Historia Rachunku Prawdopodobieństwa i Statystyki
WYKŁAD IV: Simpson, Bayes, Laplace

MiNI PW

Problem, którego nie potrafił rozwiązać de Moivre

CTG de Moivre'a nie dało mu rozwiązania problemu praktycznego:

wiedząc np. (dane z tablicy Halley'a), że z 346 mężczyzn w wieku 50 lat tylko 142 dożywa siedemdziesięciu, jaka jest wiarogodność frakcji $\hat{p} = 142/346$ jako oceny prawdopodobieństwa $\theta = P(X \geqslant 70|X \geqslant 50)$? jaka jest szansa, że to prawdopodobieństwo wynosi $1/2$?

Był to problem policzenia *odwrotny* do tego, który umiał rozwiązać de Moivre: wiemy, że $p = 1/2$. Jaka jest szansa, że obserwowana frakcja będzie blisko p ? Z tego względu taki problem nazywano problemem odwrotnego prawdopodobieństwa (*inverse probability*).
We współczesnym języku: problem obliczenia rozkładu aposteriori

$$\theta|A,$$

gdzie A= $\{$ 142 osoby z 346 dożyły 70 lat $\}$ i o θ zakładamy, że jego rozkład apriori jest jednostajny na (0,1) .

De Moivre policzył jeszcze

$$E|(X/n) - p| \approx \sqrt{2pq/n\pi}$$

ale to również nie pomagało rozwiązać problemu odwrotnego prawdopodobieństwa

Problem modelowania błędów pomiaru

Drugi problem z którym borykali się astronomowie:
Rozkład krzywej błędów w modelu obserwacji astronomicznych:

$$O = P + \varepsilon,$$

gdzie O - punkt obserwowany, P - punkt, który chcemy obserwować, ε - błąd obserwacji.
Kto dokonał przełomu ?
Simpson, Bayes, Laplace ??

Tomasz Simpson

Doctrine of Annuities and Reversions (1755)

Tomasz Simpson 1710-1761

TS był tkaczem i samoukiem matematycznym, w wieku 19 lat ożenił się, w 1735 przeniósł się z rodziną do Londynu i zarabiał na życie tkactwem w ciągu dnia i uczeniem matematyki wieczorami.
Napisał *Doctrine of Annuities and Reversions*: uproszczona i poprawiona wersja *Annuities upon Lives* de Moivre'a.
Wykorzystywał w obliczeniach tablice śmiertelności Graunta, a nie tablice śmiertelności Halleya.
de Moivre poczuł się zagrożony publikacjami Simpsona (odbierały mu źródło zysku ...)

Reakcja de Moivre'a na publikację *Doctrine of Annuities and Reversions* (z przedmowy drugiego wydania *Annuities upon Lives*)

After the pains I have taken to perfect this Second Edition, it may happen, that a certain Person, whom I need not name, out of *Compassion to the Public*, will publish a Second Edition of his Book on the same Subject, which he will afford at a *very moderate Price*, not regarding whether he mutilates my Propositions, obscures what is clear, makes a Shew of new Rules, and works by mine; in short, confounds, in his usual way, every thing with a croud of useless Symbols; if this be the Case, I must forgive the indigent Author, and his disappointed Bookseller. (De Moivre, 1743, p. xii)

Pod koniec życia Simpson został nauczycielem matematyki w Królewskiej Akademii Wojskowej w Woolwich Simpson sformułował model

$$y_i = m + \varepsilon_i,$$

gdzie ε_i są błędami o średniej 0 i oszacowanie

$$P(|\bar{\varepsilon}| \geqslant \eta) = P(|\bar{y} - m| \geqslant \eta)$$

Krzywa błędów

Simpson rozpatrzył bład mający rozkład na $\{-v, -v+1, \ldots, 0, 1, \ldots, , v\}$ postaci

$$C \times (r^{-v}, 2r^{-v+1} \ldots, (v+1)r^0, vr^1 \ldots, r^v\}$$

(dla $r = 1$ rozkład $C \times (1, 2, \ldots, v+1, v, \ldots, 1)$), a następnie policzył prawdopodobieństwo

$$R_n(k) = \frac{P(|\bar{X}_n| > k)}{P(|X_1| > k)}$$

dla $k = 1, 2$ i $r = 1, v = 5, n = 6$.
$k = 1$ $R_6(k) = 0.49$, $k = 2$ $R_6(k) = 0.099$

Konkluzja Simpsona:
Upon the whole .. it appears that the taking of the Mean of a number of observations, greatly diminishes the chances for all smaller errors, and cuts off almost all possibility of any great ones: which last consideration, alone, seems sufficient to recommend the use of the method, not only to astronomers but to all others concerned in making of experiments of any kind (to which the above reasoning is equally applicable). And the more observations are made, the less the conclusion be liable to err, provided they admit of being repeated under the same circumstances.

Tomasz Bayes (1702 - 1761)

Tomasz Bayes był pastorem anglikańskim; należał do non-konformistów (odłam kościoła anglikańskiego, przeciwny obrzędowości mszy). Tomasz Bayes nie opublikował nic za życia, swoją podstawowę pracę zostawił w testamencie T. Price'owi z prośbą o opublikowanie.
Została przeczytana przez Price'a na spotkaniu Royal Society w dwa lata po jego śmierci i opublikowana w Transactions of RS w 1764.

Model Bayesa

Podstawowe twierdzenie i model:
Proposition 3: zdarzenie E_2 zaszło po E_1. Wtedy
$P(E_2|E_1) = P(E_1 \cap E_2)/P(E_1)$.
Model: kwadrat ABCD, na który rzucono najpierw jedną kulę (jej wsp. x - parametr θ, później wykonano n rzutów drugą kulą i zarejestrowano

$$X - \mathrm{liczba\,rzutów\,o\,wsp.\,x} \leqslant \theta.$$

Bayesa interesowało prawdopodobieństwo

$$P(L < \theta < U|X)$$

W naszym języku: $X \sim Bin(n, \theta)$, ale $\theta \sim U[0,1]$.

Kluczowe założenie: $\theta \sim U[0,1]$.
Bayes argumentował, że jeśli wnioskujemy o prawdopodobieństwie θ, o ktorym nic nie wiemy to powinniśmy założyć, ze rozkład tego prawdopodobieństwa jest jednostajny.
Edgeworth: gdybyśmy wzięli pewną monotoniczną funkcję $f(\theta)$ i załozyli rozkład jednostajny dla niej, to rozkład θ nie będzie już z jednostajny ...

'Stół bilardowy' Bayesa

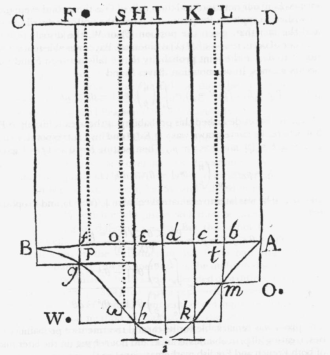

Figure 3.4. Bayes's "Billiard Table": a square table ABCD and two balls W and O. The ball W is rolled and comes to a rest on the line os, then the ball O is rolled n times and the number of times it comes to a rest in the rectangle osDA is recorded. The lower curve is meant to show the posterior density of the fraction Ao/AB, given the recorded count. Bayes's upper and lower approximations to this density are also shown. (From Bayes, 1764, p. 385.)

Bayes policzył

$$P(L < \theta < U)|X = p) = \frac{\int_L^U \binom{n}{p}\theta^p(1-\theta)^{n-p}\,d\theta}{\int_0^1 \binom{n}{p}\theta^p(1-\theta)^{n-p}\,d\theta} =: \frac{L}{M}$$

i wiedział, że $M = (n+1)^{-1}$.
(gęstość rozkładu beta $B(p+1, n-p+1) : \frac{(n+1)!}{p!(n-p)!}\theta^p(1-\theta)^{n-p}$).
Nie umiał policzyć

$$f(s) = \int_0^s \theta^p(1-\theta)^{n-p}\,d\theta$$

co (wiedział to !) rozwiązywałoby jego problem (niekompletna funkcja beta)

Prosecutor's fallacy

Twierdzenie Bayesa i prawdopodobieństwo warunkowe jest często źle interpretowane i zrozumiane. Typowy przykład to tzw. błędne przekonanie oskarżyciela: **Prosecutor's fallacy**.
E - dowody (evidence) I - niewinny (innocent)
Powinniśmy wnioskować:
$P(I|E)$ -małe $\rightarrow$: oskarżony może być winny.
Prosecutor's fallacy: Zastąpienie $P(I|E)$ poprzez $P(E|I)$. Mamy

$$P(I|E) = \frac{P(E|I)P(I)}{P(E)}.$$

Jeśli $P(I)/P(E) >> 1$ to $P(I|E) >> P(E|I)$.

Sprawa Sally Clark

Jeden najbardziej znanych (i tragicznych) przykładów: ekspertyza R. Meadowa na procesie Sally Clark (1999) dotycząca dwóch przypadków nagłego zgonu niemowląt SIS. Medows sformułował tzw **syndrom M$\ddot{u}$nhausena przez przeniesienie**: opiekunka/matka pozoruje chorobę dziecka, aby zwrócić na siebie uwagę. Dlatego:
unless proven otherwise one cot death is tragic, two is suspicious and three is a murder.
Sally Clark straciła dwóch synów (w wieku 11 i 8 tygodni). Meadow ocenił

$$P(I|E) \approx P(E|I) = (1/8500)^2 = 1/(73 \times 10^6)$$

na podstawie tego orzeczenia została skazana, wypuszczona z więzienia dopiero po 3 latach.

Inny przykład błednego rozumowania sądowego

Próbka DNA z miejsca przestępstwa jest porównanana z bazą danych zawierającą DNA 20 000 osób.

$$P(\text{dwie losowe próby są takie same}) = 1/10000.$$

Znaleziono zgodność z osobą X w bazie.
Czy to oznacza, że

$$P(\text{X niewinny}) = 1/10000$$

Oczywiście nie ! Prawdopodobieństwo, że w bazie znajdzemy przynajmniej jedną zgodność

$$1 - (1 - \frac{1}{10000})^{20000} = 86\%$$

Problem wielokrotnego testowania

Pierre Simon Laplace (1749-1827)

Pierre Simon Laplace (1749–1827), as portrayed in 1799

Pierre Simon Laplace pochodził z Normandii.
Jego ojciec był handlarzem cydru.
Studiował teologię na unwersytecie w Caen, wyjechał nie uzyskawszy stopnia z listem polecającym do d'Alamberta.
Profesor Akademii Wojskowej i Dyrektor Urzędu Miar i Wag.
W 1799 był krótko (6 tygodni) ministrem spraw wewnętrznych u Napoleona.
(Laplace "carried the spirit of the infinitesimal into administration.")
astronom, geodeta, matematyk (analiza, rachunek prawdopodobieństwa).

Osiągnięcia Laplace'a

- To co nazywa się twierdzeniem Bayesa:

$$A_1, \ldots, A_n - \text{partycja} : P(A_i|E) = \frac{P(E|A_i)P(A_i)}{\sum_j P(E|A_j)P(A_j)}.$$

- Centralne Twierdzenie Graniczne.
- Twierdzenie o zgodności rozkładu aposteriori $\theta \sim U[0,1]$-prawdopodobieństwo sukcesu w rozkładzie dwumianowym. Zaobserwowano p-sukcesów, q-porażek. Jak zachowuje się

$$P(|\theta - \frac{p}{p+q}| \leqslant w|p,q) \to 0 \quad p.n.p.$$

dla dowolnego $w > 0$.

Pokazał mianowicie, że

$$P(|\theta - \frac{p}{p+q}| \leqslant w|p,q) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-w}^{w} exp(-z^2/(2\sigma^2))\, dz,$$

gdzie $\sigma^2 = \frac{pq}{(p+q)^3} \to 0$ p.n.p.

Zgodność względem rozkładu apriori

An ingenious Friend has communicated to me a Solution to the inverse Problem, in which he has shewn what the Expectation [that is, probability] is, when an Event has happened p times, and failed q times, that the original Ratio of the Causes for the Happening or Failing of an Event should deviate in any given Degree from that of p to q. And it appears from this Solution, that where the Number of Trials is very great, the Deviation must be inconsiderable. (Hartley, 1749, vol. 1, p. 339; Stigler, 1983)