

Mirosław Szreder

Uniwersytet Gdański

DOŚWIADCZENIE I INTUICJA W BADANIACH STATYSTYCZNYCH

Streszczenie

Celem artykułu jest zaprezentowanie i przedyskutowanie najważniejszych problemów dotyczących potrzeby wykorzystania dodatkowej poza próbą statystyczną informacji w badaniach niewyczerpujących. Szczególny nacisk położono na doświadczenie i intuicję badacza, które mogą przyczynić się do wzrostu precyzji i efektywności wnioskowania o badanej zbiorowości. Ilustracją propozycji autora są przykłady wybranych badań opinii społecznej.

Słowa kluczowe: badania niewyczerpujące (próbkowe), badania opinii, informacje spoza próby.

Wprowadzenie

Wśród filozofów nauki, mimo różnic poglądów na wiele kwestii, panuje duża zgodność co do natury samej nauki jako dzieła rozumu – łac. *ratio*. Ponieważ nauka jest dziełem rozumu, więc dostarcza racjonalną wiedzę [6, s. 210]. Właśnie w racjonalności tkwi atut postępowania naukowego, także w naukach ilościowych: matematyce, statystyce, ekonometrii. Opór i nieufność budzi to, co w postępowaniu badacza jest nieracjonalne, subiektywne, intuicyjne. Intuicja, jako forma aktywności psychicznej „jest funkcją irracjonalną (w sensie – pozaracjonalną), która w formie nieświadomej komunikuje postrzeżenia” [6, s. 106]. Mimo pozornego nieprzystawania do siebie nauki i intuicji,

jako dwóch różnych kategorii – racjonalnej i nieracjonalnej – warto dostrzec obszary ich wzajemnego przenikania się w badaniach naukowych.

1. Rola intuicji w nauce

W dziełach dawnych i współczesnych filozofów nauki łatwo daje się odnaleźć relacje między intuicją i nauką. Współczesna filozof nauki, prof. Alina Motycka, stwierdza wprost: „Nauka nie jest wyłącznym dziełem rozumu, lecz także intuicji. Jedynie w kontekście uzasadniania nauka jest dziełem rozumu. Ale z kolei w kontekście tworzenia nauka jest dziełem intuicji” [6; s. 227]. W innym miejscu dodaje: „Cóż bowiem może rozum, jeśli najpierw nie olśni go intuicja” [6, s. 228]. Tę inspirującą dla nauki rolę intuicji podkreślał w swoich *Pismach filozoficznych* Albert Einstein: „Prawdziwie wielkie osiągnięcia powstały na drodze niemal diametralnie odwrotnej w stosunku do indukcji. Intuicyjne uchwycenie tego, co istotne w wielkim kompleksie faktów, prowadzi badacza do sformułowania hipotetycznego prawa lub kilku praw” [1, s. 44].

Intuicja w nauce może być inspirująca, ale może też utrudniać zrozumienie jakiegoś zagadnienia, a nawet prowadzić badacza na manowce. Percepcja otaczających nas zjawisk, a dalej utrwalenie sądów o tych zjawiskach, jest kombinacją realnej rzeczywistości oraz sposobu jej postrzegania, w tym naszej interpretacji i odczuć. Wszystko to, co jest produktem ludzkiej wyobraźni i intuicji, a także sądów opartych na indywidualnych spostrzeżeniach, w tym na ciągach wielu obserwacji, które mielibyśmy prawo nazwać doświadczeniem, powinno więc podlegać wszechstronnej i dokładnej weryfikacji naukowej. Dotyczy to zarówno prostych, jak i skomplikowanych zdarzeń. W przypadku jednych i drugich intuicja może być zarówno pomocna, jak i zwodnicza. Nawet w tak prostych kwestiach, jak pojęcie i rozpoznanie losowości, większość osób – jak dowodzą psychologowie – popełnia błędy¹. Znane w matematyce paradoksy i sofizmaty są dobrą ilustracją sytuacji, w których intuicja nie wspiera zrozumienia lub wyjaśnienia określonej sytuacji (zdarzenia). Błąd hazardzisty, pole-

¹ Przykład typowego błędu w rozumieniu losowości podaje L. Młodinow [5, s. 203]. Firma Apple wyposażyła swoje iPody w opcję losowego wyboru piosenek: prawdziwa losowość prowadzi nieraz do powtórzeń, ale użytkownicy, słuchając parokrotnie tej samej piosenki, nabrali przekonania, że dobór wcale nie był losowy. Steve Jobs, założyciel Apple, ostatecznie stwierdził, że trzeba było zmienić mechanizm na „mniej losowy po to, żeby wydawało się, że jest bardziej losowy”.

gający na przekonaniu, że długa seria przegranych zwiększa prawdopodobieństwo wygrania w następnej próbie, jest jednym z tych paradoksów. „Naprawianie błędów tego rodzaju stanowi terapeutyczną stronę teorii statystyki” – twierdził logik i filozof, profesor Uniwersytetu Harvarda Willard van Orman Quine [8, s. 77]. W samej statystyce odwoływanie się do intuicji może przynieść różne skutki. Pozytywne, na przykład w wyjaśnieniu tego, że rozproszenie cechy w populacji istotnie wpływa na liczebność próby, jaką z tej populacji należy pobrać do badania reprezentacyjnego; negatywne, gdy intuicja zdaje się sugerować, że liczebność próby w takim badaniu powinna zależeć od wielkości populacji.

W kwestii reprezentatywności obserwacji, jakich każdy z nas dokonuje na co dzień, intuicja również zdaje się często błędnie sugerować, że obserwacje te są wystarczające do wiarygodnych (bardzo prawdopodobnych) uogólnień. Ludzie zdają się wierzyć w „prawo małych liczb”², jak określili to słynni badacze Amos Tversky i Daniel Kahneman³ w latach 70. XX wieku [13]. Oczywiście, nie istnieje takie prawo. Istnieją prawa wielkich liczb, jednak w małych próbach, przeświadczeni o wadze ich losowości, dość często ulegamy złudzeniu, że zaobserwowane w nich prawidłowości można z dużym prawdopodobieństwem uogólnić na populację. Wyobraźmy sobie scenę polityczną, na której są tylko dwie partie, Y i Z, cieszące się jednakowym poparciem opinii publicznej (wyborców). Oznacza to, że losując z tej populacji wyborcę, prawdopodobieństwo, że będzie to zwolennik partii Y wynosi $\frac{1}{2}$, tyle samo, ile prawdopodobieństwo, że będzie to zwolennik partii Z (także $\frac{1}{2}$). Załóżmy, że polityk, niedowierzając tym proporcjom, postanowił wylosować próbę pięciu wyborców. Czy gdyby w próbie tej uzyskał czterech lub pięciu zwolenników tylko jednej partii, miałby prawo kwestionować jednakową popularność obu tych partii? W sensie statystycznym pytanie to brzmiałoby następująco: czy wobec wyniku: czworo lub pięcioro zwolenników jednej partii Y lub Z w próbie pięcioelementowej, należy odrzucić hipotezę o jednakowej popularności obu partii? Czy uzyskanie takiej próby ma na tyle małe prawdopodobieństwo, że raczej skłonni będziemy kwe-

² „People’s intuitions about random sampling appear to satisfy the law of small numbers, which asserts that the law of large numbers applies to small numbers as well” [13, s. 106].

³ D. Kahneman jest psychologiem, laureatem Nagrody Nobla w dziedzinie ekonomii z 2002 roku. Wspólnie z A. Tverskym opublikowali wiele cennych wyników badań na temat istoty losowości, wyobrażeń o losowości, zdolności człowieka do kwantyfikacji stopnia niepewności.

stonować, iż obie partie mają tyle samo zwolenników? Większość odpowie raczej twierdząco. Proporcja $4/5 = 80\%$ lub $5/5 = 100\%$ na korzyść jednej z partii daje dość silne przekonanie o niejednakowym poparciu dla partii. Tymczasem, wykonując proste obliczenia (por. załącznik), okazuje się, że prawdopodobieństwo takiej próby wynosi aż $3/8$, czyli jest bliskie $1/2$. Intuicja źle nam sugeruje, że prawdopodobieństwo takich zdarzeń jest bardzo małe. Większą liczbę przykładów ilustrujących znaczenie intuicji (pozytywne i negatywne) można znaleźć w złożonych zagadnieniach współzależności cech, w szczególności w problemie przyczynowości zjawisk i zdarzeń (w statystyce i w ekonometrii).

2. Współczesny pomiar niepewności – personalistyczna interpretacja prawdopodobieństwa

We współczesnym świecie, w którym trudno cokolwiek przewidywać lub opisywać w kategoriach deterministycznych, podejście probabilistyczne staje się coraz powszechniejsze w wielu naukach, a także w życiu codziennym. Charakterystyka życia gospodarczego i społecznego coraz częściej odwołuje się do ilościowego (liczbowego) opisu, który początkowo był zaledwie uzupełnieniem opisu werbalnego (jakościowego), a obecnie coraz częściej zastępuje ten jakościowy opis. Mnogość najróżniejszych wskaźników koniunktury, barometrów nastrojów, prognoz ilościowych, rankingów, sprawia że na znaczeniu zyskuje nie tylko statystyka, odpowiedzialna za właściwą konstrukcję i interpretację wszystkich tych mierników, ale także teoria prawdopodobieństwa, dzięki której jesteśmy w stanie wyrazić stopień przekonania co do prawdziwości tych szacunków. O ile do niedawna ekonomistom wystarczały prognozy punktowe, uzyskiwane z ankiet makroekonomicznych, o tyle obecnie coraz częściej oczekuje się prognoz probabilistycznych⁴. To z kolei wymaga posłużenia się odpowiednią interpretacją prawdopodobieństwa, możliwą do zastosowania w realnych warunkach funkcjonującej gospodarki.

Rzeczywistość, którą obserwujemy i analizujemy, coraz trudniej daje się opisać kategoriami **klasycznej interpretacji prawdopodobieństwa**. Wynika to z samej istoty interpretacji klasycznej, w której zakłada się, że dane zdarzenie można zdekomponować na przeliczalną liczbę jednakowo możliwych przypad-

⁴ O tym, czy istnieje potrzeba zastępowania prognoz punktowych prognozami probabilistycznymi, a także, czy te ostanie rozwiążą dotychczasowe problemy, pisze w interesującym artykule H. Kowalczyk [3].

ków. W wielu praktycznych sytuacjach nie jest to założenie adekwatne do rzeczywistości i trudno byłoby je wiarygodnie uzasadnić. Klasyczne rozumienie prawdopodobieństwa jest niewystarczające do obliczania prawdopodobieństw takich zdarzeń, jak niewypłacalność klienta po roku od udzielenia mu kredytu, uzyskanie określonej przedziałowo kwoty przychodu przez przedsiębiorstwo po zakończeniu zadania inwestycyjnego, wzrost kursu złotego względem euro do określonego poziomu, udział w referendum co najmniej połowy uprawnionych do głosowania itp.⁵ Rozwój teorii prawdopodobieństwa, gdy prześledzić go w czasie, odbywa się w kierunku wypracowania takich interpretacji i sposobów pomiaru prawdopodobieństwa, które byłyby w stanie jak najlepiej zmierzyć i wyrazić stopień niepewności badacza (decydenta) do interesującego go zdarzenia.

W wielu praktycznych sytuacjach korzysta się z **częstościowej (statystycznej) interpretacji prawdopodobieństwa**, na której oparta jest cała teoria wnioskowania statystycznego. O ile, w klasycznej interpretacji następuje odwołanie do właściwości mechanizmu losującego (symetryczność monety lub kostki), o tyle w interpretacji częstościowej prawdopodobieństwa – nawiązanie do powtarzalności zdarzeń. Prawdopodobieństwo danego zdarzenia – zgodnie z interpretacją częstościową – jest granicą częstości względnej wszystkich zdarzeń elementarnych sprzyjających temu zdarzeniu. W praktycznych zastosowaniach abstrakcyjne pojęcie „granicy częstości względnej” zostaje zastąpione estymowaną w oparciu o dostatecznie liczną próbę wartością częstości względnej zdarzeń. W przypadku tej interpretacji powtarzalność zdarzeń jest kluczowym pojęciem. Tylko bowiem w długiej serii powtarzalnych zdarzeń można wiarygodnie oszacować względną częstość (prawdopodobieństwo) wybranych, interesujących badacza przypadków (zdarzeń). Znaczenie częstościowej (statystycznej) interpretacji prawdopodobieństwa jest obecnie większe niż kiedykolwiek wcześniej. Jest to spowodowane możliwością uzyskiwania wielu rzeczywistych lub symulowanych komputerowo dużych ciągów zdarzeń losowych. Interpretacja ta nie jest jednak pozbawiona istotnych ograniczeń. Jej krytycy zwracają uwagę na to, że nawet jednokrotne powtórzenie doświadczenia w identycznych warun-

⁵ L. Młodinow przywołuje inny przykład, nie ze sfery gospodarczej, lecz prawniczej. Gdy w maju 1994 r. Paula Jones wytoczyła prezydentowi Billowi Clintonowi sprawę karną, poproszono 200 praktykujących prawników o ocenę prawdopodobieństwa, że nie dojdzie do pełnego procesu [5, s. 37]. Trudno sobie wyobrazić, aby którykolwiek z nich, podając swoje oszacowanie zakładał spełnienie założeń klasycznej interpretacji prawdopodobieństwa. Bez zaangażowania własnego doświadczenia i intuicji nie da się w tych okolicznościach podać wartościowej oceny prawdopodobieństwa.

kach jest niezwykle trudne. Traktowanie ciągu rzeczywistych zdarzeń ekonomicznych jako zaistniałych w tych samych warunkach może więc być założeniem daleko odbiegającym od rzeczywistości. To, co jest możliwe w komputerowych symulacjach, nie zawsze jest proste w rzeczywistości. Interpretacja częstościowa nie może być stosowana przede wszystkim do zjawisk rzadkich lub niepowtarzalnych (np. niewypłacalność kontrahenta, wynik wyborów prezydenckich, bankructwo kraju strefy euro) lub takich, które się co prawda powtarzają, ale w zmieniających się warunkach. W wielu takich sytuacjach istnieje mniej lub więcej informacji, które mogłyby się okazać użyteczne do określenia, na ile prawdopodobne jest wystąpienie interesującego badacza zdarzenia, ale problem tkwi w tym, że modele klasyczny i częstościowy prawdopodobieństwa są za mało elastyczne, aby wykorzystać te informacje. Dodatkowym źródłem informacji, które często może się okazać cenne w określeniu prawdopodobieństwa wielu zdarzeń, jest wiedza ekspertów. Ale i ona trudno poddaje się kwantyfikacji w formie wymaganej założeniami wskazanych dwóch interpretacji prawdopodobieństwa.

Współcześnie dąży się do wykorzystania do oceny prawdopodobieństwa całej wiedzy o zdarzeniu, nie tylko o jego prostym modelu (tak, jak w interpretacji klasycznej) i o częstości względnej realizacji tego zdarzenia (jak w interpretacji częstościowej). W ten sposób, w wyniku badań między innymi T. Bayesa, L.J. Savage'a, B. de Finettiego [12, aneks do rozdz. 3], sformułowana została interpretacja personalistyczna prawdopodobieństwa, zwana inaczej subiektywną.

Przez **subiektywne prawdopodobieństwo** (ang. *subjective probability*) tego, że jakiś sąd na temat zdarzenia A jest prawdziwy, rozumie się stopień pewności (ang. *degree of belief*) lub przekonania danej osoby o prawdziwości tego sądu. Zgodnie z tą interpretacją prawdopodobieństwo sądu na temat zdarzenia A jest przypisane do danej osoby i może być różne dla różnych osób. W przeciwieństwie do interpretacji częstościowej prawdopodobieństwo subiektywne jest traktowane personalistycznie i warunkowo ze względu na przeszłość i całe otoczenie analizowanego zdarzenia. Jedynym ograniczeniem jest to, aby zbiór prawdopodobieństw był dla danej osoby spójny i nie łamał postulatu logicznej zgodności. „Zgodne lub spójne oszacowania są dotąd dopuszczalne, dopóki dana osoba czuje, że odpowiadają one jej przekonaniom” [14, s. 1105, tłum. własne]. Danemu zdarzeniu nie musi więc koniecznie odpowiadać jedna wartość prawdopodobieństwa. Może ona być różna nie tylko dla różnych osób,

ale także dla tej samej osoby w różnych momentach, wyrażając zmieniający się stan wiedzy tej osoby na temat danego zdarzenia i jego uwarunkowań⁶. Subiektywne prawdopodobieństwo jest możliwe do ustalenia dla zdarzeń rzadkich i jednostkowych. Wydaje się, że właśnie ta cecha powoduje, iż prawdopodobieństwo – rozumiane personalistycznie – pojawia się obecnie w wielu dziedzinach życia, w których nie występowało wcześniej.

3. Rola doświadczenia i intuicji w badaniach sondażowych

Badania ankietowe i sondażowe, które na etapie analizy danych można było do niedawna kojarzyć z dość prostym obliczaniem proporcji (wskaźników struktury), z wielu powodów wymagają obecnie stosowania znacznie bardziej zaawansowanych technik statystycznych. Najważniejsze z tych powodów to częstsza odmowa respondentów, krótki czas przeznaczony na realizację badań, konieczność korekty niektórych ocen podanych przez respondentów (np. ocen obciążonych dążeniem części ankietowanych do podania opinii oczekiwanej przez ankietera). Niezależnie od tego są takie sytuacje, w których niezbędne wydaje się łączne wykorzystanie z próbą statystyczną innych źródeł informacji, aby przedstawiona ocena uwzględniała niedoskonałość próby.

Jedną z takich sytuacji jest konstrukcja prognozy wyborczej, w której do najtrudniejszych elementów należy ocena struktury głosujących w dniu wyborów. Nawet gdyby próba respondentów w badaniu sondażowym doskonale odzwierciedlała strukturę postaw wyborczych w populacji uprawnionych do głosowania, to przy konstruowaniu prognozy wyborczej prawdziwym wyzwaniem dla statystyków jest umiejętność przewidzenia, jaka będzie struktura subpopulacji wyborców, którzy wezmą udział w głosowaniu. Nie ma przesady w twierdzeniu, że tego typu badanie próbkowe jest jednym z niewielu, w którym na podstawie próby wylosowanej z danej populacji oczekuje się wniosków dotyczących innej populacji, która jeszcze nie została określona. Wśród różnych sposobów weryfikacji (przesiewu) respondentów w próbie (ang. *screening techniques*) największą dotychczas popularność zdobywa sobie technika nazy-

⁶ W ten sposób interpretacja personalistyczna narażona jest na krytykę dotyczącą niespójności sądów wielu osób, a także na zarzut utraty wartości poznawczej przez prawdopodobieństwo. Dyskusję na ten temat przytaczają m.in. M.J. Machina i D. Schmeidler [4], oraz M. Szreder [11; 12].

wana *likely voter technology* [2]. Polega ona na tym, że za pomocą dodatkowych informacji o respondencie wylosowanym do próby przypisuje się mu subiektywne (personalistyczne) prawdopodobieństwo, wyrażające przekonanie badacza co do szans, że respondent ten rzeczywiście weźmie udział w wyborach. Celem dodatkowych pytań skierowanych do respondentów jest próba oceny, jak silna jest ich motywacja do udziału w akcie wyborczym. Pytania te odnoszą się do motywów głosowania, zainteresowania „wyścigiem”, udziału w poprzednich wyborach, wiedzy o lokalu wyborczym, stopnia zainteresowania lub identyfikacji z preferowaną partią polityczną itp. Od doświadczenia i intuicji badaczy zależy, jak są sformułowane te dodatkowe pytania, ile ich jest, a także to, w jaki sposób przyporządkowuje się odpowiedziom rozkład prawdopodobieństwa. Spektakularnym sukcesem tego podejścia była prawie bezbłędna prognoza wyborcza w wyborach do Kongresu USA 5 listopada 2002 roku sporządzona przez Instytut Gallupa [7].

Niezdana frekwencja wyborcza i struktura głosujących nie są jedynymi trudnościami w opracowaniu prognozy wyborczej. Źródłem potencjalnych dużych błędów mogą też być nieprawdziwe deklaracje respondentów odnośnie do ich zamiarów zachowania się przy urnie, a także zmiana wcześniejszych preferencji wyborczych w przeddzień głosowania. Są to jedne z najbardziej kłopotliwych problemów dla statystyków. Także tutaj rozwiązań należy poszukiwać przede wszystkim w innych poza próbą źródłach danych, a także w informacjach uzyskanych od ekspertów. Praktyką ośrodków badawczych we Francji jest stosowanie technik korekty w stosunku do każdej warstwy populacji, niezależnie od liczby odmów w próbie. Jej celem jest skorygowanie rozbieżności między deklarowanymi intencjami wyborczymi respondentów a późniejszym ich głosowaniem. Najpowszechniej stosuje się wagi z wcześniejszych wyborów (*diverse weights*), oparte na porównaniach z przeszłości deklaracji wyborczej z rzeczywistymi wynikami głosowania w danej warstwie. Można by to także zrobić w odniesieniu do prognoz wyborczych w Polsce, przynajmniej w stosunku do dwóch głównych partii: PO i PiS. Bardzo prawdopodobne, biorąc pod uwagę wybory parlamentarne z 2005 roku oraz wybory prezydenckie i parlamentarne z 2010 roku, przeszacowanie w sondażach popularności PO wymaga liczbowego wyrażenia tego błędu. Nie ma na to wypracowanych formuł analitycznych. Bez pewnej odwagi badaczy, wyrażającej się przede wszystkim gotowością włączenia do tego pomiaru własnych obserwacji i intuicji, będzie to trudne do osiągnięcia. O wielkości średniego przeszacowania PO w stosunku do

PiS można na przykład wnioskować na podstawie dwóch badań *exit poll* wykonanych różnymi technikami w dniu pierwszej tury wyborów prezydenckich 2010 roku. Ankieterzy MB SMG/KRC przeprowadzili 20 czerwca 2010 roku telefoniczny sondaż na próbie 6 tys. osób, które wzięły udział w tym dniu w głosowaniu. Przy tak dużej próbie waga błędu losowania w stosunku do innych błędów jest nieznaczną. Należy sądzić, że głównym źródłem błędu, jaki mógł się w tym badaniu pojawić, był błąd systematyczny, spowodowany dążeniem do „politycznej poprawności” respondentów. W tabeli 1 podano wyniki tego sondażu, a dla porównania także wyniki klasycznie wykonanego badania *exit poll* przez TNS OBOP oraz prawdziwe wyniki ogłoszone ostatecznie przez Państwową Komisję Wyborczą (PKW).

Tabela 1. Odsetki głosów oddanych w pierwszej turze wyborów prezydenckich na dwóch głównych kandydatów (prognoza MB SMG/KRC i TNS OBOP oraz oficjalny wynik PKW)

Kandydat	MB SMG/KRC	TNS OBOP	PKW
B. Komorowski	45,7	40,7	41,5
J. Kaczyński	33,2	35,8	36,5

Źródło: TVN, TVP, PKW.

Gdyby na podstawie danych z tabeli 1 sformułować ocenę stopnia przeszacowania kandydata PO, a jednocześnie niedoszacowania kandydata PiS, to wyniosłoby ono 10%–11% wyniku uzyskanego w rzeczywistości⁷. Potrzeba tej korekty jest ważna, by przeciwdziałać rosnącej nieufności do badań sondażowych, spowodowanej częstymi w przeszłości rozbieżnościami między wynikami sondaży a rozstrzygnięciami wyborczymi. Potrzebna jest odwaga realizatorów badań i większa determinacja pod tym względem. Tego typu wskaźniki korygujące, być może nieprecyzyjne i wymagające aktualizacji, są lepsze niż próby odgadywania przez dziennikarzy, jak bardzo jedna partia jest przeszacowana, a druga niedoszacowana [11]. Są także, jak się wydaje, odpowiedniejsze

⁷ Nie należy tego wskaźnika korygującego 10% lub 11% mylić z korektą o 10 lub 11 punktów procentowych. Na przykład, z badania TNS OBOP wykonanego 16–22.06.2011 r. wynika, że poparcie dla PO wynosi 45%, a dla PiS – 27% („Gazeta Wyborcza” z 24.06.2011 r., s. 4). Uwzględniając 10-proc. korektę, trzeba by przyjąć, że w rzeczywistości jest ono w przypadku PO bliskie 40,5%, a w przypadku PiS – 30%.

od standardowego, asekuracyjnego zdania w komentarzu zamieszczanym przez jeden z dzienników przy każdym sondażu: „Wybory prezydenckie 2010 r. pokazały, że niezdecydowani częściej wybierają PiS i SLD niż Platformę, dlatego przewaga PO nad partiami opozycyjnymi może być mniejsza”.

Podobnej odwagi w korzystaniu z doświadczenia i intuicji wymaga schemat losowania lokali wyborczych w badaniach typu *exit poll*. Do próby losuje się ich w Polsce zwykle około 500. Reprezentatywność próby lokali w dużej mierze wpływa na dokładność późniejszej prognozy. Statystyka także zaleca tutaj korzystanie z różnych źródeł danych. W praktyce wielu krajów ostatecznego wyboru próby lokali wyborczych dokonuje się, biorąc pod uwagę ich reprezentatywność w poprzednich wyborach. Technika ta (ang. *tied sample*) oznacza powiązanie wyboru próby ze stopniem zgodności wyników w lokalach tej próby z ogólnokrajowymi wynikami w poprzednich wyborach. W uproszczeniu można powiedzieć, że spośród rozważanych prób lokali wybiera się tę próbę, która okazała się najbardziej reprezentatywna w poprzednich wyborach. Podobnie jak w przypadku prognoz wyborczych brakuje gotowych, optymalnych procedur postępowania. Z badań prowadzonych w Katedrze Statystyki Uniwersytetu Gdańskiego wynika, że reprezentatywność większości lokali w jednych wyborach nie przenosi się na ich reprezentatywność w kolejnych wyborach. Skonstruowanie schematu wyboru próby lokali wyborczych, który skutecznie korzystałby z wiedzy z przeszłości, nadal jest zadaniem do podjęcia przez statystyków. Doświadczenie badaczy i ich intuicja są tu jednymi z najważniejszych czynników warunkujących pozytywny efekt tych prac.

Podsumowanie

Złożoność opisywanej przez naukę rzeczywistości i dynamika jej zmian powodują, że badacze coraz częściej są zmuszeni sięgać do wielu źródeł informacji, w tym także do własnego doświadczenia i intuicji. Dotyczy to różnych grup badaczy, w szczególności statystyków. W wielu sytuacjach informacja z próby statystycznej okazuje się niewystarczająca do wiarygodnych uogólnień na całą populację. Potrzebne jest jej uzupełnienie na różnych etapach badania próbkowego, począwszy od techniki próbkowania, a skończywszy na niezbędnych korektach uzyskanych w próbie wyników. W artykule omówiono kilka

tego typu sytuacji i zwrócono uwagę na intuicję i doświadczenie badacza, które mogą być dodatkowym źródłem informacji.

W populacji istnieje jednakowa liczba zwolenników dwóch partii Y i Z. Jakie jest prawdopodobieństwo, że losując z tej populacji próbę pięcioelementową, uzyska się w niej czterech lub pięciu zwolenników jednej z tych partii (czyli uzyska się proporcję w próbie 80% lub 100% na korzyść którejś z tych partii)?

Rozwiązanie:

Obliczmy najpierw prawdopodobieństwo tego, że wśród wylosowanych pięciu osób będzie pięciu zwolenników partii Y:

$$P(Y=5) = \binom{5}{5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^0 = \frac{1}{32}.$$

Takie samo jest prawdopodobieństwo, że wśród wylosowanych pięciu osób będzie pięciu zwolenników partii Z, czyli, gdyby chcieć podać prawdopodobieństwo tego, że próba losowa, składająca się z pięciu osób będzie w całości zdominowana przez zwolenników tylko jednej partii (Y lub Z), to wynosi ono $2/32$, tj. $1/16$. Jest to prawdopodobieństwo niewielkie, ale nie bardzo małe (większe od 0,05 – typowego poziomu istotności).

Następnie obliczmy prawdopodobieństwo tego, że wśród wylosowanych pięciu osób będzie czterech zwolenników partii Y:

$$P(Y=4) = \binom{5}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^1 = \frac{5}{32}.$$

Tyle samo wynosi prawdopodobieństwo otrzymania 4 zwolenników partii Z w losowej próbie pięcioelementowej.

Odpowiedzią na zadane na wstępie pytanie jest suma prawdopodobieństw:

- otrzymania w próbie pięciu zwolenników partii Y lub pięciu zwolenników partii Z,
- otrzymania w próbie czterech zwolenników partii Y lub czterech zwolenników partii Z:

$$2 \cdot \frac{1}{32} + 2 \cdot \frac{5}{32} = \frac{12}{32} = \frac{3}{8}.$$

Prawdopodobieństwo otrzymania takiej na pozór nietypowej próby nie jest wcale małe, bo wynosi 0,375. Jest ono kilkakrotnie większe od typowego poziomu istotności, dlatego nie upoważnia (w sensie statystycznym) do kwestionowania założenia o jednakowej liczbie zwolenników obu partii.

Literatura

1. Einstein A., *Pisma filozoficzne*, tłum. K. Napiórkowski, Instytut Filozofii i Socjologii PAN, Warszawa 1999.
2. Erikson R., Panagopoulos C., Wlezien Ch., *Likely (and Unlikely) Voters and the Assessment of Campaign Dynamics*, „Public Opinion Quarterly” 2004, Vol. 68.
3. Kowalczyk H., *O eksperckich ocenach niepewności w ankietach makroekonomicznych*, „Bank i Kredyt” 2010, nr 41.
4. Machina M.J., Schmeidler D., *A More Robust Definition of Subjective Probability*, „Econometrica” 1992, Vol. 60.
5. Młodinow L., *Matematyka niepewności. Jak przypadki wpływają na nasz los*, tłum. P. Strzelecki, Wydawnictwo Prószyński i S-ka, Warszawa 2011.
6. Motycka A., *Rozum i intuicja w nauce*, Wyd. ENETEIA, Warszawa 2005.
7. Newport F., *Yes, Polling Works*, „The Gallup Poll” z 11.11.2002.
8. Orman Quine W., van, *Od bodźca do nauki*, tłum. B. Stanosz, Wyd. Fundacja Aletheia, Warszawa 1998.
9. Szreder M., *Informacje a priori w klasycznej i bayesowskiej estymacji modeli regresji*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 1994.
10. Szreder M., *Prior Probabilities – the Pleasure of Bayesian Inference*, „Przegląd Statystyczny” 2002, nr 4 (t. 49).
11. Szreder M., *2010 – rok sondażowych nieporozumień*. „Rzeczpospolita” 3.01.2011.
12. Szreder M., *Metody i techniki sondażowych badań opinii*, wyd. II zm., PWE, Warszawa 2010.
13. Tversky A., Kahneman D., *Belief in the Law of Small Numbers*, „Psychological Bulletin” 1971, Vol. 76, No. 2.
14. Winkler R., *The Quantification of Judgment: Some Methodological Suggestions*, „Journal of the American Statistical Association” 1967, Vol. 62.

INTUITION AND EXPERIENCE IN STATISTICAL SURVEYS

Summary

The aim of this paper is to address and discuss some of the major problems connected with the need for using non-sample information in statistical surveys. Special attention is paid on researcher's experience and intuition which can be used to improve the quality of the survey's findings. Examples of opinion polls are given to illustrate such cases.

Keywords: sample survey, opinion poll, non-sample information.

Translated by Mirosław Szreder

