|B)P(B)=P(B|A)P(A)$$

Jest to właśnie **twierdzenie Bayesa**. Zależność jest trywialna, i myślę że intuicyjnie zrozumiała, prawdopodobieństwo złożenia dwóch zdarzeń jest takie same niezależnie które ze zdarzeń uważamy za warunek.

Twierdzenie Bayesa

Jak to się jednak stosuje w praktyce? Przepiszmy to twierdzenie jeszcze raz, zmieniając nieco oznaczenia:

$$P(M|d)P(d)=P(d|M)P(M)$$

Zakładamy, że mamy jakiś zbiór danych \mathbf{d} , o prawdopodobieństwie opisywanym przez $P(\mathbf{d})$ -prawdopodobieństwo to w anglojęzycznej terminologii nazywa się *evidence*. Mamy też pewien zbiór modeli M , które prognozują wartości tychże danych, prawdopodobieństwo przypisywane modelom $P(M)$ nazywa się prior. Podobnie $P(M|d)$ nazywa się *posterior* zaś $P(d|M)$ *likelihood*.

Cała zabawa polega na tym, żeby wybrać taki model, przy którym prawdopodobieństwo uzyskania danych, jakie uzyskujemy w doświadczeniu, było jak największe. Innymi słowy warunkujemy taki model M , by $P(d|M)$, *likelihood*, było jak największe. Nazywa się to w literaturze anglojęzycznej *maximum likelihood approach*, zaś w literaturze polskiej *metodą największej wiarygodności*.

Likelihood można łatwo obliczyć przekształcając twierdzenie Bayesa:

$$P(d|M)=P(M|d)P(d)/P(M)$$

przy czym ponieważ d oznacza tu zbiór wszystkich danych, zatem trzeba czynnik $P(M|d)P(d)$ wysumować po wszystkich danych. Trzeba jeszcze jakoś oszacować prior. Zwykle, gdy nie ma żadnej informacji sugerującej, który model jest lepszy, zakłada się po prostu losowe prawdopodobieństwo dla wszystkich.

Ten model, który ma największy *likelihood* (czyli dane najbardziej do niego pasują), wygrywa.

Statystyka Bayesowska zdobyła sobie niezwykłą popularność w krajach anglosaskich. Moja prywatna opinia jest taka, że może to mieć związek, z podobieństwem do ichnich systemów wyborczych: zwycięzca bierze wszystko. Nawet jak głosowała na niego mniejszość. Bo statystyka Bayesowska ma także swoje słabe punkty. Jednym z nich jest właśnie to, że *likelihood* zwycięskiego modelu, pomimo iż największe ze wszystkich, **wcale nie musi być duże**. Co oznacza, paradoksalnie, że bardziej prawdopodobne jest to, że ten model nie jest poprawny, poprawny jest któryś z innych modeli, choć indywidualnie każdy z nich przegrywa ze zwycięzcą.

Drugim zastrzeżeniem, które należy poczynić w stosunku do statystyki Bayesowskiej, jest to że niejako wymusza ona wybór któregoś z modeli, które, jak zakładamy, muszą wyczerpywać wszystkie możliwości. Któryś model, musi być prawdziwy. A jest to założenie, które wcale nie musi odzwierciedlać rzeczywistość -raczej to jak zdajemy się ją postrzegać. Głosujesz na Republikanów czy Demokratów? Na Konserwatystów, Liberalów, czy Partię Pracy?

Trzecie zastrzeżenie odnosi się właśnie do tego, w jaki sposób postrzegamy tę rzeczywistość. A postrzegamy ją w świecie Bayesowskim przez owe dane. Które możemy dobrać jak nam się tylko podoba, jakie tylko uznamy za istotne. Nie istnieją jednoznaczne kryteria probabilistyczne mówiące, jaki dobór danych (w sensie jakie wielkości uznajemy za istotne) musimy wybrać, podobnie jak nie istnieją kryteria odnośnie doboru modeli -poza warunkiem, że te modele muszą dawać nam przewidywania odnośnie danych.

Co pozwala manipulować tym na wiele sposobów. Rozważmy następujące przykłady.

Mamy trzy modele:

A) Jeśli poprawną religią jest buddyzm, to wysokość Mount Everest dana jest przez rozkład normalny wycentrowany na wysokości 9 km, z odchyleniem standardowym 1 km.

Twierdzenie Bayesa

B) Jeśli jest nią hinduizm, to wysokość Mount Everest jest dana przez rozkład normalny o odchyleniu 1 km, wycentrowany na wysokości 10 km

C) Jeśli jest nią mormonizm, to wysokość Mount Everest jest opisana przez rozkład normalny, centrum na wysokości 8 km, odchylenie 1 km (wszystkie trzy odchylenia są takie same).

Dane: wysokość Mount Everest 8848 m, zatem największy likelihood ma opcja A), buddyzm, na drugim miejscu są mormoni.

Drugi przykład: trzy modele:

A) Jeśli poprawną religią jest katolicyzm, to wysokość Rysów dana jest przez rozkład normalny wycentrowany na wysokości 1999 m, z odchyleniem standardowym 1000 m.

B) Jeśli poprawną religią jest prawosławie, to wysokość Rysów dana jest przez rozkład normalny wycentrowany na wysokości 2999 m, z odchyleniem standardowym 1000 m.

A) Jeśli poprawną religią jest islam, to wysokość Rysów dana jest przez rozkład normalny wycentrowany na wysokości 999 m, z odchyleniem standardowym 1000 m.

Dane: wysokość Rysów 2499 m.

W tym przypadku mamy remis między katolicyzmem a prawosławiem -odchylenie od modelu jest w obu przypadkach takie same.

Oczywiście oba te przykłady są zupełnie bezsensowne. Jednak **z matematycznego punktu widzenia, są one sformułowane prawidłowo**. Co pokazuje, że wszystko tak naprawdę zależy od sformułowania problemu. Oba przykłady używają innego doboru danych i innego doboru modeli, co pokazuje, że bezpośrednie porównanie między nimi jest tak naprawdę niemożliwe.

W praktyce wszelkie dysputy prowadzone za Oceanem za pomocą narzędzi bayesowskich kończą się na tym, że jeśli wnioskowanie takie przeprowadza wojujący ateista, to prawdopodobieństwo, że np. Jezus nie istniał wychodzi mu bliskie 1, a jak chrześcijański apologeta, to prawdopodobieństwo że Jezus istniał, zmartwychwstał i był Bogiem, też mu wychodzi bliskie 1.

Statystyka Bayesowska sprawdza się bardzo dobrze w sytuacji, gdy mamy dobrze określony problem, taki jak np. gdy mamy pewien gatunek ptaszków, w których występują dwie maści upierzenia i dwa kształty dzioba, i jakie jest prawdopodobieństwo że ptaszek o danym upierzeniu, ma dany kształt dzioba. Gdy problem jest otwarty, można sformułować go na różne nierównoważne sposoby, każdy ma inne do niego podejście, statystyka Bayesowska ma swoje ograniczenia. Jest ona niewątpliwie bardzo istotnym narzędziem -podstawowym, jeśli chodzi np. o klasyfikację, czy redukcje olbrzymich ilości danych doświadczalnych dostarczanych przez dzisiejsze instrumenty badawcze (akceleratory, teleskopy, kamery itp.). Ale musi być stosowana z głową. Gdy tak nie jest, to jest to prosta droga do katastrofy -nadużywanie instrumentów bayesowskich stawało się nieraz przyczyną baniek giełdowych i kryzysów finansowych.

Styczeń 2017

Author: O.K.

<http://ok.apologetyka.info/article/970/>