DYSKUSJE I POLEMIKI

Archeologia Polski, t. L:2005, z. 1-2
PL ISSN 0003-8180

ADAM WALANUS

ARCHEOLOGIA AWANGARDĄ NAUK ŚCISŁYCH. O FORMIE KALIBROWANYCH DAT ^{14}C

Pojęcie nauki ścisłej nie jest jednoznaczne, zapewne obejmuje ono również matematykę, na pewno nauki techniczne i przyrodnicze, z których, przede wszystkim fizykę. Wyróżnikiem ścisłości mogłoby być stosowanie metody ilościowej, liczbowej, co jednak nie rozstrzygałoby chyba przynależności matematyki. Wydaje się, że zupełnie jednoznaczne będzie wskazanie na pomiar jako narzędzie, którego istotne stosowanie nadaje dziedzinie aktywności poznawczej pewną cechę ścisłości. Z nauk humanistycznych, w socjologii wykonuje się badania, które nazwać można pomiarami, a na pewno ich wyniki analizuje się stosunkowo zaawansowanymi metodami. Humaniści korzystają z pomiarów, lecz wykonywanych zwykle przez przyrodników.

Pomiar jest czynnością poznawczą zawierającą koniecznie dwa elementy: porównywanie i zliczanie. Przykładając linijkę do naczynia porównujemy położenie kresek milimetrowych z pozycją krawędzi naczynia, po czym "liczymy" kreski milimetrowe. Mierząc liczbę jabłek w koszyku porównujemy kolejne sztuki z wzorcem jabłka i zliczamy. Woltomierz cyfrowy i inne zaawansowane technicznie mierniki również działają podobnie. Nieco bliższy wgląd w pomiar ujawnia jednak pewien kłopot. W jaki sposób w pomiarze otrzymuje się liczbę (w zasadzie naturalną) nie podlega kwestii, problem jest jednak z poznawczym statusem otrzymanej liczby - wyniku pomiaru. Otóż wynik pomiaru (w postaci liczby [z wymiarem]) jest wątpliwym obrazem prawdy. Z bardzo wielu powodów nie całkiem wierzymy, że naczynie ma 345 mm wysokości, jabłek jest 34, a napięcie wynosi 543 mV. Chodzi tu o możliwość błędów, mniejszych, większych, a nawet tzw. grubych.

Nie jest oczywiste czy tzw. ocena niepewności wyniku pomiaru jest elementem pomiaru, czy też nie dałoby się w niektórych wypadkach doszukać w procesie wyznaczania wielkości błędu elementów cechujących właśnie pomiar. Na pewno z sensem można mówić o błędzie błędu. Ze względu na charakter pomiarów, które są tu zasadniczym tematem, warto, być może, przytoczyć przykład dość ścisłego wyznaczenia wielkości niepewności pomiarowej. Niech pomiarowi podlega liczba atomów ^{14}C w drewnianym szczątku (metoda ^{14}C, patrz np. A. Walanus, T. Goslar 2004). Atomy liczy się jednak nie w całym obiekcie, a w małym jego kawałku. Atomów zliczono, powiedzmy 1000. Ponieważ atomy ^{14}C rozłożone są losowo pomiędzy „zwykłymi” atomami ^{12}C, to stwierdzić można, że w określonej części jest ich zapewne 1000, ale z błędem równym pierwiastkowi z 1000, czyli wynoszącym około 32 (wynik ma postać: 1000±32 sztuk). Podstawą obliczeń jest tu tzw. rozkład prawdopodobieństwa Poissona (patrz np. P.J. Durka 2003.). Podstawą wnioskowania najczęściej jest twierdzenie Bayesa (T.R. Bayes 1763; C.E. Buck, J.B. Kenworthy, C.D. Litton, A.F.M. Smith 1991; C.E. Buck, W.G. Cavanagh, C.D. Litton 1996; D. J. Michczyńska, M. F. Pazdur, A. Walanus 1989; A.

Walanus 1983; 2000). Nie wchodząc w zawsze budzące wątpliwości zagadnienie rozkładu *a priori* (przed pomiarem), można przyjąć, że wynikiem pomiaru jest rozkład prawdopodobieństwa (Poissona, z parametrem 1000). W pomiarze wysokości naczynia czy napięcia nieco inne jest źródło błędu, a właściwie źródła, których wielość jest istotna. Tu wynik opisywany jest rozkładem prawdopodobieństwa normalnym, który tym głównie różni się od rozkładu Poissona, że jego „szerokość" (odchylenie standardowe, sigma) nie jest związana z wartością oczekiwaną (środkową). Poza tym, w typowych sytuacjach rozkład Poissona można traktować jak rozkład normalny.

W naukach ścisłych wynikiem pomiaru jest normalny rozkład prawdopodobieństwa; krzywa dzwonowa o szerokości określonej mniej lub bardziej dokładnie. Warto zauważyć, że błąd pomiarowy (sigma), będący wyznacznikiem jakości, a więc i wartości pomiaru, podlega, niestety, zwykłemu prawu rynkowemu i jest często zaniżany. Tak więc w technice i w naukach przyrodniczych wiadomo, że zapis 1000±32 oznacza, że z prawdopodobieństwem 0,68 rzeczywista wartość wielkości mierzonej mieści się między 1000-32, a 1000+32, natomiast, dla osiągnięcia większej pewności, na poziomie 0,95, rozszerzyć trzeba przedział do ±2 sigma, itd. Prawdopodobieństwa równego jeden nie osiąga się nigdy, bo osiąga się je w nieskończoności.

Archeolodzy są w znacznie trudniejszej sytuacji (A. Walanus 2004). Banalna krzywa dzwonowa odeszła w przeszłość w momencie, gdy fizycy udoskonalili metodę ^{14}C pomiaru wieku. Źródłem niepewności pomiarowej pozostaje Poissonowska natura zliczania atomów; na pewnym etapie pomiaru wynik nadal opisywany jest rozkładem normalnym i jego forma (5670±50 lat) odnosi się do dwóch parametrów tego rozkładu. Jednak obecnie fizycy dysponują kalibracją swej metody pomiarowej (P.J. Reimer i inni 2004; C. Bronk Ramsey 1998). Kalibracja oznacza, że np. zamiast 5670 lat przyjąć trzeba jako poprawną wartość raczej 6450 lat. Różnica jest dość duża (w porównaniu z ±50 lat), jednak nie o to chodzi. Otóż kalibracja nie oznacza po prostu korygującego dodawania czy mnożenia. Nie wystarcza też żadna „zwykła" funkcja, na przykład jakiś wielomian. Zależność kalibrująca jest empiryczną (otrzymaną w pomiarach) funkcją mającą formę tabeli lub zębatej, nieregularnej krzywej. Postać funkcji kalibracyjnej odzwierciedla złożoność przyrody (powstawania ^{14}C, obiegu węgla itp.). Niepewność pomiarowa krzywej kalibracyjnej, w większości typowych sytuacji nie ma praktycznego znaczenia i nie będzie tu omawiana.

Wynik pomiaru wieku obiektu archeologicznego, po obowiązkowym zastosowaniu kalibracji, jest rozkładem prawdopodobieństwa, którego kształt przypomina pasmo górskie typu alpejskiego. Gęstość prawdopodobieństwa, w pewnym zakresie wieku wielokrotnie wzrasta i opada, osiągając maksimum w środku lub na skraju. Nierzadko też w środku rozkładu (mediana) przyjmuje wartość bliską zeru. Skomplikowany graficzny obraz tego rozkładu sugeruje jakoby niósł on ogromną ilość informacji. Tymczasem jest raczej odwrotnie.

Funkcja gęstości prawdopodobieństwa ma interpretację dość prostą, pozwala powiedzieć, jakie jest prawdopodobieństwo, że wiek wynosi np. 6444 lata.

W tym miejscu należałoby, być może, sprecyzować rozumienie pojęcia prawdopodobieństwa. Traktując jako darwinowską, wytyczną dążenia do prawdy zgodzić się wystarczy, że archeolog zobowiązany jest przyjąć rozwiązanie (wiek) najbardziej do prawdy podobne. Oczywiście po Bayesowskim uwzględnieniu (przez pomnożenie) swych prawdopodobieństw *a priori*, o czym będzie jeszcze mowa niżej.

Obserwacja praktyki wykorzystania dat ^{14}C skłania do zajęcia się kwestią ilości informacji zawartej w bardzo skomplikowanym rozkładzie prawdopodobieństwa, który archeolodzy czują się zobowiązani publikować *in extenso*. W zwykłych pomiarach, gdzie funkcjonuje rozkład normalny, opisywany wyczerpująco dwoma liczbami, również można by podawać wynik pisząc, że prawdopodobieństwo wieku (by pozostać przy tej wielkości) t1=p1, t2=p2, ..., tn=pn. Jednak zwykła oszczędność, a właściwie nawet uczciwość, nakazuje odwołać się do powszechnie znanej funkcji Gaussa (rozkładu normalnego). Krzywa kalibracyjna ^{14}C nie ma tak mocnego statusu jak funkcja rozkładu normalnego, ma charakter empiryczny, ale globalny (w przybliżeniu). W każdym pomiarze wieku stosuje się tę samą krzywą kalibracyjną. Dwa wyniki identyczne przed kalibracją pozostaną takie po kalibracji. Jasne jest więc, że publikując wynik pomiaru wieku wystarczyłoby podać dwa parametry rozkładu normalnego (tzw. datę konwencjonalną) i, być może, przypomnieć, że istnieje kalibracja. Tak też dzieje się w geologii, gdzie krótkowzrocznie, nie zaprzątając sobie głowy kalibracją, przyjęto tzw. radiowęglową skalę czasu (byt ewidentnie zbędny), bo metoda ^{14}C dominuje tam na rynku pomiarów wieku. Archeolog jednak nie może sobie na takie uproszczenie pozwolić gdyż zna wiele dat z innych źródeł. Tu dąży się do prawdy o zwykłym wieku kalendarzowym.

Z otrzymanego, wydawałoby się precyzyjnego, bo skomplikowanego, alpejskiego w kształcie, rozkładu prawdopodobieństwa, archeolog rezygnuje niechętnie. Fizycy sugerują użycie przedziału wieku jako końcowego wyniku pomiaru, przedziału obejmującego rzeczywisty wiek z prawdopodobieństwem 0,95 (stosuje się też wartość 0,68, odpowiadającą konwencji podawania wieku Gaussowskiego). Przedział określają dwie liczby, czyli jest to minimum informacji. Jednak, naturalna dążność do „precyzji" rozczłonkowuje czasem ten przedział w imię minimalizacji sumarycznej długości przedziałów cząstkowych. Dzieje się tak, gdy w środku jest minimum prawdopodobieństwa, które warto ominąć. Podaje się wtedy dwa, trzy, rzadziej więcej przedziałów wraz z odpowiadającymi im prawdopodobieństwami (sumującymi się do 0,95). Względną klarowność takiej operacji mącą dwa fakty. Jeden wynika z ułomnego charakteru wiedzy empirycznej w ogóle, którego skutki tu właśnie podlegają, wydaje się, pewnemu wzmocnieniu. Otóż zawsze większe od zera (oceniane dla metody ^{14}C, na co najmniej 0,1) prawdopodobieństwo błędów „grubych" radykalnie psuje precyzyjny obraz przedziałów, nawet, gdy pomyłka, jaka zdarzyć się może od momentu odsłonięcia przez archeologa obiektu, poprzez wieloetapowy proces pomiaru, nie była zbyt wielka. Wystarczy, że błąd pomiarowy zaniżono np. dwukrotnie, co według krytycznych pomiarowców jest standardem. Drugi kłopot wynika ze zdrowej w zasadzie dążności do wydobycia maksimum wiedzy z wyniku pomiaru. Archeolog woli wyszukać odpowiadający mu wierzchołek góry prawdopodobieństwa i wskazać na jedną wartość wieku. Ma do tego prawo, ale jeśli chce uzasadniać dalej swą tezę wynikiem pomiaru, to zobowiązany jest do konsekwentnego przestrzegania reguł ilościowych.

Trudno dziwić się tendencji do uzyskania jako wyniku pomiaru jednej tylko wartości. Nawet fizyk zapytany o masę elektronu, w pierwszym odruchu poda jedną wartość i dopiero w lepszych tablicach znajdzie aktualną wartość błędu pomiarowego. Jednak w przemyśle, w praktyce, która podlega ciągłym silnym naciskom w kierunku „dobrej roboty", duże znaczenie przywiązuje się do poprawnej oceny błędu pomiarowego. A wszędzie tam funkcjonuje w zasadzie rozkład normalny, który jako tako umożliwiałby podanie jednej liczby.

W literaturze archeologicznej twierdzenie Bayesa stosowane jest nawet na bardzo zaawansowanym poziomie (być może właśnie tu wykorzystywane jest najintensywniej, patrz: C.E. Buck, J.B. Kenworthy, C.D. Litton, A.F.M. Smith 1991; C.E. Buck, W.G. Cavanagh, C.D. Litton 1996). Sformułowanie własnej wiedzy *a priori* w postaci rozkładu prawdopodobieństwa nie musiałoby być trudne, wymaga jednak pewnego skoku myślowego. Przede wszystkim pogodzenia się z „rozmyciem" tej wiedzy, a po drugie ujęcia liczbowego tego „rozmycia", podania prawdopodobieństwa (typu: 0,8, 0,2) konkretnych dat. W przypadku ciągłych rozkładów *a priori* będzie trudniej. We wzorze Bayesa po prostu mnoży się prawdopodobieństwo *a priori* przez prawdopodobieństwo z rozkładu otrzymanego w pomiarze (końcowy wynik normuje się do

1). Jeżeli archeolog proponuje przed pomiarem dwie daty, z których, po pomiarze tylko jedna leży w sensownym zakresie rozkładu, to wynikiem *a posteriori* będzie jedna data i cel będzie osiągnięty. Nie ulega wątpliwości, że w takim wypadku, jak i w innych przypadkach realnego stosowania twierdzenia Bayesa, tzn. stosowania rozkładów *a priori* istotnie zmieniających rozkład pomiarowy, konieczne jest ujawnianie wkładu wiedzy *a priori*. Nie można, jako wyniku pomiaru podawać, bez komentarza, wyłącznie rozkładu *a posteriori*. Nie można więc, z wielomodalnego (alpejskiego) pomiarowego rozkładu prawdopodobieństwa wieku wybierać jednej wartości, nawet tej spod Mont Blanc, i podać jej jako wyniku pomiaru.

Zarysowane tu kłopoty z podstawowym w archeologii pomiarem wieku metodą ^{14}C występują w codziennej praktyce. Stąd, ironiczny tytuł tej wypowiedzi w jakimś wąskim sensie mógłby być zdaniem prawdziwym. Wymaga się od archeologów, z natury humanistów, tak zaawansowanego podejścia ilościowego, z jakim nieczęsto spotykają się nawet fizycy. Archeologia źle znosi tę sytuację.

Bynajmniej nie wyjątkowym przykładem będzie tu data radiowęglowa podana w następujący sposób (V. Mihailescu-Birliba, M. Szmyt 2003): BC 1 sigma (68,2%): 3500-3460 (9,5%), 3380-3330 (27,8%), 3220-3180 (14,5%), 3160-3120 (16,3%); BC 2 sigma (95,4%): 3500-3430 (13,6%), 3380-3260 (37,5%), 3240-3090 (44,3%). To jest jedna data (podawanie prawdopodobieństwa z dokładnością do 0,1% nie ma nigdzie uzasadnienia w naukach dedukcyjnych). W przekonaniu autora, należy podać: 3500-3090 (95%) (A. Walanus, T. Goslar 2004).

Wiedza zdobyta przy zastosowaniu kosztownych pomiarów bywa zarówno nadużywana, jak i nie w pełni wykorzystana. Konieczne, wobec nadużyć, częste podważanie dat nie sprzyja współpracy interdyscyplinarnej, która przy dystansie dzielącym naukę instytucjonalnie humanistyczną od fizyki, nauki najściślejszej, jest i tak materią delikatną. Być może problem ma naturę socjologiczną i wystarczyłoby gdyby archeolodzy zechcieli więcej energii włożyć w zrozumienie tego, co otrzymują od fizyków. Autor nie jest jednak kompetentny by dokonać bilansu społecznych kosztów takiej, ewentualnej redystrybucji sił. Również nauki przyrodnicze, jak biologia, geografia, geologia, przymuszone wymogiem aktualności metod stosują zaawansowane procedury obliczeniowe w sposób przynoszący więcej szkody niż pożytku (A. Walanus 2000). Prosty postulat rzetelności w postępowaniu, który powinien tu wystarczyć, staje się coraz trudniejszy do spełnienia wobec rozwoju metod analitycznych. Nie można być jednocześnie archeologiem i fizykiem. A przecież i nie każdy fizyk wie, co to jest funkcja gęstości prawdopodobieństwa.

WYKAZ CYTOWANEJ LITERATURY

Bayes T.R.
1763 *An essay towards solving a problem in the doctrine of chances* "Philosophical Transactions of the Royal Society" t. 53, s. 370-418.

Buck C.E., Kenworthy J.B., Litton C.D., Smith A.F.M.
1991 *Combining archaeological and radiocarbon information: a Bayesian approach to calibration*, "Antiquity" t. 65, nr 249, s. 808-821.

Buck C.E., Cavanagh W.G., Litton C.D.
1996 *Bayesian approach to interpreting archaeological data*, Wiley.

Bronk Ramsey C.
1998 *Probability and dating*, "Radiocarbon", t. 40, nr 1, s. 461-474.

Durka P. J.
2003 *Wstęp do współczesnej statystyki*, Warszawa.

Michczyńska D. J., Pazdur M. F., Walanus A.
1989 *Bayesian approach to probabilistic calibration of radiocarbon ages*, PACT 29, s. 69-79.

Mihailescu-Birliba V., Szmyt M.
2003 *Radiocarbon chronology of the Moldavian (Siret) subgroup of the Globular Amphora culture, [w:] The foundations of radiocarbon chronology of cultures between the Vistula and Dnieper: 4000-1000 BC*, "Baltic-Pontic Studies" t.12, s. 82-112.

Reimer P.J., Baillie M.G.L., Bard E., Bayliss A., Beck J.W., Bertrand C., Blackwell P.G., Buck C.E., Burr G., Cutler K.B., Damon P.E., Edwards R.L., Fairbanks R.G., Friedrich M., Guilderson T.P., Hughen K.A., Kromer B., McCormac F.G., Manning S., Bronk Ramsey C., Reimer R.W., Remmele S., Southon J.R., Stuiver M., Talamo S., Taylor F.W., van der Plicht J., Weyhenmeyer C.E.
2004 „Radiocarbon" t. 46, s. 1029-1058.

Walanus A.
1983 *Zagadnienia podstawowe interpretacji wyników pomiarów fizycznych na przykładzie datowań metodą ^{14}C,* „Archeologia Polski", t. 28, z. 1, s.7-17.

2000 *Istotność statystyczna wniosków z analiz ilościowych na przykładzie badań górnego czwartorzędu*, „Kwartalnik AGH, Geologia", t. 26, z. 4, s. 469-527.

2004 *Wyznaczanie wieku metodą ^{14}C,* „Czasopismo Archeologiczne Menhir", nr. 4, s. 72-73.

Walanus A, Goslar T.
2004 *Wyznaczanie wieku metodą ^{14}C dla archeologów*, Rzeszów.

ADAM WALANUS

ARCHEOLOGIA AWANGARDĄ NAUK ŚCISŁYCH.

O FORMIE KALIBROWANYCH DAT ^{14}C

S t r e s z c z e n i e

Datowanie radiowęglowe jest typowym przykładem fizycznego pomiaru stosowanego w archeologii. Jak w każdym pomiarze, tak i tu ważny jest losowy czynnik błędu pomiarowego. Znana jest postać tzw. dat konwencjonalnych typu: 5670±50 BP, gdzie wyraźnie podany jest błąd pomiarowy (niepewność pomiarowa), który, w tym wypadku wynosi 50 lat. Sens błędu określa rozkład prawdopodobieństwa normalny (Gaussa), jakiemu podlega rzeczywisty wiek próbki (który, choć ustalony, to jednak jest nieznany). Sens ten jest, miedzy innymi, taki, że z prawdopodobieństwem 0,68 rzeczywisty wiek (radiowęglowy) próbki zawiera się w przedziale ±50 lat, wokół daty 5670 BP.

Taka sytuacja, typowa dla nauk przyrodniczych, fizyki i techniki, jest łatwa, a nawet banalna, w porównaniu z tym, z czym archeolodzy zostali skonfrontowani w momencie wprowadzenia kalibracji metody radiowęglowej. Obecnie, na skutek skomplikowanego kształtu tzw. krzywej kalibracyjnej, równie skomplikowany jest kształt rozkładu prawdopodobieństwa rzeczywistego wieku próbki (tu już zwykłego, kalendarzowego wieku). Rozkład ten przypomina pasmo górskie typu alpejskiego, miewa wiele maksimów, doliny, odseparowane od głównego masywu wyspy prawdopodobieństwa. Tak trudnego do interpretacji wyniku pomiaru można długo szukać w naukach uznawanych za ścisłe. Stąd ironiczny tytuł niniejszego artykułu.

Nie przypadkiem w archeologii stosuje się zaawansowane metody ścisłego uwzględniania niepewnej informacji *a priori*, bazujące na twierdzeniu Bayesa. Zachowując wysoką, matematyczną dyscyplinę wnioskowań, można w pełni wykorzystać całą informację zawartą w dacie kalibrowanej. Jednak informacji tej nie jest tam tak wiele, jak wydawałoby się patrząc na skomplikowany kształt rozkładu prawdopodobieństwa wieku kalibrowanego, często przytaczany *in extenso* w pracach archeologicznych. Sytuacja staje się więc trudna.

Komputerowe programy do kalibracji dat radiowęglowych, poza rozkładem w postaci graficznej podają też przedziały wieku, w których rzeczywisty wiek próbki mieści się z zadanym prawdopodobieństwem. Przy tym, ze względu na tradycję rozkładu normalnego, stosuje się dwa prawdopodobieństwa: 0,68 i 0,95.

Otóż, archeologom, którzy nie poczuwają się do bycia awangardą nauk ścisłych zalecić można używanie wyniku datowania w najprostszej postaci, jako przedziału odpowiadającego prawdopodobieństwu 0,95, przy tym, niech będzie to przedział ciągły, nie poprzerywany odcinkami o nieco niższych wartościach prawdopodobieństwa. Prawdopodobieństwo 0,68, to zaledwie 2/3; w jednym przypadku na trzy rzeczywisty wiek będzie poza podawanym przedziałem. Nawet 0,95, to jednak nie 1. Zakładając, że wszelkie warunki poboru próbki i pomiaru w laboratorium były idealne, i tak, średnio, w co dwudziestym przypadku rzeczywisty wiek będzie poza podawanym przedziałem (o czym, oczywiście nie będzie wiadomo).