



DROGI I BEZDROŻA STATYSTYKI W BADANIACH NAUKOWYCH

prof. zw. dr hab. inż. Ryszard Tadeusiewicz⁷

Seminarium, którego finalnym elementem ma być ten referat, poświęcony jest głównie temu, aby zapoznać P.T. Uczestników ze stosowaniem statystyki jako metody naukowej i z programem *STATISTICA* jako jej praktyczną implementacją. Wystąpienia skoncentrowano więc na tym, jak można rozwiązywać różne problemy naukowe przy użyciu metod i technik statystycznych. W badaniach naukowych statystyka rzadko występuje jako element samodzielny, znacznie częściej wspomaga badacza w poprawnym wnioskowaniu na tematy gospodarcze, medyczne, techniczne, przyrodnicze, społeczne i inne. Słuchając wcześniejszych referatów poznaliście Państwo konkretne metody wykorzystania statystyki w badaniach naukowych – w naukach społecznych, w medycynie, w chemii. Posługując się tymi metodami można uzyskać odpowiedzi na wiele ważnych pytań, gdyż dobrze stosowana statystyka potrafi ujawniać prawidłowości i współzależności, tkwiące w zebranych informacjach, ale niezauważalne w przypadku analizowania bezpośrednio nie przetworzonych danych źródłowych.

W tym końcowym referacie chciałbym oderwać się od konkretnych zastosowań, pokazując zamiast tego pewne ogólne właściwości statystyki jako narzędzia w instrumentarium naukowca prowadzącego badania naukowe. Zacznę od skrótowego przeglądu najważniejszych zalet i atutów technik statystycznych, które bez wątplenia zasługują na pozytywną ocenę, ponieważ pozwalają ustalać i wykrywać fakty naukowe, których bez sięgania do tych technik nie dało by się wcale stwierdzić, lub stwierdzenie takie obciążone było by ogromnym bagażem niepewności. Potem jednak spróbuję uwrażliwić Czytelników na fakt, że statystyka nie zwalnia badacza od obowiązku myślenia i rozumienia problemów, które analizuje, zaś pozorna łatwość użycia technik statystycznych (zwłaszcza przy korzystaniu z zaawansowanych narzędzi informatycznych, które uwalniają naukowców od trudu „ręcznego” przetwarzania zgromadzonych danych), może łatwo stać się pułapką, w której wśród masy tabel, eleganckich wykresów i mądrze nazywanych wskaźników - utracony zostanie bezpowrotnie sens i cel badań. Spróbuję w związku z tym podzielić się z Państwem moimi doświadczeniami, które zdobyłem podczas z górą 30 lat stosowania statystyki w różnych obszarach badań naukowych, koncentrując się głównie na tym, jak prawidłowo eksploatować tę dziedzinę. Serce mi bowiem krwawi, gdy widzę wartościowe prace empiryczne, w których statystyka jest stosowana wyłącznie jako zwyczajowy ozdobnik, konieczny do tego, by mogła zostać zaprezentowana jako publikacja naukowa, ale w najmniejszym stopniu nie wzbogacający merytorycznej zawartości. Nieomal regułą jest, że z ogromnym trudem zdobyte obserwacje albo kosztowne pomiary (koszt ten trzeba rozważać nie tylko w sensie ekonomicznym – często ma on także wymiar deontologiczny, np., w przypadku doświadczeń dokonywanych na zwierzętach!) – są potem marnowane, gdyż uzyskany wynik jest podawany w postaci nieomal zupełnie surowej, jedynie bardzo „płytko” przetworzonej statystycznie. Takie postępowanie przypomina politykę krajów zacofanych cywilizacyjnie, które rabunkowo eksploatują swoje surowce mineralne, eksportując je potem w postaci praktycznie nie przetworzonej!

⁷ Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie.



Jeśli w wyniku badań naukowych zgromadzone zostały nowe wyniki, jeśli zdobyto cenne dane, unikatowe obserwacje, pomiary, których nikt nigdy wcześniej nie wykonał ani nie opisał – to **obowiązkiem** badacza jest wyciągnięcie z nich jak największej liczby wartościowych naukowo konkluzji. Inaczej my wykonujemy kosztowne badania – a inni dochodzą na ich podstawie do wartościowych (także w sensie gospodarczym) wyników naukowych. Tak być nie powinno, a często jest...

Oczywiście statystyka nie jest tu jakimś uniwersalnym „panaceum”. Przeciwnie – często tak bywa, że użycie statystyki **zamiast** kreatywnego myślenia powoduje dodatkowe spłaszczenie i spłycenie naukowej wartości pracy. Jednak spośród posiadanych metod naukowego (a zwłaszcza matematycznego) opracowywania wyników badań naukowych (zwłaszcza przyrodniczych, ale także społecznych, ekonomicznych, technicznych i uzyskiwanych w naukach ścisłych) **mańdrze stosowana** statystyka jest zdecydowanie najbliższa ideału uniwersalnego narzędzia pomagającego konwertować **dane** naukowe na **wiedzę** naukową.

Praca ta nie rości sobie pretensji do roli kompletnego i wyczerpującego studium wszystkich problemów występujących podczas prób stosowania statystyki w badaniach naukowych. Jej rola może polegać wyłącznie na tym, że powinna pobudzić do myślenia innych badaczy, sprowokować pewien intelektualny niepokój, uświadomić, że prowadzenie badań naukowych bez żadnej próby zastosowania zaawansowanych metod statystycznych do analizy ich wyników, jest zwykłym marnotrawstwem, chociaż nadużywanie statystyki (nie tylko zresztą w nauce...) bywa zwykłą głupotą.

Chciałbym bowiem również przestrzec Czytelników przed statystyką traktowaną niekiedy jako swoisty „fetyusz” albo – co gorsza – jako parawan, za którym ukrywa się rzeczywista miałkość i mizéria pseudonaukowego „wyniku”. Jeśli brak stosowania statystyki określono wyżej jako swoisty błąd w sztuce prowadzenia badań naukowych, a jej nieumiejętne (i wskutek tego zbyt płytkie) użycie wskazano jako marnotrawstwo, to użycie statystyki do ukrywania prawdy (zamiast jej eksponowania) najłagodniej określić można jako nadużycie. Czasem, chociaż chciałbym wierzyć, że w mniejszości przypadków, jest to nadużycie świadome i celowe.

Zacznijmy jednak od zalet. Otóż ogromną zaletą statystycznych metod analizy danych jest fakt, że uwalniają one umysł badacza, poszukującego rozwiązań pojawiających się przed nim problemów, od trzech głównych źródeł dyskomfortu.

- ◆ Pierwsze z tych źródeł dyskomfortu związane jest z faktem, że danych, które dostarcza prowadzony eksperyment naukowy, jest zwykle bardzo dużo. Nowoczesne przyrządy, rejestratory i analizatory mogą w krótkim czasie dostarczyć ogromnej ilości danych empirycznych, co stwarza wrażenie, że o obiekcie badań wiemy bardzo dużo. Jest to jednak mądrość pozorna, na to bowiem, by surowe dane empiryczne przekształcić w użyteczną informację, a informację potem w wiedzę – potrzeba jeszcze bardzo wiele pracy. W uproszczeniu (wystarczającym w wielu praktycznych sytuacjach) można powiedzieć, że nie wystarczy zebrać dane, aby naukowo zbadać interesujące zjawisko lub proces. Aby uzyskać wymaganą wiedzę trzeba zebrane dane rozważyć i przeanalizować. Tymczasem danych bywają tysiące, a z psychologii procesów kognitywnych wynika, że człowiek może brać pod uwagę (w dowolnym procesie analizy sytuacji i rozumowania na podstawie danych) jedynie bardzo ograniczoną liczbę faktów (niektórzy twierdzą, że taką liczbą równocześnie współuświadamianych faktów jest liczba siedem). Bardzo rzadko odpowiedź na pytanie nurtujące naukowca da się wywieść z tak niewielu empirycznych faktów!



Statystyka (i służące do jej praktycznej implementacji oprogramowanie komputerowe) są najłatwiej dostępnymi środkami, pozwalającymi zamienić duży zbiór nie uporządkowanych danych doświadczalnych w zwartą i treściwą reprezentację, możliwą do ogarnięcia myślą, a dzięki odpowiedniej wizualizacji – także łatwą do percepcyjnej oceny. W ten sposób statystyka pomaga zrozumieć dane, chociaż zwykle odbywa się to kosztem utraty szczegółowości prowadzonych analiz.

- ◆ Drugie źródło dyskomfortu, dla którego statystyka jest w zasadzie jedynym dostępnym remedium, jest kwestia pewnego wnioskowania w oparciu o niepewne dane. Jest trywialnie oczywiste, że stałym celem naukowców jest dążenie do odkrywania prawd pewnych, bo tylko takie mają wartość naukową. Jednak cała metodologia badań empirycznych, służących do praktycznej realizacji tego postulatu odkrywania prawdy, zmusza nas do uznania faktu, że dane na których się opieramy, są wysoce niepewne. Niepewność ta wynika z różnych błędów pomiarowych, z niejednorodności badanego fragmentu rzeczywistości, wreszcie – z niedoskonałości modeli, jakie stosujemy do interpretacji gromadzonych informacji doświadczalnych.

Statystyka pozwala wyeliminować lub ograniczyć przynajmniej niektóre spośród wyżej wymienionych czynników zmienności, zatem jest bardzo użyteczna, bo na wynikach obliczeń statystycznych można prowadzić rozumowanie i wnioskowanie znacznie bardziej pewne i znacznie lepiej kontrolowane (także w myśl kryteriów formalnych), niż by to było możliwe przy korzystaniu wyłącznie z wejściowych (oryginalnych) danych badawczych. Wymieniona zaleta znajduje zastosowanie głównie dzięki temu, że dysponujemy obecnie odpowiednimi programami pozwalającymi dokonać automatycznie wszystkich niezbędnych obliczeń, jako że w nowoczesnej statystyce droga od empirycznych przesłanek do statystycznie zweryfikowanego wyniku bywa niekiedy bardzo długa i uciążliwa numerycznie. Jednak w sytuacji, gdy odpowiednio mądre oprogramowanie nie tylko wykonuje całkiem automatycznie wszystkie wymagane obliczenia, ale także kontroluje stosowalność odpowiednich wzorów i samodzielnie sugeruje użycie najbardziej właściwych (metodologicznie) technik i form analiz statystycznych – znika większość ograniczeń utrudniających (jeszcze niedawno) stosowanie technik statystycznych przez mniej doświadczonych użytkowników, zatem można uznać, że **zaniechanie** statystycznej analizy empirycznych danych jest swoistym „błędem sztuki”.

- ◆ Trzecie źródło dyskomfortu wiąże się (poniekąd) z tym wcześniej omówionym, jednak ze względu na jego znaczenie i dokuczliwość - celowe jest jego osobne wskazanie i oddzielne omówienie. Otóż przy analizowaniu problemu naukowego badaczowi przeszkadza nie tylko nadmierna **ilość** gromadzonych w badaniach naukowych danych empirycznych, a także ich niska **jakość** (w sensie „zaszumienia” przypadkowymi błędami lub artefaktami metodologicznymi) ale także niska **relevancja** zdobywanych danych w odniesieniu do zadania, które trzeba rozwiązać.

Chodzi o to, że dane „surowe”, zdobyte w trakcie obserwacji lub eksperymentu, zawierają wprawdzie w sobie odpowiedzi na pytania nurtujące badacza, jednak odpowiedzi te są w nich zawarte w formie „uwikłanej”. Na przykład na ogół mało przydatna jest informacja, jakie wartości mają kolejno rejestrowane obserwacje x_i oraz y_i , natomiast ważne jest, czy są one ze sobą związane albo czy się od siebie różnią w sposób istotny. Statystyka dostarcza w tym przypadku bardzo użytecznych modeli pozwalających zbadać, co właściwie **wynika** ze zgromadzonych danych. Czasem modele te są bardzo proste (np. wystarczy określić korelację lub zbadać istotność różnicy z pomocą testu Studenta), czasem jednak mogą to być bardzo zaawansowane techniki, odwołujące się np. do wielowymiarowych modeli regresyjnych, analizy czynnikowej lub dyskryminacji wielokrotnej. Nie jest jednak nadmiernie ważne, czy wybrany model jest prosty, czy bardziej złożony – jego użycie jest bowiem zawsze oparte na wywołaniu



stosownej gotowej procedury w odpowiednim pakiecie oprogramowania, a skuteczne zastosowanie jest głównie uzależnione od wiedzy i wyobraźni badacza. Mówiąc prościej – można z pomocą statystyki uzyskać dowolną odpowiedź (choć czasami może to być odpowiedź informująca o tym, że nie ma odpowiedzi...), pod warunkiem, że umie się postawić pytanie.

Przytoczono wyżej **zalety** statystyki w badaniach naukowych, chociaż są zapewne ogólnie znane i chociaż zostały na użytek tej pracy skomprimowane do zaledwie trzech punktów. Trzeba jednak powiedzieć także o tym, że statystyka ma w badaniach naukowych także zdecydowane wady. Chcąc gorąco rekomendować wszystkim słuchaczom i wszystkim czytelnikom statystykę jako jedno z najdoskonalszych narzędzi badawczych – muszę także wskazać na jej ograniczenia, a także na możliwe nadużycia, jakie mogą się pojawić (i pojawiają się!) przy korzystaniu z narzędzi statystycznych w badaniach naukowych.

Zacznę od tego, że o ile poważnym błędem jest pominięcie statystycznej analizy wtedy, gdy jest ona potrzebna (pisałem o tym wyżej), o tyle jeszcze poważniejszym nadużyciem jest wypychanie jej tam, gdzie potrzebna nie jest. Czasem jest to nadużycie świadome, gdy badacz np. podaje tylko wartości średnie pomierzonych danych po to, aby ukryć i rzeczywisty rozrzut, uniemożliwiający jakiegokolwiek wnioskowanie przed wykonaniem nowych, znacznie dokładniejszych pomiarów. Częściej jednak zdarza się, że naukowiec odwołuje się do terminów i pojęć statystycznych nie w pełni świadomie, na zasadzie przeniesienia do publikacji naukowej zwrotów i sformułowań zaczerpniętych z życia codziennego. Zaskakujące jest bowiem stwierdzenie, że wszyscy ludzie stale posługują się pojęciami i terminami mającymi związek ze statystyką. Zwłaszcza politycy wręcz uwielbiają przeróżne *średnie*, *korelacje*, *trendy* itp., które - zależnie od tego, czy należą do ugrupowania sprawującego władzę, czy też wchodzą w skład opozycji pozwalają dowiedzieć, że jest wspaniale i coraz lepiej, albo przeciwnie - że jest fatalnie i wszystko zmierza do katastrofy. Pozostawmy jednak na boku demagogiczne gry polityków i zajmijmy się naszą własną codziennością. Jednak po bliższym przyjrzeniu się stwierdzimy, że także zwykli ludzie, po prostu każdy z nas, w życiu codziennym i w pracy, chętnie posługują się pojęciami pochodzącymi ze statystyki, czyniąc to często wręcz nieświadomie. Niezwykle często mówimy na przykład o wartości średniej albo o przeciętnej – nie mając przy tym bynajmniej dokładnego poglądu na temat ich rzeczywistej wartości, a tylko wyrabiając sobie ogólny jakościowy pogląd na podstawie pewnej liczby dość arbitralnie zestawionych obserwacji. Wynika to ze wspomnianego wyżej dyskomfortu, jaki odczuwamy skonfrontowani z koniecznością wytworzenia sobie pewnego syntetycznego sądu na podstawie wielu danych szczegółowych. W takim przypadku werbalnie odwołujemy się zwykle do średniej, podczas gdy w rzeczywistości częściej prowadzimy nasze (nie do końca uświadomione) rozumowanie na podstawie innych miar tendencji centralnej, na przykład mediany albo wartości modalnej - chociaż czasem nie znamy nawet nazw tych pojęć. Jednak wypowiadając sądy w rodzaju „w tym sklepie jest drogo” albo „w tamtej firmie dobrze zarabiają” czy też „to jest bogaty kraj” w istocie sięgamy właśnie do heurystycznego uogólnienia szczegółowych obserwacji, którego korzenie tkwią w statystycznych technikach ustalania wskaźników tendencji centralnej.

Postępując tak w życiu codziennym nie musimy przy tym zwykle wykonywać jakichkolwiek obliczeń matematycznych. Wystarczy sam tylko ogólny ogląd surowych danych (cen na półkach sklepowych, zarobków ludzi, widoku domów i samochodów) by nasz mózg - bez wkładania w to świadomego wysiłku z naszej strony! - dostarczył nam pewnego syntetycznego poglądu. Ten sam proceder w przypadku badań naukowych ma jednak inny wydźwięk. Tu nie może być niedopowiedzeń czy nadinterpretacji. Jeśli w artykule naukowym mówi się, że jakaś wartość **średnia** wynosi x , to istotnie sąd ten musi być oparty na operacji uśredniania, a nie na domniemaniu, że tak



by było, gdyby tę średnią policzyć. Co więcej, mówiąc i myśląc o średniej (a nie np. medianie lub wartości modalnej) też powinniśmy zawsze maksymalnie precyzyjnie wskazywać, o jaką średnią nam chodzi – pojęcie to bowiem nie jest wcale jednoznaczne. Zwykle (choć nie znaczy to, że zawsze!) mamy do czynienia ze zwykłą średnią arytmetyczną „czystych” pomiarów. Jednak przy braku wyraźnego określenia, że tak jest w istocie, może się okazać, że właściwsza była w tym przypadku średnia ważona, uwzględniająca np. zróżnicowaną wiarygodność poszczególnych danych, a może była to średnia moderowana (powstająca po odrzuceniu krańcowych, często „odstających” obserwacji) albo średnia geometryczna, lepiej nadająca się do opisu rozkładów skośnych, albo np. średnia harmoniczna, bardzo użyteczna np. w badaniach czasu przeżycia itp. Jedno proste i pozornie oczywiste pojęcie – a tak wiele znaczeń! W wypowiedziach potocznych, w formułowaniu sądów na użytek doraźny (np. dla celów zarządzania) można się tymi subtelnościami nie trapić. W tych przypadkach ważniejszy jest wniosek, jaki się wyciąga z wartości średniej, niż to, jak dokładnie była ona liczona. Jednak praca naukowa podlega pewnym rygorom, spośród których na czoło wysuwa się rygor ścisłości. Każdy element publikacji naukowej musi być tak skonstruowany, by dokładnie informował, co właściwie badacz stwierdził – i w jaki sposób. Ważne i potrzebne jest zwłaszcza to, by w innych laboratoriach możliwe było odtworzenie drogi uzyskania rozważanego wyniku naukowego ze wszystkimi jej szczegółami – tak aby rezultat naukowy mógł być potwierdzony (lub falsyfikowany) w doświadczeniach kontrolnych. Skoro pojęcia i metody statystyczne stanowią element instrumentarium naukowego, to muszą być dookreślone maksymalnie dokładnie, gdyż użycie niewłaściwej miary statystycznej będzie równie skutecznie utrudniało potwierdzenie wyniku badania, jak podanie np. niewłaściwego ciśnienia albo użycie innego gatunku zwierząt doświadczalnych.

Przytoczone wyżej uwagi odniesiono do najprostszego (poznajmniej, jak się okazuje!) pojęcia statystycznego, jakim jest średnia. Pojęciem jeszcze bardziej nadużywanym, i to zwłaszcza w niektórych pracach naukowych, jest **korelacja**. Ilekroć badacz wykryje fakt współzależności jakichś procesów lub zjawisk – to chętnie mówi o tym, że są one skorelowane, przy czym w zdecydowanej większości przypadków wcale nie bada się, czy istotnie obserwacje współbieżności obserwacji potwierdza formalne wyliczenie współczynnika korelacji – tylko zapożycza się (by nie powiedzieć – nadużywa się) statystycznego terminu bez żadnego uzasadnienia. Tymczasem korelacja ma w statystyce kilka znaczeń, wszystkie jednak są bardzo dokładnie określone. W szczególności najpopularniejszy w praktyce współczynnik korelacji Pearsona, obliczany dla danych ilościowych skalarnych oznacza miarę siły związku **liniowego**. Jeśli więc mówimy, że jakieżś obserwacje wskazują na korelację pewnych zjawisk, a w istocie zauważyliśmy tylko ich współzależność – to sugerujemy więcej, niż mamy prawo.

Skoro już jesteśmy przy korelacji, to warto odnotować znany fakt, że stałym obszarem nadużyć tego pojęcia statystycznego jest nagminne zacieranie różnicy, jaka występuje pomiędzy związkiem wykazany na podstawie korelacji, a związkiem przyczynowym. Aż wstyd o tym przypominać, bo sprawa dotyczy jednej z elementarnych zasad używania statystyki i była setki razy dyskutowana w podręcznikach, ale wciąż trzeba powtarzać, że *wykrycie korelacji nie jest równoznaczne z ustaleniem związku przyczynowego pomiędzy zjawiskami*, podobnie zresztą jak brak dającej się statystycznie potwierdzić korelacji nie dowodzi braku związku przyczynowego. Ta ostatnia uwaga ma zresztą w statystyce znacznie szerszy zasięg stosowalności: Każde badanie, każda statystyczna ewaluacja – jest obarczona możliwością błędu. Poprawnie stosując metody statystyczne można często ten błąd zminimalizować, a w większości przypadków można go przynajmniej oszacować, jednak wyniki badań statystycznych **nigdy** nie są pewne! Dlatego często robi się w statystyce znane „łamańce” (na przykład przy weryfikacji testów statystycznych), stawiając hipotezę odwrotną do tej, która nas naprawdę interesuje i podejmując próbę jej obalenia. W wyniku takiego pokrętnego



postępowania za sukces uznajemy, paradoksalnie, kiedy obalenie się nie uda, a dokładniej, kiedy prawdopodobieństwo tego, że postawiona hipoteza jest prawdziwa, będzie wystarczająco małe. Czynimy tak, ponieważ w statystyce brakuje skutecznych i bezpiecznych technik generowania rozstrzygnięć pozytywnych. Stosując taką „okrężną” metodę dochodzenia do prawdy naukowej trzeba jednak zawsze być bardzo ostrożnym w ferowaniu na tej podstawie zbyt kategoriycznych sądów, powtarzając sobie stale starą, ale bardzo mądrą maksymę: *Brak dowodu istnienia nie jest dowodem nieistnienia!* Jakże często jednak spotyka się prace mało doświadczonych adeptów statystyki, którzy wynik testu statystycznego traktują bez mała jak prawdę objawioną, pewną, gwarantowaną, nie podlegającą dyskusji. Zjawisko to nasiliło się wraz z upowszechnieniem stosowania zaawansowanych narzędzi komputerowych, takich jak pakiet *STATISTICA*. Wynika to z faktu, że w następstwie bardzo daleko posuniętego udoskonalenia tych narzędzi obecnie użytkownikowi statystyki (lub pakietu *STATISTICA*) nie jest już potrzebna sterylnie *żadna* wiedza statystyczna. Dawniej badacz, który chciał „zaprzęgać” statystyczne wnioski do swej pracy badawczej, musiał trochę się natrudzić. Na ogół sam znał statystykę i wtedy miał świadomość, na jakich teoretycznych fundamentach (i na jakich założeniach odnośnie do danych) opierają się takie czy inne wykorzystywane przez niego techniki analizy. Z kolei jeśli badacz sam statystyki nie znał (było to nagminne zwłaszcza wśród lekarzy) – to zlecał opracowanie wyników specjalście, który (jeśli poważnie traktował swoje obowiązki) przekazywał mu wyniki wraz z odpowiednim komentarzem, skrótkowo mówiąc wskazującym, co z tych wyników naprawdę wynika. Nowoczesna technika zastąpiła takiego „wynajmowanego eksperta” tylko połowicznie. Stosując odpowiedni program można całkiem *bezboleśnie* uzyskać wyniki nawet najbardziej wyrafinowanych obliczeń i analiz. Jednak żaden program komputerowy nie zawiera niestety opcji, która pozwalała by powiedzieć użytkownikowi: *Wiesz co? Z takimi danymi to sobie możesz te wyniki do... butów wsadzić!* Zresztą nikt z użytkowników nie przyjął by takiego werdyktu od komputera, chociaż z własnych doświadczeń pamiętam, że wielu ludzi dość gładko przetykało taką informację przekazaną w przyjacielskiej formie przez wykonującego obliczenia kolegę.

Niestety, program komputerowy (zwłaszcza jeśli jest „*user-friendly*” lub jak kto woli „*idiot-proof*”), dostarcza nie tylko wyników, ale i złudnego przekonania, że się te wyniki rozumie. Spotkałem się nawet z bardzo butną postawą takich zadufanych badaczy, którzy wszelkie próby merytorycznej dyskusji na temat sensowności niektórych elementów ich rozumowania naukowego (opartego na statystyce) odpierali magicznym zaklęciem: tak wyliczył komputer, więc w tym nie może być błędu, bo komputer się nie myli. Istotnie, komputer się nie myli w obliczeniach, jednak odpowiada tylko na takie pytania, jakie mu postawiono. Dawni profesorowie mieli zwyczaj powiadać (zachęcając studentów do aktywności na wykładach) *Nie ma głupich pytań, są tylko głupie odpowiedzi*. Wiele lat doświadczeń ze stosowaniem statystyki na użytek różnych badaczy upewniło mnie, że jednak istnieją głupie pytania.

Kolejnym zagadnieniem, które wiąże się z trudną sztuką stosowania statystyki w badaniach naukowych, jest sposób prezentacji wyników statystycznego przetwarzania wejściowych danych. Pozornie technika komputerowa daje tu znowu nieograniczone wręcz możliwości. Nie ma nic miłszego, jak obejrzeć swój zbiór danych naukowych (pomiarowych, ankietowych lub pochodzących z obserwacji) po przetworzeniu statystycznym w formie eleganckiej tabeli albo perfekcyjnie dopracowanego wykresu. Okres sporządzania wykresów na papierze milimetrowym (czy naukowa młodzież wie chociaż, co to takiego?) należy już szczęśliwie do przeszłości i dzisiaj nie ma nic łatwiejszego, niż uzyskać wykres żądanych wyników, często dodatkowo w kolorze lub nawet trójwymiarowych. Jednak znowu pojawia się podobny problem, jak sygnalizowany w poprzednim akapicie. Techniczna łatwość graficznej prezentacji uzyskanego wyniku nie skłania do pogłębionej refleksji nad tym, co właściwie należy uznać za godny pokazania wynik? Tymczasem wbrew



pozorem sprawa nie jest banalna, gdyż zła forma prezentacji wyników może doprowadzić do błędnej ich interpretacji – nawet wtedy, gdy sama procedura uzyskania wyników była absolutnie poprawna. Chodzi o to, że dobór tego, co się prezentuje, a także wybór formy, w jakiej się prezentuje wynik – wpływa na sposób rozumienia (przez odbiorcę pracy naukowej) zarówno rozważanego problemu, jak i jego rozwiązania. Wychodząc od trochę górnolotnego stwierdzenia, że celem badań naukowych jest zawsze odkrycie jakiejś prawdy, można w odniesieniu do dyskutowanego tym momencie problemu sformułować postulat, że prawdy tej (po jej odkryciu) nie należy wyrażać w języku niezrozumiałym lub wieloznacznym.

Oczywiście w każdym konkretnym zadaniu chodzi o coś trochę innego, zatem trudno tu udzielać szczegółowych wskazówek, jednak przynajmniej na kilka ogólnych aspektów warto tu zwrócić uwagę. Pierwszy z nich polega na tym, by maksymalnie uczciwie ostrzegać odbiorcę danych naukowych, że prezentowany mu wynik ma ograniczoną dokładność. Statystyka wykazuje w tym zakresie wyjątkowo dużą naturalną skłonność do stwarzania wrażenia dokładności i pewności tam, gdzie jej w istocie nie ma. Być może z tego właśnie bierze się znane powiedzenie: „*male kłamstwo, łgarstwo, statystyka*”, mające obrazować, jak perfidnie metody statystyczne zakrywają niekiedy istotę rzeczy. Tymczasem statystyka jako taka jest tu niewinna – obliczenia statystyczne dostarczają danych precyzyjnych i ścisłych, gdyż jest ona takim samym działem matematyki jak trygonometria, algebra, rachunek różniczkowy i całkowy itp. Niedokładne (a czasem także tendencyjnie spreparowane, ale o tym tu nie mówimy) są natomiast z reguły wprowadzane do statystyki dane, zaś sposób prezentacji wyników może albo akcentować dokładność obróbki tych danych – albo ich naturalną, obniżającą precyzję wnioskowania, różnorodność i zmienność.

Problem polega między innymi na tym, że generalnie łatwiej i wygodniej jest formułować na podstawie statystyki sądy lub komunikaty fałszywie jednoznaczne i nadmiernie kategoriyczne. Taka jest bowiem – z samej swojej natury – istota obliczeń statystycznych. Znacznie trudniej jest zdobyć się na to, by na sąd kategoriyczny i pozornie ścisły – nałożyć dodatkowe informacje wskazujące na jego ograniczoną wiarygodność. Jak już wyżej wspomniano, jednym z głównych celów statystycznego opracowania danych jest uzyskanie informacji w bardziej zwartej postaci, a więc łatwiejszej do interpretacji. Jednak ta zwarta informacja sama z siebie nie wyraża faktu, że została uzyskana na podstawie danych, których różnorodność skłaniałaby (gdyby była znana) do bardzo ostrożnego formułowania sądów na temat tego, co właściwie z tych danych wynika. Posłużmy się najprostszym przykładem – wartości średniej. Jeśli dla jakichś danych podamy tylko samą tę wartość, to czytelnik naszego doniesienia (powiedzmy, że naukowego, ale problem pojawia się też np. w polityce czy administracji) będzie w istocie zdezinformowany, a nie poinformowany. Jeśli rozrzut danych wejściowych, dla których liczonego średnią jest bardzo duży – to średnia w istocie słabo reprezentuje te dane i w gruncie rzeczy wnioskowanie na podstawie takiej średniej jest mocno utrudnione (na przykład wykrycie faktu, że dwie średnie różnią się od siebie, może w istocie być naukowo bezwartościowe wobec potencjalnie możliwej zmienności samych średnich). Z kolei obecność w zbiorze danych wejściowym wartości „odstających” – to znaczy takich, które są istotnie różne od pozostałych, może w znacząco zdeformować wartość średniej i pozbawić ją zdolności prawidłowej reprezentacji zgromadzonych danych. Wiadomość o tym powinna także w jakiejś formie dotrzeć do odbiorcy wyniku statystycznego.

Przytoczone uwagi dotyczyły najprostszego i najpopularniejszego narzędzia statystycznego – wartości średniej, ale są one prawdziwe praktycznie dla wszystkich statystyk, jakie sobie tylko można wyobrazić. Jeśli więc nie chcemy, by opracowane i prezentowane w pracy naukowej dane statystyczne służyły tylko jako ozdóbki lub kamuflaż naukowości - to musimy zagwarantować ich przydatność do tworzenia i ferowania różnych ocen. Oznacza to, że dbając o zwartość statystycznej informacji musimy zadbać o jej pełną reprezentatywność, gdyż inaczej interesująca nas prawda



naukowa, zamiast wydobycia i ujawnienia - zostaje całkowicie zagubiona w chwili prezentowania takiego właśnie statystycznie wyznaczonego skrótu. Jak jednak zapewnić sobie pogodzenie zwartości opisu z jego kompletnością? Otóż warto pamiętać, że w statystyce dla prawie wszystkich wielkości, które mogą podlegać statystycznemu wyznaczaniu istnieją tak zwane **miary punktowe i miary przedziałowe**.

Miara punktowa to po prostu pojedyncza wyliczona wartość. Stwierdzamy, że średnia wynosi tyle, albo że korelacja jest taka – i koniec. Właśnie użycie miary punktowej jest zwykle przyczyną różnego rodzaju nieporozumień czy nawet nadużyć statystycznych. Tymczasem miara przedziałowa mówi znacznie więcej, a przy użyciu komputera jest równie łatwa do wyznaczenia. Miara przedziałowa wskazuje także najbardziej prawdopodobną wartość interesującego parametru statystycznego, nie porzeczając jednak na tym i podaje również przedział, w jakim parametr ten się mieści (z założonym prawdopodobieństwem). Jeśli zatem zobaczymy, że najbardziej prawdopodobna wartość (powiedzmy dyskutowana wyżej *średnia*) mieści się z prawdopodobieństwem 95% w przedziale, który odchyła się od jej podanej wartości o ułamek procenta w każdą stronę – to możemy na takiej wiadomości budować bardzo pewne i wiarygodne rozumowanie, gdyż mamy pewność, że średnia dobrze reprezentuje zgromadzone dane. Jeśli jednak szerokość przedziału, w którym udało się zlokalizować średnią, jest porównywalna z jej wartością – to każde rozumowanie oparte na takim wyniku musi być prowadzone z bardzo dużą ostrożnością oraz powinno być poprzedzone sprawdzeniem, dlaczego nasz wynik statystyczny jest tak kiepski – może zebrane dane są w istocie niejednorodne? Może są obciążone błędami? Może warto je przesortować, usuwając najbardziej zakłócone pomiary? Każdą z tych rzeczy warto zrobić, gdyż dopiero w sumie mogą one doprowadzić do sensowniejszego wyniku. Nie wolno z prostego wygodnictwa pozwolić na to, by opublikowana praca naukowa była pełna różnego rodzaju liczb, wskaźników, wykresów diagramów – ale nie miała sensu.

Działając w statystyce (i nie tylko w statystyce...) warto bowiem stale mieć na uwadze stwierdzenie, które przypisuje się Newtonowi: *Celem obliczeń nie są liczby, ale zrozumienie prawdy...*

Literatura

1. Marczewski W., Tadeusiewicz R.: Antropomotoryka biocybernetyczna, AWF, Kraków 1993
2. Moczko J. M., Bręborowicz G. H., Tadeusiewicz R.: Statystyka w badaniach medycznych, Springer PWN, Warszawa 1998
3. Tadeusiewicz R., Wajs W.: Informatyka medyczna. Wyd. Naukowe AGH, nr 1612, Kraków 1999
4. Tadeusiewicz R., Białas S., Szmuc T., Szymkat M. (red.): Metody i systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim, CCATIE, Kraków 1999
5. Tadeusiewicz R., Izvorski A., Majewski J.: Biometria. Skrypt Uczelniany AGH, Kraków, 1993
6. Tadeusiewicz R.: Problemy biocybernetyki, PWN, Warszawa 1993 (II wydanie poprawione)