

Marcin Szymkowiak

Podjęcie kalibracyjne w badaniach społeczno-ekonomicznych

WYDAWNICTWO UEP



UNIWERSYTET
EKONOMICZNY
W POZNANIU

$D_4(d, w) = d^T G_4(g)$

Wstęp

W ostatnich latach obserwowany jest wzrost zapotrzebowania na informacje statystyczne dotyczące różnych zjawisk o charakterze społeczno-ekonomicznym. Popyt na aktualne i wiarygodne dane zgłaszany jest przez wiele instytucji na szczeblu państwowym, regionalnym i lokalnym. Są one niezbędne z punktu widzenia podejmowania odpowiednich decyzji przez władze rządowe i samorządowe, a także w kontekście monitorowania i oceny procesów społeczno-ekonomicznych w celu wzbogacenia wiedzy o otaczającej nas rzeczywistości.

Konieczność zaspokojenia potrzeb szerokiego grona odbiorców danych statystycznych wymusiła na organach statystyki publicznej dostarczanie informacji o wysokiej jakości, na którą składają się bardziej precyzyjne dane, o większym znaczeniu dla oceny analizowanych zjawisk oraz udostępniane na bieżąco (Witkowski, 2010). Statystyka publiczna, aby sprostać tym oczekiwaniom i wyzwaniom, podlega nieustannym przeobrażeniom, które obejmują wykorzystanie nowoczesnych technologii pozyskiwania i przetwarzania informacji oraz zaawansowanych metod badawczych, a także stosowanie danych pochodzących z różnych źródeł statystycznych. Dostrzegalne zmiany w zakresie organizacji badań statystycznych oznaczają konieczność wykorzystywania danych nie tylko ze spisów i badań reprezentacyjnych, ale w coraz większym zakresie z różnego rodzaju rejestrów i systemów administracyjnych. Większą rolę będą odgrywać również inne źródła w postaci internetu oraz tzw. big data (Beręsewicz i Szymkowiak, 2015; Beręsewicz, 2016; Daas i Puts, 2014; Daas, Puts, Buelens i van den Hurk, 2015; Szreder, 2015a; 2017; Zeelenberg, 2016).

Pomimo długiej tradycji wykorzystywania danych pochodzących ze spisów powszechnych, częstszego użycia rejestrów administracyjnych oraz innych źró-

deł danych, jak wspomniane dane z internetu czy big data, w dalszym ciągu jest oczywiste, że doniosłą rolę w statystyce publicznej odgrywają badania reprezentacyjne. Wynika to z kilku zasadniczych powodów. Po pierwsze dane pozyskiwane ze spisów powszechnych czy rejestrów administracyjnych pokrywają tylko pewien wycinek rzeczywistości. Istniejącą lukę w wielu innych obszarach, związanych z ubóstwem, rynkiem pracy, niepełnosprawnością, jakością życia, zdrowiem czy edukacją w dalszym ciągu wypełniają informacje pochodzące z odpowiednich badań reprezentacyjnych. Po drugie, jak zauważa Szreder (2016, s. 1–2), ważnymi przesłankami stosowania badań reprezentacyjnych są dostrzegalny postęp w rozwoju teorii próbkowania i wnioskowania statystycznego oraz powszechna świadomość niedostatków badań wyczerpujących. W nurt związany z ewolucją teorii estymacji wpisuje się również podejście kalibracyjne w rozumieniu zaprezentowanym w pracy Deville'a i Särndala (1992), któremu poświęcona jest niniejsza książka, a które w badaniach reprezentacyjnych odgrywa szczególną rolę. Kluczowym aspektem jest wreszcie relatywnie niski koszt ich realizacji, przetworzenia oraz udostępnienia danych.

Istotnym czynnikiem przemawiającym za wykorzystaniem badań reprezentacyjnych w procesie pozyskania odpowiedniej jakości danych jest także ich bogata historia i ugruntowana pozycja, wsparta przez odpowiednio wypracowaną metodologię. Jak podają Wright i Farmer (2000, s. iv), początków badań reprezentacyjnych można się doszukiwać już w XIX wieku, kiedy to wybitny matematyk francuski Pierre Simon de Laplace przekonał rząd francuski, aby na podstawie wylosowanej próby złożonej z okręgów administracyjnych zwanych gminami dokonać oszacowania liczebności populacji Francji. Wśród innych wczesnych zastosowań metody reprezentacyjnej, jak podają Rao i Fuller (2017, s. 146), należy wymienić pracę Kiaera (1897), który może być postrzegany jako pierwszy statystyk postulujący wykorzystanie badań częściowych opartych na próbie losowej kosztem badań wyczerpujących. Warto również wspomnieć o pracach Bowleya, który – jak podają Wright i Farmer (2000, s. iv) – wykorzystał losowanie systematyczne do oceny stopnia ubóstwa w mieście Reading w Anglii w 1912 roku.

W rozwoju metody reprezentacyjnej należy także uwzględnić pracę wybitnego polskiego matematyka i statystyka Jerzego Sławy-Neymana (1934), twórcy nowoczesnego pojęcia przedziału ufności, który wprowadził go do teorii testowania hipotez statystycznych i wytyczył kierunki rozwoju w obszarze badań reprezentacyjnych na dziesiątki lat. Wśród pozostałych, kluczowych opracowań, które odegrały fundamentalną rolę w rozwoju metody reprezentacyjnej i praktyce badań próbkowych, należy wymienić prace Cochran (1977), Deminga (1950), Hansena,

Hurvitza i Madowa (1953), Kisha (1965), a wśród polskich prace Brachy (1996), Pawłowskiego (1972), Steczkowskiego (1995), Wywiata (1999; 2010) czy Zasepy (1962; 1972).

W rozwoju badań reprezentacyjnych szczególnie ważną rolę odegrała także kalibracja, której podwaliny zostały stworzone przez Deville'a i Särndala (1992). Jest ona techniką powszechnie wykorzystywaną w praktyce badań reprezentacyjnych, której idea polega na takim skorygowaniu wag wynikających z przyjętego planu losowania próby, aby odtworzyć znane wartości globalne wszystkich zmiennych pomocniczych. Mimo że teoria podejścia kalibracyjnego została sformalizowana we wspomnianej już pracy Deville'a i Särndala (1992), jej początków należy upatrywać w metodzie rakingu¹ zaproponowanej przez Deminga i Stephana (1940).

W tym miejscu należy podkreślić, że w książce zarówno termin kalibracja, jak i podejście kalibracyjne będą używane zamiennie, mimo że to drugie w rozumieniu autora ma nieco szerszy kontekst. Podejście kalibracyjne (ang. *calibration approach*) będziemy definiować jako cały proces estymacji nieznanymi parametrami w populacji generalnej bądź w odpowiednio zdefiniowanych subpopulacjach (domenach), który obejmuje dobór zmiennych pomocniczych, kalibrację wag, statystyczną ocenę uzyskanych wag, konstrukcję odpowiednich estymatorów kalibracyjnych i ich wariancji oraz analizę otrzymanych wyników. W wąskim rozumieniu kalibracja jest zatem jednym z etapów podejścia kalibracyjnego i oznacza korektę wag, tak aby odtworzyć znane wartości globalne zmiennych pomocniczych. W szerszym ujęciu obejmuje również pozostałe etapy opisanego procesu estymacji z wykorzystaniem estymatorów kalibracyjnych.

Kalibracja nie jest jedyną techniką statystyczną stosowaną w konstruowaniu wag, które mogą być wzięte pod uwagę w procesie uogólniania wyników. Można bowiem tutaj zastosować inne metody wyznaczania wag, w tym m.in. regresję logistyczną, gdzie wagi są odwrotnością prawdopodobieństwa określanego jako skłonność respondenta do udzielenia odpowiedzi na pytania zawarte w kwestionariuszu ankiety (Kalton i Flores-Cervantes, 2003). W wypadku kalibracji uzyskujemy jednak wagi, które odtwarzają znane wartości globalne zmiennych pomocniczych, co jest niezwykle pożądaną własnością w praktycznych zastosowaniach. Co więcej, inne znane metody konstruowania wag, takie jak wspomniany już raking czy estymacja typu GREG², są szczególnymi przypadkami kalibracji.

¹ Metoda korygowania liczebności w odpowiedniej tablicy kontyngencji z wykorzystaniem algorytmu iteracyjnego proporcjonalnego dopasowania, której celem jest odtworzenie znanych wartości brzegowych. Więcej informacji na temat tej techniki statystycznej można znaleźć w podrozdziale 1.5.

² Będzie ona przedmiotem rozważań w podrozdziale 1.4.

Współcześnie kalibracja jest techniką stosowaną w badaniach reprezentacyjnych zarówno przez krajowe urzędy statystyczne, jak i ośrodki badania rynku. Jej zastosowania nie ograniczają się przy tym do badań próbkowych, ale dotyczą także spisów realizowanych metodą mieszaną (Szymkowiak, 2014), badań pełnych bazujących na rejestrach i systemach administracyjnych (A. Wallgren i B. Wallgren, 2014) czy badań pełnych obejmujących całą populację, w których poważnym problemem są odmowy (Klimanek i Szymkowiak, 2017).

Warto nadmienić, że kalibracja jako metoda korygowania wag nie tylko jest wykorzystywana na potrzeby odtworzenia znanych wartości globalnych zmiennych pomocniczych, co ma się przyczynić do zachowania zgodności z wynikami innych badań statystycznych, ale odgrywa szczególną rolę w takich obszarach badawczych jak problematyka braków odpowiedzi i błędów pokrycia (Särndal i Lundström, 2005) czy zagadnienie wartości odstających związane z występowaniem skrajnie asymetrycznych rozkładów cech (Duchesne, 1999). Najnowszy nurt w podejściu kalibracyjnym stanowią jej zastosowania w badaniach, w których próby mają nielosowy charakter (Chen i in., 2017; Chen, Valliant i Elliott, 2018a; 2018b). Wynika z tego, że chociaż metody oferowane przez teorię kalibracji są już szeroko omawiane w literaturze przedmiotu, to nadal zauważalna jest potrzeba ich dalszego rozwoju. Dotyczyć to może wspomnianych już prób nielosowych, a także konieczności jej zastosowania w spisach realizowanych metodą mieszaną czy w badaniach, w których poważnym problemem są występujące braki danych. Odpowiednia korekta wag ma nie tylko się przyczynić do zachowania spójności ze strukturami znanymi z innych badań, takich jak spisy powszechne czy rejestry, ale także ma docelowo wpłynąć na redukcję obciążenia i poprawę jakości precyzji oszacowań na skutek istniejących błędów nielosowych.

Celem niniejszej pracy jest przedstawienie podejścia kalibracyjnego w bardzo szerokim ujęciu, z uwzględnieniem aktualnego dorobku naukowego, doświadczeń będących pokłosiem zastosowań tej metody w praktyce badań statystycznych oraz autorskich rozwiązań i propozycji w kontekście rozwoju statystyki publicznej w Polsce. Ukazano w niej różnorodność zagadnień obejmujących wykorzystanie kalibracji w badaniach z brakami odpowiedzi czy w spisach realizowanych metodą mieszaną. Analizie poddano również tzw. podejście funkcyjne w zagadnieniu wyznaczania wag kalibracyjnych oraz wspomagane modelem, z uwzględnieniem najnowszych prac i rozwiązań metodologicznych w tym zakresie. Tak sformułowany cel ogólny opracowania składa się z następujących celów szczegółowych:

- dyskusji podstawowych pojęć z zakresu metody reprezentacyjnej i podejścia kalibracyjnego,

- wskazania ewolucji teorii kalibracji wraz z obszarem jej zastosowań w praktyce badań statystycznych,
- konstruowania estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości i ograniczeń nakładanych na wagi wraz z dyskusją najważniejszych ich własności,
- budowania estymatorów kalibracyjnych wartości globalnej w kontekście podejścia funkcyjnego wraz z omówieniem kluczowych ich własności,
- budowania estymatorów modelowo-kalibrowanych wartości globalnej, w tym z uwzględnieniem modeli regresji grzbietowej i adaptacyjnego modelu regresji LASSO oraz hybrydowych, wraz z dyskusją ich najważniejszych własności,
- próby odpowiedzi na pytanie o możliwość wykorzystania podejścia kalibracyjnego na niższym poziomie agregacji przestrzennej w badaniach z brakami danych na przykładzie Badania Aktywności Ekonomicznej Ludności (BAEL),
- próby odpowiedzi na pytanie o możliwość zastosowania kalibracji zintegrowanej w spisach realizowanych metodą mieszaną na przykładzie Narodowego Spisu Powszechnego Ludności i Mieszkań z 2011 roku (NSP 2011).

Realizacji celów podporządkowana została struktura książki. Rozdział pierwszy ma charakter wprowadzający. Sformułowano w nim niezbędne pojęcia z zakresu metody reprezentacyjnej, kluczowe z punktu widzenia rozważanej teorii kalibracji. Dokonano także studiów literaturowych w świetle aktualnego stanu wiedzy, wskazano na rozwój podejścia kalibracyjnego oraz przedstawiono obszary najważniejszych zastosowań kalibracji. Uwzględniono przy tym własny dorobek autora w zakresie wykorzystania omawianej metody w różnych badaniach statystycznych.

Rozdział drugi poświęcono przedstawieniu podejścia kalibracyjnego w bardziej zmatematyzowanym ujęciu. Dyskusji poddano sposób konstruowania wag kalibracyjnych oraz estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości. Zostały przy tym omówione szczegółowo ich najważniejsze zalety oraz wady.

W rozdziale trzecim przedstawiono z kolei za literaturą przedmiotu inny sposób konstruowania wag kalibracyjnych, które uzyskuje się w tzw. podejściu funkcyjnym wykorzystującym koncepcję zmiennych instrumentalnych.

Rozdział czwarty poświęcono estymatorom kalibracyjnym wartości globalnej wspomaganym odpowiednim modelem. Oprócz przeglądu najważniejszych estymatorów prezentowanych w literaturze, uwzględniających również podejście hybrydowe oraz wspomagane modelem regresji grzbietowej i adaptacyjnym mo-

delem regresji LASSO, w badaniu symulacyjnym zaproponowano wykorzystanie nowych estymatorów hybrydowo-kalibrowanych, które w swej konstrukcji wykorzystują wspomniane modele regresji grzbietowej i LASSO.

W rozdziale piątym uwaga została skupiona na możliwości wykorzystania podejścia kalibracyjnego w badaniach statystycznych z brakami danych. Od strony teoretycznej przedstawiono uogólnienie metody konstruowania estymatorów kalibracyjnych wartości globalnej w podejściu jednokrokowym oraz dwukrokowym typu A i B rozważanej przez Särndala i Lundströma (2005) na przypadek dowolnie zdefiniowanej funkcji odległości. Zaproponowano również wykorzystanie podejścia kalibracyjnego w obszarze rynku pracy w przekroju podregionów, tj. na poziomie, dla którego brakuje oficjalnych szacunków publikowanych przez Główny Urząd Statystyczny. O jego znaczeniu niech świadczy to, że wśród podregionów mamy największe miasta w Polsce, dla których zawsze lokalni politycy i działacze gospodarczy dopominają się o informacje z zakresu rynku pracy. Zastosowanie i specyfikację związanych z tym problemów badawczych omawianej metody przedstawiono na danych z Badania Aktywności Ekonomicznej Ludności, w którym jednym z kluczowych problemów są występujące braki odpowiedzi.

Ostatni, szósty rozdział pracy zawiera rozważania dotyczące możliwości zastosowania kalibracji w spisach powszechnych realizowanych metodą mieszaną. Ten typ spisu był wykorzystywany w rundzie spisowej 2010–2011 w wielu państwach europejskich, w tym w Polsce (Gołata, 2018; Valente, 2010). W pracy przedstawiono sposób konstruowania wag kalibracyjnych w kontekście ludności faktycznie zamieszkałej zgodnie z wypracowaną w 2011 roku w spisie koncepcją korygowania wag. Wagi te zostały wyznaczone w Ośrodku Statystyki Małych Obszarów Urzędu Statystycznego w Poznaniu, a autor książki kierował pracami grupy metodologów odpowiedzialnych za proces kalibracji. W książce przedstawiono również propozycję konstruowania tzw. wag zintegrowanych w spisach mieszanych na przykładzie spisu powszechnego z 2011 roku. Należy podkreślić, że tego typu rozwiązanie nie zostało zastosowane w tym badaniu. Wyniki przeprowadzonej analizy wskazują, że podejście zintegrowane może nabrać szczególnego znaczenia w kontekście kolejnego spisu, który zostanie przeprowadzony w 2021 roku zgodnie z mieszanym modelem realizacyjnym. Na zakończenie dyskusji poświęconej zawartości każdego z rozdziałów warto zaznaczyć, że kończą się one podsumowaniami, w których ujęto najważniejsze wnioski.

Uwzględniając sformułowane cele oraz strukturę opracowania, należy podkreślić, że prezentowana praca ma charakter teoretyczny, empiryczny i aplikacyjny. W warstwie teoretycznej dokonano matematycznego opisu estymatorów kalibracyjnych wartości globalnej w ujęciu klasycznym, w podejściu funkcyjnym, wspo-

magany modelem, w badaniach z brakami odpowiedzi oraz w obszarze tzw. kalibracji zintegrowanej. Niektóre z rozważanych podejść miały przy tym autorski charakter, co należy postrzegać w charakterze novum metodologicznego.

W warstwie empirycznej głównym celem była analiza najważniejszych własności rozważanych estymatorów kalibracyjnych z uwzględnieniem oceny uzyskanych wag kalibracyjnych. Realizacji tego celu służyły odpowiednio zaprojektowane badania symulacyjne przedstawione w rozdziałach 2–4, w których wykorzystano rzeczywiste dane pochodzące z Badania Budżetów Gospodarstw Domowych (BBGD). Warto podkreślić, że w rozdziałach tych zastosowane zostały własne kody napisane w języku R (R Core Team, 2018), w których centralną rolę odgrywały pakiety **survey** (Lumley, 2004; 2017), **glmnet** (Friedman, Hastie i Tibshirani, 2010) oraz zmodyfikowana wersja funkcji *calibWeights* z pakietu **laeken** (Alfons i Templ, 2013). Modyfikacja polegała na dodaniu funkcji odległości w postaci sinusa hiperbolicznego i zastosowaniu pakietu do analizy macierzy rzadkich **Matrix** (Bates i Maechler, 2018) w celu przyspieszenia czasochłonnych obliczeń.

W warstwie aplikacyjnej dokonano zastosowania omawianych metod kalibracji w Badaniu Aktywności Ekonomicznej Ludności oraz w spisie powszechnym z 2011 roku, który został przeprowadzony metodą mieszaną. W tym celu ponownie skorzystano ze wspomnianych powyżej pakietów i funkcji języka R. Jedyne na potrzeby oceny precyzji oszacowań dla danych spisowych z wykorzystaniem otrzymanych wag zintegrowanych użyto programu SAS i procedury *survey-means*. Egzemplifikacja podejścia kalibracyjnego w bardzo ważnych badaniach statystycznych stanowi niewątpliwe novum poznawcze.

Książka powstała w ramach grantu naukowego: Estymacja pośrednia w zakresie badania niepełnosprawności na podstawie NSP 2011. Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/B/HS4/01472.