Biblioteka
WIADOMOŚCI STATYSTYCZNYCH

PROBLEMY BADAŃ STATYSTYCZNYCH METODĄ REPREZENTACYJNĄ

Tom 36 WARSZAWA 1989

POLSKIE TOWARZYSTWO STATYSTYCZNE
KOMITET STATYSTYKI I EKONOMETRII PAN

PROBLEMY BADAŃ STATYSTYCZNYCH METODĄ REPREZENTACYJNĄ

KONFERENCJA NAUKOWA
ZWARTOWO, 26—28 WRZEŚNIA 1988 R.

GŁÓWNY URZĄD STATYSTYCZNY

Opracowanie merytoryczne
prof. dr hab. Ryszard Zasępa

Redakcja techniczna
mgr Mirosława Zagrodzka

Okładkę projektował
mgr Krzysztof Dobrowolski

Zakład Wydawnictw Statystycznych. 00-925 Warszawa, al. Niepodległości 208.
Nakład 1300 egz, ark. druk. 12,5. Papier druk. kl. V 70 g, 61 × 86.
Przekazano do składu w kwietniu 1989 r. Druk ukończono w sierpniu 1989 r.
Cena 1600,— zł
Druk ZWS, Warszawa, zam. 152. St-2.
Informacje w sprawach zbytu publikacji — tel. 25-48-86.

OD REDAKCJI

W prezentowanym tomie 36 Biblioteki Wiadomości Statystycznych publikowane są referaty wygłoszone na Konferencji Naukowej, zorganizowanej przez Polskie Towarzystwo Statystyczne oraz Komitet Statystyki i Ekonometrii PAN, która odbyła się w Zwartowie w dniach 26–28 września 1988 r.

W przedstawionych referatach omówione zostały ważniejsze zagadnienia metodologii badań przeprowadzanych metodą reprezentacyjną. Tematyka referatów jest z konieczności dość różnorodna – począwszy od teorii, problemów estymacji do praktycznego zastosowania metody reprezentacyjnej w badaniach. Tym niemniej Redakcja Wiadomości Statystycznych, biorąc pod uwagę kryterium tematyczne zgromadzonych materiałów, starała się zaprezentować referaty w zestawach problemowych.

Pierwszy zestaw – to referaty poświęcone teorii metody reprezentacyjnej.

Drugi zestaw – to problematyka dotycząca estymacji w badaniach reprezentacyjnych.

Trzeci zestaw zagadnień – to konkretne doświadczenia ze stosowania metody reprezentacyjnej w badaniach gospodarstw domowych, gospodarstw rolniczych i wielu badań ankietowych.

Z konieczności Redakcja ograniczyła zawartość tomu do treści wygłoszonych referatów, nie podając wypowiedzi uczestników konferencji w ożywionej dyskusji.

Redakcja żywi nadzieję, że prezentowane w tomie 36 referaty z konferencji przybliżą czytelnikom, statystykom i użytkownikom badań metodę reprezentacyjną, a tym samym przyczynią się do dalszego jej rozwoju i szerokiego stosowania w badaniach statystycznych.

SPIS TREŚCI

Str.

Od Redakcji 3

Słowo wstępne 6

Lista obecności 9

Teoria metody reprezentacyjnej

Ryszard Zasępa — Teoria badań reprezentacyjnych i jej zastosowanie w praktyce statystycznej 10

Jan Steczkowski — Kilka myśli o metodzie reprezentacyjnej w badaniach zjawisk społeczno-ekonomicznych 20

Andrzej Balicki — Wybrane problemy badań reprezentacyjnych w światowym piśmiennictwie statystycznym ostatnich lat . . 30

Józef Bielecki — Zagadnienie struktury badań reprezentacyjnych w Zintegrowanym Systemie Badań Gospodarstw Domowych . 50

Czesław Bracha — Stosowanie wybranych testów statystycznych do danych z prób nieprostych 66

Problemy estymacji w badaniach reprezentacyjnych

Richard Platek — Niekompletne dane, ich korekta i efekty . . 77

Jan Kordos — Konieczność uwzględnienia błędów losowych i nielosowych w planowaniu badań reprezentacyjnych i wykorzystaniu ich wyników 98

Bogdan Stefanowicz — Wybrane zagadnienia jakości informacji statystycznych 110

Janusz Wywiał — Współczynniki rzetelności danych wielowymiarowych 120

Zofia Mielecka Kubień — Możliwość zastosowania badań reprezentacyjnych do oceny stopnia rozprzestrzeniania się alkoholizmu w Polsce 127
Henryk Hillar — Błędy badania reprezentacyjnego 146
Jan Paradysz — O błędach nielosowych w badaniu dzietności kobiet w ramach Narodowego Spisu Powszechnego 1970 154

Metoda reprezentacyjna w praktyce statystycznej

Antoni Kubiczek — Zagadnienia braku odpowiedzi w badaniach gospodarstw domowych 160
Wiesław Łagodziński — Jednorazowe badania ankietowe w Zintegrowanym Systemie Badań Gospodarstw Domowych . . . 170
Stanisława Ryckowska — Problematyka badania budżetów gospodarstw domowych metodą rotacyjną oraz innych badań społecznych w świetle doświadczeń Wojewódzkiego Urzędu Statystycznego w Gdańsku 179
Florian Ruszkowski — Zastosowanie metody reprezentacyjnej w statystyce rolniczej 186

Słowo końcowe 197

SŁOWO WSTĘPNE

W imieniu prof. dra Wiesława Sadowskiego, prezesa GUS, przewodniczącego Komitetu Statystyki i Ekonometrii PAN oraz Rady Głównej Polskiego Towarzystwa Statystycznego pragnę przekazać uczestnikom obecnej konferencji naukowej pozdrowienia i życzenia owocnych obrad. Jest to już czwarta konferencja naukowa, której patronuje Polskie Towarzystwo Statystyczne.

Pierwsza konferencja nt. organizacji badań statystycznych odbyła się w 1985 r. w Gdańsku, druga — poświęcona jakości informacji statystycznych odbyła się w 1986 r. w Szczyrku. W 1987 r. obchodziliśmy 75-lecie powstania Polskiego Towarzystwa Statystycznego, które połączone zostało z konferencją naukową na temat spisów ludności. Dzisiejsza konferencja, poświęcona problematyce badań reprezentacyjnych, dotyczy ważnego działu metodologii badań statystycznych.

Wyniki obecnej konferencji naukowej, tj. przygotowane referaty, zostaną opublikowane w wydawnictwie GUS w serii „Biblioteka Wiadomości Statystycznych".

W słowie wstępnym do naszych obrad pragnę podzielić się kilkoma refleksjami na temat zastosowań metody reprezentacyjnej w badaniach statystycznych. Warto bowiem pamiętać o pewnych faktach z przeszłości, aby można było je wykorzystać w naszych kolejnych działaniach. Szkoda, że na dzisiejszą konferencję nie przygotowano krytycznej oceny zastosowania metody reprezentacyjnej w dotychczasowych badaniach statystycznych lub nie dokonano analizy tych badań. Nie dokonano krytycznej oceny żadnej publikacji wydanej w ostatnich latach z dziedziny badań reprezentacyjnych. Wydaje się celowe przypomnienie pewnych faktów, dotyczących badań reprezentacyjnych. Pragnę zwrócić uwagę na dwa wydarzenia, które zaważyły w znacznym stopniu na zastosowaniu metody reprezentacyjnej w badaniach statystycznych, a mianowicie na działalność Komisji Matematycznej GUS oraz seminarium międzynarodowe, poświęcone badaniom reprezentacyjnym, które odbyło się w 1979 r. w Warszawie.

DZIAŁALNOŚĆ KOMISJI MATEMATYCZNEJ GUS

Komisja Matematyczna GUS (nazwa jej w różnych okresach ulegała pewnym zmianom) powstała w 1949 r. pod przewodnictwem prof. Stefana Szulca, ówczesnego prezesa GUS. Znaczną część swej działalności poświęciła ona zastosowaniom metody reprezentacyjnej w badaniach statystycznych. Zajmowała się również szkoleniem pracowników GUS z zakresu teorii badań reprezentacyjnych, przygotowaniem odpowiednich pomocy naukowych, a także pracami naukowo-badawczymi w tej dziedzinie. Wiele z tych prac zostało opublikowanych w postaci artykułów lub opracowań zwartych. Warto przypomnieć kolejnych przewodniczących Komisji Matematycznej GUS, a mianowicie: prof. Marka Fisza, autora podręcznika pt. „Rachunek prawdopodobieństwa i statystyka matematyczna"; prof. Ryszarda Zasępę obecnego na dzisiejszej konferencji, autora dwóch podręczników z teorii metody reprezentacyjnej i wielu artykułów, nestora polskich badań reprezentacyjnych i wieloletniego przewodniczącego Komisji Matematycznej; prof. Zbigniewa Pawłowskiego, autora podręcznika metody reprezentacyjnej pod skromnym tytułem „Wstęp do statystycznej metody reprezentacyjnej"; w popularny sposób przybliżył on zagadnienia badań reprezentacyjnych oraz prof. Wiesława Welfego, który kierował całokształtem zagadnień związanych z zastosowaniem metod matematycznych w statystyce. Należy również wymienić prof. Jerzego Grenia, zmarłego przed trzema laty, który prowadził intensywne prace badawcze z zakresu teorii i praktyki badań reprezentacyjnych. Był autorem kilku interesujących artykułów o metodzie reprezentacyjnej oraz wielu zastosowań tej metody w praktyce.

Komisja Matematyczna zakończyła swoją działalność na początku lat osiemdziesiątych, a prace prowadzone przez nią kontynuuje Zakład Badań Statystyczno-Ekonomicznych GUS i PAN.

MIĘDZYNARODOWE SEMINARIUM POŚWIĘCONE BADANIOM REPREZENTACYJNYM

Zwołanie międzynarodowego seminarium było wynikiem aktywnej działalności Komisji Matematycznej GUS w latach sześćdziesiątych. Prof. dr hab. Wincenty Kawalec, ówczesny prezes GUS, z inicjatywy tejże Komisji zgłosił na posiedzeniu Stałej Komisji Statystycznej RWPG temat: *Zbadanie możliwości szerszego zastosowania metody reprezentacyjnej w badaniach statystycznych krajów-członków RWPG.*

Kierowałem wówczas pracami grupy przygotowującej ten temat i miałem możliwość zapoznania się z dorobkiem zastosowania metody reprezentacyjnej w poszczególnych krajach socjalistycznych. Seminarium odbyło się w Warszawie w dniach 21—24 kwietnia 1970 r. Wyniki seminarium zostały opublikowane w dwóch tomach „Biblioteki Wiadomości Statystycznych":

— *Badania statystyczne metodą reprezentacyjną w krajach socjalistycznych*, tom 14, Warszawa 1971, s. 220,

— *Wybrane problemy metodologiczne badań reprezentacyjnych*, tom 15, Warszawa 1971, s. 152.

Publikacje te dostarczyły bogatych materiałów do studiowania teorii badań reprezentacyjnych i jej zastosowania w badaniach statystycznych. Były tłumaczone i wykorzystywane w różnych krajach. Warto tu przypomnieć niektóre wnioski z omawianego Seminarium, które wynikały ze zgłoszonych referatów i bogatej dyskusji:

- Centralne urzędy statystyczne krajów-członków RWPG powinny posiadać specjalne stałe oddziały, zajmujące się badaniami reprezentacyjnymi.
- Pracowników terenowych organów statystyki należy na specjalnych kursach szkoleniowych zapoznać z teorią metody reprezentacyjnej.
- W celu zwiększenia efektywności badań, prowadzonych metodą reprezentacyjną, konieczne jest oparcie ich na wynikach odpowiednich prac badawczych.
- Istnieje konieczność rozszerzenia badań eksperymentalnych.
- Zachodzi konieczność wymiany doświadczeń w zakresie badań reprezentacyjnych między różnymi krajami.
- Wśród użytkowników danych statystycznych należy popularyzować badania reprezentacyjne.

Każdy z podanych wniosków został szczególnie uzasadniony i udokumentowany. Nie wszystkie z nich zostały w pełni zrealizowane, ale niewątpliwie miały znaczący wpływ na badania statystyczne w urzędach statystycznych krajów socjalistycznych.

Po Seminarium kontynuowano wiele badań statystycznych metodą reprezentacyjną w poszczególnych krajach socjalistycznych. Z postępami w tej dziedzinie mogłem się zapoznać z racji uczestnictwa w konferencjach międzynarodowych, które dotyczyły głównie badań gospodarstw domowych, a te należą niewątpliwie do najtrudniejszych badań statystycznych.

Pomimo trudności dokonano w tej dziedzinie znacznego postępu, ale nie jest on znany szerszemu ogółowi statystyków i użytkowników tych badań. Dlatego wydaje się celowe podjęcie próby syntezy dotychczasowych zastosowań metody reprezentacyjnej w badaniach społeczno-ekonomicznych, prowadzonych przez urzędy statystyczne poszczególnych krajów. Taka publikacja monograficzna przyczyniłaby się niewątpliwie do dalszego postępu badań w tej dziedzinie.

Sądzić należy, że dzisiejsza konferencja naukowa przybliży szerszemu ogółowi statystyków i użytkowników badań statystycznych problematykę zastosowań metody reprezentacyjnej w praktyce statystycznej, a dyskusja poruszy ważkie problemy, dla dalszego rozwoju badań.

PREZES POLSKIEGO
TOWARZYSTWA STATYSTYCZNEGO

doc. dr hab. Jan Kordos

LISTA OBECNOŚCI

Lista uczestników konferencji: „Problemy badań statystycznych metodą reprezentacyjną" w Zwartowie, 26—28 IX 1988 r.

1. Doc. dr hab. Balicki Andrzej, Uniwersytet Gdański
2. Mgr Bahr Iwona, WUS Przemyśl
3. Dr Barańska Zygmunta, Uniwersytet Gdański
4. Dr Bielecki Józef, Uniwersytet Gdański
5. Doc. dr hab. Bracha Czesław, SGPiS Warszawa
6. Mgr Fedak Roman, WUS Zielona Góra
7. Dr Godziniec Romuald, Uniwersytet Gdański
8. Doc. dr hab. Hillar Henryk, Uniwersytet Warszawski
9. Mgr Hryniewiecka Izabela, Uniwersytet Gdański
10. Dr Jurkiewicz Bogdan, Uniwersytet Gdański
11. Mgr Kasprowicz Armand, Uniwersytet im. M. Kopernika, Toruń
12. Doc. dr hab. Kordos Jan, GUS Warszawa
13. Dr Karol Mirosława, Uniwersytet Szczeciński
14. Prof. dr hab. Krzysztofiak Mirosław, Uniwersytet Gdański
15. Mgr Lasota Barbara, Uniwersytet Szczeciński
16. Mgr Lednicki Bronisław, ZBSE GUS i PAN, Warszawa
17. Mgr Łagodziński Wiesław, GUS Warszawa
18. Dr Mielecka-Kubień Zofia, Akademia Ekonomiczna, Katowice
19. Mgr Mindak Jolanta, Uniwersytet im. M. Kopernika, Toruń
20. Dr Muzalewski Jerzy, Akademia Ekonomiczna, Poznań
21. Dr Niewczas Ryszard, Uniwersytet im. M. Kopernika, Toruń
22. Doc. dr hab. Paradysz Jan, Akademia Ekonomiczna, Poznań
23. Mgr Pietrzak Zbigniew, Uniwersytet Gdański i WUS Gdańsk
24. Mgr Plenikowska-Ślusarz Teresa, Uniwersytet Gdański
25. Mgr Ruszkowski Florian, GUS Warszawa
26. Mgr Ryckowska Stanisława, WUS Gdańsk
27. Prof. dr hab. Steczkowski Jan, Akademia Ekonomiczna Kraków
28. Doc. dr hab. Stefanowicz Bogdan, SGPiS Warszawa
29. Dr Szreder Mirosław, Uniwersytet Gdański
30. Doc. dr Urbanek-Krzysztofiak Danuta, Uniwersytet Gdański
31. Doc. dr Wierzchosławski Stanisław, Akademia Ekonomiczna, Poznań
32. Dr Wywiał Janusz, Akademia Ekonomiczna, Katowice
33. Prof. dr hab. Zasępa Ryszard, ZBSE GUS i PAN, Warszawa
34. Prof. dr hab. Zawada Eugeniusz, Akademia Spraw Wewnętrznych, Warszawa
35. Mgr Zeb Elżbieta, WUS Kielce

TEORIA METODY REPREZENTACYJNEJ

Teoria badań reprezentacyjnych i jej zastosowanie w praktyce statystycznej

prof. dr hab. Ryszard Zasępa
Zakład Badań Statystyczno-Ekonomicznych GUS i PAN
wiceprezes Polskiego Towarzystwa Statystycznego

1. Pięćdziesiąt cztery lata temu Jerzy Neyman wygłosił w Królewskim Towarzystwie Statystycznym Wielkiej Brytanii słynny referat [7], przedstawiając w nim nowe idee teorii i praktyki badań reprezentacyjnych, zastosowane we wcześniej zaprojektowanym przez niego badaniu reprezentacyjnym ludności robotniczej [6]. Źródłem badania były materiały drugiego Powszechnego Spisu Ludności w Polsce (1931). Przedtem czołowi statystycy uważali (czego dowodem jest opublikowane memorandum A. L. Bowleya w 22 biuletynie Międzynarodowego Instytutu Statystycznego), że losowy wybór próby z populacji skończonej to wybór indywidualny z jednakowymi prawdopodobieństwami, dający w efekcie próbę prostą. Wobec częstych trudności w możliwości dokonania takiego wyboru w praktyce, za racjonalne rozwiązanie uznano wybór celowy grup jednostek badania. Moim zdaniem, pod wpływem tych opinii włoscy statystycy C. Gini i L. Galvani w 1926 r. zastosowali wybór celowy materiałów włoskiego Spisu Powszechnego z 1921 r. do uzyskania reprezentacyjnej próby dla całego kraju, pod względem rozkładów podstawowych cech demograficznych. Wybór celowy próby zastosował także w 1934 r. O. Anderson w Bułgarii.

Wysunięte idee J. Neymana były następujące:

a) nie ma potrzeby ograniczania pojęcia próby losowej do próby prostej. Losową próbę uzyskamy również stosując wybór zespołów nie tylko z całej populacji, lecz także w drodze wyboru warstwowego. Przy takim wyborze można konstruować, przy odpowiednio wysokim współczynniku ufności, przedziały ufności dla szacowanej wielkości;
b) do oceny jakiejś charakterystyki populacji (w rozważaniach chodziło o średnie ważone), posługując się metodą najmniejszych kwadratów A. A.

Markowa, można określić estymator najlepszy, tj. najmniejszej wariancji z klasy nieobciążonych estymatorów liniowych. Taki estymator dostarcza najkrótsze przedziały ufności;

c) porównując wariancję najlepszego estymatora liniowego średniej wartości badanej cechy okazuje się, że losowanie bez zwracania jest efektywniejsze niż losowanie ze zwracaniem oraz losowanie warstwowe jest z reguły efektywniejsze niż losowanie nieograniczone;

d) w szacowaniu średniej, rozdział próby pomiędzy warstwy jest optymalny, gdy liczba jednostek losowania (*jl*) z warstwy jest proporcjonalna do iloczynu z liczby *jl* w warstwie, przez odchylenie standardowe badanej cechy. Taki rozdział próby będzie również dobry, gdy szacujemy średnie cech ze sobą skorelowanych. Jeśli nie ma możliwości oceny odchylenia standardowego w warstwach, wówczas najlepszym rozwiązaniem jest schemat, w którym liczba *jl* z warstwy jest proporcjonalna do wielkości warstwy (mierzonej liczbą *jl*);

e) analizując wyniki badania Giniego i Galvaniego dochodzi się do wniosku, że w przypadku dużej próby (ten warunek jest zwykle spełniany w badaniach reprezentacyjnych) wybór losowy próby jest lepszy niż wybór celowy.

2. Przedstawiona przez Neymana nowa teoria wnioskowania wywołała początkowo zastrzeżenia, czego dowodem są wypowiedzi Bowleya, Isserlisa, Fishera w czasie dyskusji nad jego referatem. Jednak, w okresie około dziesięciu lat, idee Neymana uzyskały powszechne uznanie i zaowocowały rozwojem teorii metody reprezentacyjnej. Chociaż Neyman ograniczył się w swoich rozważaniach do losowania jednostopniowego i z jednakowymi prawdopodobieństwami wyboru, to z prezentowanych przez niego zasad wynikały logicznie możliwości rozszerzenia schematów losowania próby. Sam Neyman (1938)[1]) wprowadził dwufazowy schemat losowania. M. H. Hansen i W. N. Hurwitz (1943) wprowadzili losowanie dwustopniowe oraz losowanie z różnymi prawdopodobieństwami wyboru (ze zwracaniem). W. G. i L. H. Madowowie (1944) rozpoczęli rozwój teorii losowania systematycznego. Głównie dzięki W. G. Cochranowi (1942) rozpoczęły się badania estymatorów ilorazowych i regresyjnych. W Indii P.C. Mahalanobis (1944, 1946) wysunął ideę podprób oraz technik zmniejszania obciążenia estymatora metodą „obcinania". H. D. Patterson (1950) opracował teorię planowania i analizy badań powtarzalnych. D. G. Horvitz i D. J. Thompson (1952) zbudowali ogólną teorię losowania z różnymi prawdopodobieństwami bez zwracania. Nieco wcześniej W. G. Madow (1949) zaproponował wykorzystanie losowania systematycznego z różnymi prawdopodobieństwami, w celu uniknięcia możliwości wylosowania wielokrotnego tej samej

[1]) W celu ograniczenia wymiaru referatu bibliografia została w nim ograniczona. Odpowiednie pozycje literatury można znaleźć np. w [1], [2], [3], [8] lub [9].

jednostki. To spowodowało dużą liczbę innych procedur losowania, zapoczątkowanych propozycją R. D. Naraina (1951). W przeglądzie M. Hanifa i K. R. W. Brewera (1980), opublikowanym w International Statistical Review 48, wyróżniono pięćdziesiąt schematów wysuniętych, do tego czasu, przez różnych statystyków.

3. Wyniki studium empirycznego, we wspomnianej rozprawie Horvitza i Thompsona (1952), stały się stymulatorem badań teoretycznych, uściślających dotychczasowe pojęcia teorii metody reprezentacyjnej, nawiązujących do tradycyjnych pojęć statystyki matematycznej. Pionierem tych studiów jest V. P. Godambe (1955). Przede wszystkim sprecyzowano definicje podstawowych pojęć. W badaniu reprezentacyjnym populacja badana jest skończonym zbiorem N **identyfikowalnych obiektów,** które możemy uporządkować liniowo. Zastępując adresy (identyfikatory) poszczególnych obiektów kolejnymi numerami, w ustalonym uporządkowaniu, populację możemy krótko utożsamić ze zbiorem numerów $U=\{1, 2, ..., N\}$. Badając cechę Y z każdym numerem $k\in U$ związana jest pewna liczba Y_k oznaczająca wartość tej cechy dla danego obiektu. Populację ze względu na cechę Y określa pewien wektor $\boldsymbol{Y}=[Y_1, Y_2, ... Y_N]$, którego współrzędnych nie znamy, ale wiemy, że $\boldsymbol{Y}$ należy do pewnej podprzestrzeni $\Omega c R_N$ przestrzeni euklidesowej. Rozkład cechy Y zależy od współrzędnych wektora $\boldsymbol{Y}$, który pełni analogiczną rolę, jak parametr w rozkładzie ciągłym i jest wobec tego nazywany **parametrem** (populacyjnym), będącym przedmiotem szacunku. W praktyce badań reprezentacyjnych jesteśmy zainteresowani oceną pewnej **funkcji parametrycznej** $T(\boldsymbol{Y})$, np. wartości globalnej cechy, $\hat{Y}=\sum_k^N Y_k$, bądź wartości średniej $\bar{Y}=N^{-1}\hat{Y}$. Jakikolwiek podzbiór $s\in U$, $s=\{k_1, k_2, ... k_{n(s)}\}$ nazywamy próbą o wymiarze $n(s)$[2]). Zbiór wszystkich możliwych prób s oznaczmy przez S. Miarę prawdopodobieństwa p: $S\to[0,1]$ taką, że $\sum_{s\in S} p(s)=1$ nazywamy **planem**[3]) **losowania (wyboru probabilistycznego)** (ang. sampling design). Mechanizm losowy, realizujący dany plan losowania nazywamy **schematem losowania**. Badanie próby s dostarcza informacji $d=[(k, Y_k) : k\in s]$. Dane „k" identyfikujące obiekty populacji, należące do próby mogą być w badaniu zaniechane i wówczas $d=(Y_k : k\in s)$. Estymatorem funkcji parametrycznej $T(\boldsymbol{Y})$ jest pewna funkcja rzeczywista $t : S\times U\to R_1$, która zależy od wektora $\boldsymbol{Y}$ jedynie przez składowe należące do próby. Badania teoretyczne koncentrują się na zagadnieniu

---

[2]) To pojęcie próby można rozszerzyć biorąc pod uwagę kolejności (i) losowania do próby obiektów (k_i), definiując próbę jako ciąg $s=(k_1, k_2, ... k_{n(s)})$ taki, że $k_i\in U$ dla $i=1, 2, ... n(s)$; próbę nazwiemy wówczas „uporządkowaną", w przeciwieństwie do wyżej określonej, jako „nie uporządkowanej".

[3]) Inaczej „nie uporządkowanym planem ..." przyjmując „uporządkowany plan ..." w przypadku definicji z notki 2).

szacowania średniej (lub globalnej) wartości cechy Y oraz na estymacji liniowej. Estymator t średniej Y jest liniowy, jeśli:

$$t = b_0(s) + \sum_{k \in s} b(s,k)\, Y_k \tag{1}$$

gdzie: — b_0, b są funkcjami rzeczywistymi, odpowiednio, na S oraz na $S \times U$, nie zależą od wartości cechy Y, ale mogą zależeć od dodatkowej informacji.

Klasa L estymatorów liniowych określonych w (1) jest znacznie szersza od klasy rozważanej przez Neymana dla szacowania średnich. W (1) uwzględnia się identyfikowalność obiektów populacji skończonej i możliwość wykorzystania identyfikatorów przy estymacji. Neyman brał pod uwagę identyfikowalność tylko dla warstw. W praktyce wykorzystuje się identyfikatory w przypadku losowania z różnymi prawdopodobieństwami wyboru, jak to ma miejsce dla estymatora średniej t_{HH} Hansensa-Hurwitza (1943) czy — dla dowolnego planu próbkowania, w przypadku estymatora t_{HT} Horvitza-Thompsona (1952). Ograniczając się do klasy L_{0u} nieobciążonych estymatorów liniowych jednorodnych ($b_0(s) = 0$) Godambe (1955) udowodnił, że w tej klasie **nie istnieje** estymator średniej, w pewnym sensie najlepszy, mianowicie **estymator o jednostajnie najmniejszej wariancji** (UMV)[4]). Godambe i Joshi (1965) wykazali również nieistnienie estymatora UMV dla klasy A_u nieobciążonych estymatorów średniej[5]), wykorzystujących identyfikatory. Przypomnijmy, że w teorii Neymana, który nie brał pod uwagę identyfikatorów wewnątrz warstwy, istniał estymator UMV średniej. Poszukując szerszej klasy estymatorów, dla której istnieje estymator UMV (Watson, 1964; Royall, 1968; Särndal, 1972, 1976) udowodniono następujące:

Twierdzenie 1. Dla uporządkowanych planów próbkowania, spełniających warunki:

(*i*) $p(s) > 0 \Rightarrow s$ jest ciągiem n różnych identyfikatorów k_i, $1 \leqslant k_i \leqslant N$, $i = 1, 2, \ldots, n$,

(*ii*) wszystkie $n!$ ciągów s będących permutacjami tych samych identyfikatorów mają to samo prawdopodobieństwo,

(*iii*) prawdopodobieństwo α_k wylosowania do próby k-tego obiektu, $\alpha_k > 0$ dla $k = 1, 2, \ldots, N$.

Oznaczając $Z_k = n(Y_k - e_k)/N\alpha_k$ ($k = 1, 2, \ldots, N$), uogólniony estymator różnicowy t_{GD},

$$t_{GD} = \sum_{s \in S} \frac{Y_k - e_k}{N\alpha_k} + \bar{e} = \bar{Z}_s + \bar{e}$$ (gdzie $e \in R_N$ jest dowolnym stałym wektorem,

---

[4]) Termin „jednostajnie" oznacza „dla wszelkich $Y \in R_N$."

[5]) W dalszym ciągu referatu będę używał terminu „estymator" w odniesieniu do „estymatora średniej", gdyż prace teoretyczne dotyczą tego przypadku (bądź co jest równoważne, estymatora wartości globalnej cechy).

$\bar{e} = N^{-1} \sum_k e_k$) jest estymatorem UMV w klasie wszystkich nieobciążonych estymatorów średniej $\bar{Y}$ zależnych od danych próby jedynie przez ciąg nieidentyfikowanych wielkości $(Z_{k_1}, Z_{k_2}, ..., Z_{k_n})$.

Z (ii) wynika, że estymatory pomijają porządek, w jakim są losowane do próby poszczególne obiekty. Estymatory należą do klasy, w której identyfikatory uwzględniane są częściowo. W praktyce będzie to miało miejsce, gdy identyfikator k nie wnosi dodatkowych elementów do statystycznego wnioskowania, o ile tylko znamy Z_{k_i} wchodzące do próby.

4. Dążąc do maksymalnej redukcji informacji z próby bez utraty istotnej jej części do wnioskowania statystycznego, wprowadzono w statystyce matematycznej pojęcie **dostateczności** statystyki, ograniczając poszukiwanie „dobrych" estymatorów do funkcji takiej statystyki. Badanie tej koncepcji, w przypadku metody reprezentycyjnej, prowadził głównie D. Basu (1958, 1969). Wykazał, że statystykę dostateczną otrzymamy usuwając z danych próby informacje o porządku w jakim wylosowano do próby poszczególne obiekty populacji oraz o krotności ich w próbie. Uzyskamy wówczas, jak się okazuje, nawet minimalną statystykę dostateczną. Zatem losowanie ze zwracaniem nie daje o średniej żadnej dodatkowej informacji. Ten rezultat jest intuicyjnie oczywisty, chociaż w podręcznikach metody reprezentacyjnej rozpatruje się średnią z wszystkich obserwacji (z powtórzeniami), a nie średnią $\bar{Y}_d$ po odrzuceniu krotności występowania obiektu w próbie z uwagi na dość skomplikowany wzór na wariancję $V(\bar{Y}_d)$.

5. W statystyce matematycznej ważną rolę odgrywa idea Fisherowska wiarogodności. Niech $d = \{(k, Y_k) : k \in s\}$ będą danymi z próby. Ogół wektorów $Y \in R_N$ można podzielić na dwie klasy: 1) wektorów zgodnych z danymi z próby, tj. takich, których współrzędne Y_k dla $k \in s$ są identyczne z współrzędnymi z próby oraz 2) pozostałych wektorów przestrzeni R_N. Oznaczmy podprzestrzeń wektorów zgodnych Ω_d. Funkcja wiarogodności $L_d(Y)$ jest to taka funkcja parametru Y, która dla dowolnego $Y \in R_N$ jest prawdopodobieństwem otrzymania d. Stąd dla dowolnego planu próbkowania p funkcja wiarogodności dana jest wzorem:

$$L_d(Y) = \begin{cases} p(s) \text{ dla } Y \in \Omega_d \\ 0 \text{ dla } Y \notin \Omega_d \end{cases} \tag{3}$$

Wynika z tego, że funkcja wiarogodności nie wyróżnia w Ω_d żadnego wektora Y.

Do problemu wnioskowania, opartego o funkcję wiarogodności, podeszli nieco inaczej: Royall (1968) oraz, niezależnie, Hartley i Rao (1969). Rozważali zagadnienie przyjmując założenie, że identyfikatory danych z próby po jej otrzymaniu, w estymacji pomija się. Abstrahując od sprawy słuszności takiej idei, otrzymali dla schematu losowania prostego bez

zwracania postać funkcji wiarogodności, pozwalającą na otrzymanie estymatora średniej $\bar{Y}$ metodą największej wiarogodności.

6. Szerokie studia prowadzone były co do użyteczności takich kryteriów w estymacji — stosowanych w statystyce matematycznej — jak dopuszczalność, minimalność i niezmienniczość estymatorów. Mówimy, że dla danego planu próbkowania p estymator t_1 jest co najmniej tak samo dobry, jak inny estymator t_2, jeśli średni błąd kwadratowy estymatora t_1, $MSE(t_1) \leqslant MSE(t_2)$ dla wszystkich $Y \in R_N$. Jeśli ponadto co najmniej dla pewnego $Y_o \in R_N$ zachodzi nierówność ostra, stwierdzamy, że estymator t_1 jest **lepszy** niż t_2. W klasie C estymatorów estymator t_0 nazywamy **dopuszczalnym** przy danym planie p wtedy i tylko wtedy, gdy nie istnieje w C estymator lepszy niż p_0. Badania Joshiego (1965, 1966) wykazały, że dla dowolnego danego planu próbkowania w klasie A wszystkich estymatorów średniej $\bar{Y}$ estymatorem dopuszczalnym jest zarówno średnia z danych dla **różnych** jednostek w próbie, $\bar{y}_s = \sum_s Y_k : \nu(s)$, $\nu(s)$ — liczba różnych jednostek w próbie), jak i estymator ilorazowy $t_R = \bar{X}\bar{y}_s/\bar{x}_s$; ($X$ — cecha dodatkowa), a ograniczając plany próbkowania do planów FES (n), w których $\nu(s) = n$ jest stałe dla dowolnych $s \in S$, wszystkie estymatory typu $t = \sum_s b_k Y_k$ spełniające 3 warunki: 1) $b_k \geqslant \frac{1}{N}$, *dla* $k = 1, 2, ..., N$; 2) $b_k = \frac{1}{N} \Rightarrow \alpha_k = 1$ (α_k jest prawdopodobieństwem wejścia do próby); 3) $\sum_{k=1}^{N} b_k^{-1} = nN$, są dopuszczalne. Godambe i Joshi udowodnili, że estymator t_{HT} Horvitza-Thompsona[6]), dla dowolnego danego planu próbkowania *nie-FES* (istnieją próby o różnych liczebnościach różnych jednostek) w klasie estymatorów liniowych jednorodnych jest niedopuszczalny. Jeśli plan próbkowania jest $FES(n)$, to t_{HT} w klasie A jest dopuszczalny. Z faktu dopuszczalności w klasie A estymatora $\bar{Y}_s$ wynika jego dopuszczalność w węższej klasie, np. A_u estymatorów nieobciążonych; w tej ostatniej klasie, przy jakimkolwiek danym planie próbkowania, nie można więc wykorzystywać identyfikatorów do konstrukcji lepszego estymatora średniej $\bar{Y}$.

Badano również dopuszczalność strategii, tj. par (p, t_i), gdzie p jest planem próbkowania, a t estymatorem, a ponadto zagadnienie estymacji minimaksowej oraz niezmienniczości estymatora. Badania te pozwoliły wniknąć głębiej w strukturę problemu wnioskowania o średniej $\bar{Y}$ w populacji skończonej.

7. W powyższych badaniach przyjmowano założenie Neymana, że badana populacja jest stała, skończona, a jej obiekty identyfikowalne. Dany jest więc

---

[6]) $t_{HT} = \sum_s Y_k/N\alpha_k$. $\alpha_k > 0$ jest prawdopodobieństwem wylosowania k-tego obiektu populacji do próby ($k = 1, 2, ..., N$).

wektor $Y=(Y_1, Y_2, ..., Y_n)$ chociaż jego współrzędnych nie znamy. Do jego oszacowania (właściwie dla oszacowania średniej lub wartości globalnej, co jest równoważne) losujemy próbę, przyjmując pewien plan próbkowania. Problemem jest znalezienie „najlepszej" strategii próbkowania i estymacji. Wyniki badań nie doprowadziły do jakiejś jednoznacznej odpowiedzi, pomimo stosowania różnych metod klasycznego wnioskowania statystycznego. Badane klasy estymatorów okazały się zbyt szerokie. Przypuszczalnie przyczyną jest to, że parametr Y składa się z N składowych parametrów $Y_1, ..., Y_N$, a więc znacznie większej ich liczby niż wynosi liczebność n próby, $n<N$. Nasuwa to sugestię, że dla redukcji zagadnienia należałoby nałożyć pewne warunki na składowe wektora parametrów. To prowadziło do założenia, że populacja jest próbą losową z pewnej **nadpopulacji,** inaczej: $Y=[Y_1, Y_2, ..., Y_N]$[7]), na pewien rozkład ξ prawdopodobieństwa, a populacja $y=(y_1, y_2, ..., y_N)$ jest próbą losową z nadpopulacji. Ponieważ zakłada się pewien „model" nadpopulacji, czyli pewien układ warunków określających klasę rozkładów, do której ξ przypuszczalnie należy, wnioskowanie z danych próby s nosi nazwę podejścia **modelowego** (model-dependent, model-based); gdy nie korzystamy z żadnego modelu wnioskowanie nazywamy podejściem **randomizacyjnym.** W podejściu randomizacyjnym występuje jeden rodzaj losowości, mianowicie losowości powodowanej próbkowaniem probabilistycznym (losowym wyborem próby). W podejściu modelowym dochodzi drugi rodzaj losowości: populacja jest próbą prostą z nadpopulacji. Mamy wówczas p-losowość i ξ-losowość. Optymalność można rozpatrywać przy założonej klasie (p, t) strategii, np. p-nieobciążonych, bądź gdy plan próbkowania uważa się za mniej ważny, badać konsekwencje, wywodzące się głównie z założonego modelu ξ nadpopulacji. Wnioskowanie statystyczne może być prowadzone klasycznymi metodami statystyki, bądź metodami bayesowskimi. Zauważmy, że obecnie $y=(y_1, y_2, ..., y_N)$ nie odgrywa roli parametru, jak w podejściu randomizacyjnym, natomiast rodzinę ξ rozkładów można parametryzować indeksując ją parametrem θ z odpowiedniej przestrzeni parametrycznej.

8. Klasę rozkładów ξ określa się najczęściej specyfikując[8]) wartość oczekiwaną $\mu_k=\mathcal{E}(y_k)$, wariancję $\sigma_k^2=V(Y_k)$ oraz kowariancję $\sigma_{kl}=C(Y_k, Y_l)$, $k\neq l$; $k, l=1, 2, ..., N$. W użyciu można wyróżnić 11 podstawowych modeli [2, str. 90]. Jednym z nich, często stosowanym, jest model regresji G_R. Przyjmuje się w nim, że $Y_1, Y_2, ..., Y_N$ są zmiennymi losowymi niezależnymi oraz $\mu_k=\beta x_k$, $\sigma_k^2=\sigma^2 v(x_k)$ dla $k=1, 2, ..., N$, β oraz σ^2 są nie znane, v (.) jest znaną funkcją, a $x_1, x_2, ..., x_N$ są znanymi liczbami dodatnimi. Na ogół przyjmuje się $v(x_k)=x_k^g$, g jest znane (w praktyce raczej nie znane) $0\leqslant g\leqslant 2$.

---

[7]) Przy podejściu modelowym dużymi literami (Y, Y, Y_k) oznaczamy zmienne losowe, które określa model nadpopulacji, małymi literami (y, y_1, ...) dane dla populacji lub próby.

[8]) E(.), V(.), C(.) oznaczają operatory w odniesieniu do modelu, natomiast E(.), V(.), MSE(.) w odniesieniu do planu próbkowania.

W podejściu modelowym statystykę T, zastosowaną do wnioskowania o średniej $\bar{y}=N^{-1}\sum_{k}^{N} y_k$, nazywamy **predyktorem zmiennej losowej** $\bar{Y}=N^{-1}\sum_{k}^{N} Y_k$. Wartość predyktora T dla danych z próby jest oceną wartości $\bar{y}$. Między statystykami występuje różnica opinii co do tego, który aspekt jest istotniejszy, modelowy czy randomizacyjny? Jeśli za ważniejszy uważa się aspekt randomizacyjny, poszukuje się najlepszej strategii (p, T), przyjmując za kryterium minimalizację $E\,MSE\,(p, t)$ lub ograniczając się do estymatorów *p-nieobciążonych minimalizację* $E\,V(p, T)$. W przypadku przeciwnym, gdy na pierwszym miejscu stawia się model, za cel uważa się uzyskanie predyktora T dla danej próby s, *minimalizującego* $E(T-\bar{Y})^2$. Uprzeciętnienie, ze względu na plan próbkowania, uznaje się za mniej istotne. Stąd w podejściu modelowym problem wyboru dobrej strategii (p, T) jest mniej ważny niż problem wyboru predyktora T, który jest dobry dla próby aktualnie otrzymanej.

9. Dla próby s obejmującej $v(s)$ różnych identyfikatorów, $s=\{k_1, k_2, ..., k_{v(s)}\}$ oznaczmy $f_s=v(s):N$ oraz $\bar{Y}_s$, $\bar{X}_{\bar{s}}$ — średnie dla elementów próby oraz dla pozostałych elementów populacji, odpowiednio. Predyktor T dla danej próby ma formę:

$$T=f_s\,\bar{Y}_s+(1-f_s)U \tag{4}$$

przy czym U jest predyktorem średniej $\bar{Y}_{\bar{s}}$. Najlepszym predyktorem liniowym ξ-nieobciążonym (*tzw. ξ-BLU* predyktor), T_{BR}, jest taki, który spełnia kryterium minimalizacji $E\,MSE$. W przypadku wyżej określonego modelu G_R regresji dla dowolnego planu próbkowania p,

$$T_{BR}=f_s\,\bar{Y}_s+(1-f_s)\,\hat{\beta}\,\bar{x}_{\bar{s}} \tag{5}$$

gdzie:

$$\bar{x}_{\bar{s}}=\sum_{\bar{s}} x_k : [N-v(s)],\quad \hat{\beta}=\sum_{s}\frac{x_k\,Y_k}{v(x_k)}\Big/\sum_{s}\frac{x_k^2}{v(x_k)}$$

Poszukując optymalnej strategii (p_{opt}, T_{BR}) Royall (1970) wykazał, że w klasie planów próbkowania $FES(n)$, jeśli $v(x)$ jest funkcją x niemalejącą, natomiast $[v(x) : x^2)]$ jest funkcją nierosnącą, to najlepszym planem p_{opt} jest wybór do próby n obiektów o najwyższych wartościach x_k, czyli **wybór celowy**. Dla innych modeli otrzymamy inne T_{BR} oraz p_{opt}. Jeśli model G_R jest fałszywy, próba p_{opt} może dawać wynik szacunku $\bar{y}$ o dużym błędzie predykcji. Zwykle szacujemy średnie dla różnych cech i dla nich mogą być słuszne różne modele nadpopulacji.

10. Wyniki studiów przy podejściu modelowym są dyskusyjne. Wywołały krytykę części statystyków. Krytykę tę przedstawili m.in. czołowi statystycy amerykańscy, współpracujący przy planowaniu badań reprezentacyjnych

Biura Spisów USA, M. H. Hansen, W. G. Madow i B. J. Tepping [4]. Opierając się na bogatym doświadczeniu uważają, że podejście modelowe może być pomocne przy konstrukcji planu próbkowania, ale jest niewłaściwe przy analizie wyników badania reprezentacyjnego. Wysuwają tezę, że dla próbkowania probabilistycznego istotna jest zgodność randomizacyjna estymatorów. Zgodność taka może być zachowana, gdy próba jest odpowiednio duża, wówczas wnioskowanie jest niezależne od rozkładu cech w skończonej populacji. Nie ma więc potrzeby opierania wnioskowania o model nadpopulacji. Autorzy, dla pokazania na konkretnej nadpopulacji zachowania się estymatorów zwykle stosowanych w badaniach reprezentacyjnych oraz *BLU* estymatorów podejścia modelowego, przeprowadzili eksperyment metodą *Monte-Carlo*. Analizując wyniki tego eksperymentu oraz toczących się dyskusji na ten temat można wysunąć następujące wnioski:

▲ Nie uzyskamy „najlepszych" estymatorów nie ograniczając przesadnie klasy rozważanych estymatorów. „Najlepszy" estymator przy podejściu modelowym jest takim jedynie, gdy założony model jest w pełni trafny; w przypadku niedużej trafności modelu otrzymujemy mylne przedziały ufności dla szacowanej wielkości.

▲ W porównaniu z „najlepszymi" estymatorami podejścia modelowego korzystniej jest stosować „dobry" estymator podejścia randomizacyjnego, gdyż w przypadku dostatecznie dużej próby losowej dostarczy on prawidłowych przedziałów ufności.

▲ Modelowo otymalne plany próbkowania zawierają ryzyko poważnego zaniżania średniego błędu kwadratowego, nawet jeśli model jest zadowalający. Gdy model okazuje się niezgodny z danymi próby może występować jeszcze znaczniejsze zaniżenie długości przedziału ufności.

▲ Korzystanie z modelu może pomagać przy ustalaniu schematu losowania próby i jego efektywności, ale wówczas, gdy próba jest duża, wnioskowanie z niej nie powinno być zależne od modelu nadpopulacji.

▲ Podejście modelowe może być słuszne w przypadku małych prób losowych, dla których wnioskowanie randomizacyjne nie jest dość precyzyjne.

▲ Nie można podać reguły ustalającej, kiedy próba jest dostatecznie duża i pozwala na uznanie randomizacyjnego rozkładu, estymatora zgodnego jako w przybliżeniu normalnego, co wiąże się z konstrukcją przedziałów ufności. To zależy od badanej populacji, prawidłowego stosowania takich mechanizmów, jak podział populacji na warstwy, losowanie zespołowe jedno- lub wielostopniowe, stosowanie odpowiednich prawdopodobieństw wyboru jednostek losowania, wykorzystanie do estymacji informacji dodatkowej itd. Zwykle uznaje się próbę poniżej 25 jednostek jako małą, natomiast powyżej 100 jednostek, jako dużą. Te wielkości odnoszą się do liczby jednostek pierwszego stopnia (*jlps*) w próbie bądź do liczby *jlps* dla dziedziny studiów, bądź jednostek przyczyniających się poważnie do oceny jakichś wielkości, dla pewnych podzbiorów badanej populacji.

11. Studia teoretyczne podstaw metody reprezentacyjnej, w zastosowaniu do populacji skończonych, dotyczą z reguły zagadnienia wnioskowania o średniej bądź — co jest logicznie równoważne — wartości globalnej **jednej** cechy. Badania reprezentacyjne dotyczą z reguły wielu cech, przy czym szacowane są nie tylko wartości globalne, ale i inne charakterystyki równocześnie. Wybór planu próbkowania w przypadku konkretnej populacji uwarunkowany jest możliwościami technicznymi i kosztami badania. Planując badanie reprezentacyjne szuka się rozwiązań możliwie najtańszych, przy założonym zakresie badania i pożądanej precyzji ocen szacowanych wielkości. Badania teoretyczne abstrahują od tych aspektów. Teoretyczna optymalność estymatora zakłada pewne ograniczenia, które mogą nie występować w praktyce. Tzw. „niedopuszczalna" strategia może być w pewnych warunkach użyteczniejsza niż strategia „dopuszczalna".

12. Na całym świecie ustawienie badań reprezentacyjnych zależy w dużej mierze od istniejących warunków technicznych. W masowych badaniach reprezentacyjnych Głównego Urzędu Statystycznego obserwuje się różne rozwiązania, różne plany próbkowania probabilistycznego, przykrojone do celów i możliwości organizacyjnych, w zgodzie z teorią takich badań. Podam dwa przykłady. Badając migracje w czasie *NSP 1978* r. plan próbkowania był poprzedzony wstępnymi studiami pięciu możliwych do realizacji planów (B. Lednicki, [5], str. 106—127). Ostatecznie zastosowano jeden z nich, który wydawał się najlepszy, a polegał na systematycznym losowaniu mieszkań z obwodów spisowych. Było to więc losowanie jednostopniowe warstwowe, w którym warstwami były obwody spisowe. Wprawdzie nie było to rozwiązanie optymalne, gdyż obwody spisowe nie są wewnętrznie bardziej jednorodne niż populacja, ale nie było realnego lepszego postępowania. Prostą postać miał estymator wartości globalnej. Był to estymator nieobciążony, a jego wariancję łatwo było oszacować. W innych warunkach był przeprowadzony *Mikrospis 1984* (A. Szarkowski, [10], str. 22—38). Na wsi zastosowano losowanie zespołowe (zespoły stanowiły obwody spisowe), warstwowe (warstwami były województwa). Lepszym rozwiązaniem byłby wybór dwustopniowy próby, ale nie wyrazili na to zgody organizatorzy spisu, obawiając się niekorzystnego wpływu na spis rozproszenia próby w terenie, co mogło wywołać częstsze błędy obserwacji. W miastach plan próbkowania był dwustopniowy oparty o strategię Rao-Hartleya-Cochrana (1962). Na pierwszym stopniu losowanie było warstwowe, jak na wsi, ale z różnymi prawdopodobieństwami wyboru, proporcjonalnymi do wielkości obwodu spisowego, mierzonej liczbą mieszkań (oszacowaną); na drugim stopniu w każdym z wylosowanych obwodów spisowych losowano mieszkania. Losowanie było proste bez zwracania. Strategia RHC[2, str. 154] jest w przypadku modelu G_R (str. 10) strategią optymalną gdy $v(x_k)=x_k^2$, natomiast nie ma zalet optymalności, gdy $v(x_k)=x_k^g$ dla $g\neq 2$. W *Mikrospisie* rolę cechy X pełniła liczba mieszkań w obwodzie.

BIBLIOGRAFIA

[1] Bracha Cz. (1987): *Wykorzystanie informacji o cechach dodatkowych w badaniach reprezentacyjnych*, ZBSE, Warszawa.
[2] Cassel C. M., Särndal C. E., Wretman J. H. (1977): *Foundations of Inference in Survey Sampling* J. Wiley, New York.
[3] Cochran W. G., (1977): *Sampling Techniques, third edition*, J. Wiley New York.
[4] Hansen M. H., Madow W. G., Tepping B. J. (1983): *An Evaluation of Model–Dependent and Probability–Sampling Inferences in Sample Surveys*, JASA, 78, 776—793.
[5] *Metoda reprezentacyjna w masowych badaniach statystycznych. Teoria i praktyka* (1981), Z Prac ZBSE, z. 122.
[6] Neyman J. (1933): *Zarys teorii i praktyki badania struktury ludności metodą reprezentacyjną*, Warszawa.
[7] Neyman J. (1934): *On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection*, JRSS 97, 558—625.
[8] Steczkowski J. (1988): *Zastosowanie metody reprezentacyjnej w badaniach społeczno-ekonomicznych*, PWN, Warszawa.
[9] Zasępa R. (1972): *Metoda reprezentacyjna*, PWE, Warszawa.
[10] *Zastosowanie metody reprezentacyjnej w masowych badaniach statystycznych GUS* (1987), Z Prac ZBSE, z. 166.

Kilka myśli o metodzie reprezentacyjnej w badaniach zjawisk społeczno-ekonomicznych

prof. dr hab. Jan Steczkowski
Akademia Ekonomiczna w Krakowie

Nie chciałbym w tej wypowiedzi zamieszczać podręcznikowej wiedzy na temat zalet, jakie posiada metoda reprezentacyjna, gdyż zrobiłem to już choćby w świeżo wydanej książce (*Zastosowanie metody reprezentacyjnej w badaniach społeczno-ekonomicznych*, PWN, Warszawa 1988). Również nie chciałbym wyręczać pracowników Głównego Urzędu Statystycznego w przedstawianiu przykładów zastosowań tej metody w prowadzonych tam przez nich badaniach. Zresztą i na ten temat pojawiła się ostatnio odpowiednia publikacja (*Zastosowanie metody reprezentacyjnej w badaniach statystycznych GUS 1981—1986*, Zakład Badań Statystyczno-Ekonomicznych GUS i PAN, z. 166, Warszawa 1987).

Chciałbym jedynie podzielić się kilkoma refleksjami związanymi z poruszoną problematyką, które — być może — dadzą asumpt do szerszej dyskusji.

Wypada rozpocząć od dość oczywistego stwierdzenia, iż metoda reprezentacyjna wiąże statystykę matematyczną z jej konkretnymi zastosowaniami, przy czym jest to szczególnie wyraźne na odcinku badań społeczno-gospodarczych. Metoda ta nie tylko odwołuje się do teorii próby, ale też i do teorii pomiaru, teorii błędu oraz metodyki badań, rozwiniętej na terenie nauk empirycznych. Stąd obok zagadnień, wchodzących w zakres rozważań formalno-matematycznych, występują zagadnienia przynależne metodologii właściwej danej dyscyplinie naukowej, jak też problematyce związanej z merytoryczną stroną dociekań, co pozwala na uwzględnienie również specyfiki analizowanego zjawiska. Właśnie w tym skomplikowanym powiązaniu różnych dziedzin wiedzy leży siła poznawcza metody reprezentacyjnej, ale stwarza ono także szereg dylematów tak teoretycznej, jak i aplikacyjnej natury; w praktyce bowiem nie wystarczy ograniczyć się do znajomości właściwych reguł, trzeba także wykazać się w danych warunkach umiejętnością prawidłowego ich zastosowania.

W. Skrzywan, profesor Uniwersytetu Jagiellońskiego, już wiele lat temu sformułował następującą konstatację: „Statystyka rozbudowuje metodę indukcyjnego rozumowania w odniesieniu do danych empirycznych i uznaje prymat teorii przedmiotowej. Statystyk posługuje się metodami matematycznymi, które nie są dla niego celem, jak dla matematyka, gdyż celem tym jest poznanie rzeczywistości. Jest to cel wspólny z każdą nauką przedmiotową i dlatego wspólna praca badawcza jest dla statystyka zadaniem dnia powszedniego, własnym zadaniem naukowym"[1]).

Rola statystyka, jako dostarczyciela metod badawczych, jest bardzo niewdzięczna. Z jednej strony bardzo często gromiony jest przez matematyków za brak pełnej poprawności formalnej w zakresie stosowanych procedur, z drugiej strony użytkownicy tych metod zarzucają statystykom abstrakcyjność proponowanych rozwiązań i w zbyt małym stopniu uwzględnianie specyfiki tej dziedziny wiedzy, w której metody statystyczne są stosowane. Te dwa rodzaje przekonań dostarczają sobie nawzajem argumentów, co do słuszności reprezentowanych postaw.

Wyżej przedstawione tendencje w szczególny sposób ujawniają się w zakresie metody reprezentacyjnej. Zachowanie w tej mierze optymalnej dla badań równowagi nie jest zadaniem prostym i dlatego dadzą się wyróżnić jakby dwie grupy statystyków.

Jedni zasadniczy wysiłek kierują na otrzymywanie coraz efektywniejszych estymatorów tylko poprzez formalną analizę zagadnienia. Często dochodzi do takich komplikacji, że trzeba posługiwać się coraz mocniejszymi założeniami, których utrzymanie w konkretnych badaniach jest coraz mniej realne. Prowadzi to do swoistej „matematyzacji" badań statystycznych, kończących się „de facto" na konstrukcji modelu matematycznego niezbyt spójnego z opisywaną przy jego pomocy rzeczywistością. Wiadomym zaś

[1]) *Sprawozdanie z Konferencji Antropologów*, Przegląd Antropologiczny, tom XIX, Poznań 1953, s. 156.

jest, że zbyt długie ostrzenie siekiery, bez użycia jej w pewnym momencie do rąbania, jest zajęciem wysoce jałowym.

Inni znów statystycy wykorzystują w badaniach estymatory — rzec można — na wyrost, tzn. nie odpowiadające całkowicie lub tylko częściowo zastosowanemu schematowi losowania (mówiąc inaczej — warunkom podjętych badań). Z drugiej strony jednak często słabości, będące udziałem danej dyscypliny naukowej, w której stosuje się metodę reprezentacyjną przypisuje się właśnie statystyce i wysuwa się pod jej adresem nieuzasadnione żądania, gdy wiadomo, że koncepcję badań formułuje się zgodnie z merytoryczną ich potrzebą i dopiero po zapoznaniu się z nią statystyk może podjąć próbę wyrażenia jej w terminologii i symbolice właściwej probabilistyce oraz wykorzystać właściwe twierdzenia rachunku prawdopodobieństwa i oparte na nich metody statystyczne. Należy przy tym zauważyć, że w ostatnich czasach, dla rozwiązania przedstawionego zadania, dużą pomoc stanowią tzw. eksperymenty symulacyjne wspomagane komputerowo oraz rozwój dociekań nad tzw. odpornością estymatorów.

Na kanwie sformułowanych stwierdzeń o charakterze ogólnym trzeba z kolei przejść do omówienia kilku konkretnych przypadków, stanowiących ilustrację trudności, na jakie napotyka teoria próby w jej praktycznych zastosowaniach.

Przede wszystkim bardzo ważny problem wiąże się z tzw. „twierdzeniami granicznymi", a przede wszystkim z „prawem wielkich liczb", które w metodzie reprezentacyjnej ma kluczowe znaczenie. W matematyce obowiązuje „indukcja zupełna", tzn. jeden kontrprzykład obala wypowiedziane twierdzenie. Gdyby obowiązywała ona również w badaniu zjawisk, wchodzących w zakres nauk empirycznych, te ostatnie nigdy nie mogłyby się pojawić, bowiem w naukach empirycznych zawsze można znaleźć choć jeden przykład zaprzeczający wypowiedzianemu sądowi o rzeczywistości. Jeżeli jednak w miarę zwiększania liczby zdarzeń losowych (niezależnych od siebie obserwacji) ciąg ocen charakteryzuje się stochastyczną zbieżnością do rzeczywistej wartości szacowanego parametru, wtedy można mówić o występującej prawidłowości, jako o fakcie naukowym, który realizuje się z określonym prawdopodobieństwem i którego predykcja jest obciążona pewnym błędem. W naukach empirycznych poszukuje się więc warunku koniecznego pojawienia się pewnego wyróżnionego zdarzenia losowego (stanu jakiegoś zjawiska), gdyż badacz nie może określić dla niego warunku dostatecznego. Metoda reprezentacyjna, jak cała statystyka, zajmuje się „zjawiskami masowymi", w których prawidłowości ujawniają się po uwzględnieniu odpowiedniej liczby zdarzeń i nie jest ona zainteresowana w rozpatrywaniu poszczególnych, nietypowych przypadków. Mimo tego w metodzie reprezentacyjnej staramy się, aby próba nie była zbyt „masowa" i sprowadzała się do koniecznej, ale i wystarczającej liczby rozpatrywanych zdarzeń. Przedstawione podejście na pewno nie odpowiada kilku dyscyplinom humanistycznym (a może tylko kilku „szkołom" w nich reprezen-

towanym), ale dla nauk społeczno-ekonomicznych, które wciąż jeszcze przyznają się do rodowodu nauk empirycznych, jest to droga postępowania najbardziej właściwa, choć może nie jedyna. To, co zostało w tej chwili wypowiedziane zatrąca truizmem, ale nie zawsze i nie wszyscy w pełni zdają sobie z tego sprawę.

Wypowiedziane dotąd myśli bezpośrednio prowadzą do koncepcji nadpopulacji (universum). Skończona liczebność zbiorowości generalnej oraz wynikające stąd skomplikowanie formuł, określających poszczególne estymatory i ich błędy, stwarza zarówno trudności przy definiowaniu pewnych pojęć, jak też przy porównywaniu efektywności różnych schematów pobierania próby.

Jeżeli chodzi o pierwszą sprawę, to wiąże się ona np. z pojęciem „losowości", gdyż w przypadku wystąpienia zbiorowości skończonych można ściśle mówić tylko o niezależności jednostek nań składających się. Nie będę szczegółowiej omawiał tego zagadnienia, dodam tylko, że ciekawy artykuł poruszający zasygnalizowany problem został napisany przez Z. Hellwiga, B. Puzdrowską i A. Smoluka („Losowość a niezależność", Przegląd Statystyczny 3—4, 1983 r.).

Również definicja „estymatora zgodnego" tylko wówczas nie budzi kontrowersji, gdy ma się do czynienia z próbą prostą, stanowiącą efekt losowania nieograniczonego, indywidualnego, ze zwracaniem. W innych przypadkach, dla uniknięcia niedogodności spowodowanej skończonością zbiorowości generalnej, niektórzy teoretycy odwołują się do tzw. podejścia modelowego (based approach) odmiennego od klasycznego podejścia zwanego randomizacyjnym (randomization approach). W ostatnich latach ukazało się szereg prac na ten temat, wymienimy tylko kilka z nich:

— C. M. Cassel, C. E. Särndal, J. H. Wretman: *Foundation of Inference in Survey Sampling,* Wiley, New York 1977;
— M. H. Hansen, W. G. Madow, B. J. Tepping: *An Evaluation of Model — Dependent and Probability — Sampling Inferences in Sample Surveys,* Journal of the American Statistical Association, 78, 1983 r.;
— T. M. Smith: *Present Position and Development: Some Personal Views Sample Surveys,* Journal of the Royal Statistical Society, A 147,2, 1984;
— A. Chandhuri, I.W.E. Vos: *Unified Theory and Strategies of Survey Sampling,* North-Holland, Amsterdam 1988.

Najogólniej rzecz traktując (i niestety wydatnie ją upraszczając) nadpopulację kreuje się na drodze reprodukowania konkretnej zbiorowości generalnej dowolną liczbę razy, przy czym związany z tym wzrost liczebności jest wyłącznie powtarzaniem pierwotnych jej elementów, stąd rozkład (struktura) nadpopulacji jest identyczny, jak zbiorowości wyjściowej. Może jednak wystąpić i inne rozumienie nadpopulacji, jako rozkładu o tych samych parametrach. Przy tak rozumianej multiplikacji mogą być zastosowane eksperymenty symulacyjne z wykorzystaniem liczb pseudolosowych i elektronicznej techniki obliczeniowej. W tym przypadku nadpopulac-

ja jest modelem w ścisłym słowa tego znaczeniu. Ustalając prawidłowości o charakterze naukowym (czyli ogólnym) w zasadzie nie można zebrać i poznać całego „universum" związanego z badanym zjawiskiem, a więc trzeba przyjąć, że każda skończona zbiorowość generalna stanowi część większej całości. Można tedy zadać sobie pytanie, czy tę część należy traktować jako próbę wielostopniową uzyskaną z nadpopulacji? Jest to chyba ściśle związane z realizowaną koncepcją badań.

Podobny problem leży u podstaw, szeroko stosowanych w naukach przyrodniczych, powtórzeń badań (replikacji). Mimo wygody takiego postępowania, łatwiejszej interpretacji wyników, jak też większej możliwości oszacowania błędów, pochodzących z różnych źródeł, w badaniach zjawisk społeczno-gospodarczych powtórzenia oparte na wielu mniej liczebnych próbach (wziętych z danej przestrzeni prób) wciąż nie znajduje szerszego uznania, a raczej preferowane jest wnioskowanie oparte o jedną próbę, posiadającą stosunkowo dużą liczebność. Na domiar są to często badania ogólnopolskie, dające w efekcie zbyt ogólną charakterystykę zjawiska, nie przystającą do żadnej konkretnej sytuacji. Przykład na to daje meteorologia, że w miarę zwiększania zakresu przestrzennego badań wyniki są coraz mniej adekwatne do rzeczywistości. Czasem tak bywa i na odcinku zjawisk społeczno-gospodarczych. Gwoli sprawiedliwości wypada podkreślić, iż raczej nauki społeczne aniżeli ekonomiczne rozwijają się i stosują badania powtarzalne w czasie, tj. badania sukcesywne i ciągłe z rotacją i bez rotacji. Najpopularniejsza z nich to metoda panelowa zaproponowana w 1928 r. przez R. A. Rice'a, a następnie rozwinięta — między innymi — przez P. E. Lazarsfelda. Z tego, co zostało dotąd powiedziane, należy wysnuć wniosek, iż metoda reprezentacyjna stwarza na omawianym odcinku nauki warunki dla bardziej pogłębionych i szerszych badań. Wydaje się, że w tym zakresie dyscypliny społeczno-ekonomiczne nie wykorzystały wszystkich tkwiących w metodzie reprezentacyjnej możliwości. Uwaga ta odnosi się przede wszystkim do nauk ekonomicznych. Trzeba także dodać, że koncepcja nadpopulacji dobrze współgra właśnie z rzeczywistością społeczną, w której zawsze i wszędzie ma się do czynienia z dużymi lub nawet bardzo dużymi, reprodukującymi się populacjami ludzkimi.

Dotychczas przeprowadzone rozważania prowadzą do następującego stwierdzenia. Wbrew temu, co się dość powszechnie uważa, omawiana metoda służy nie tylko zbieraniu danych statystycznych, ale znajduje przy tym dodatkowe zastosowanie, stanowiąc istotny element samej koncepcji badawczej. W socjologii czy ekonomii posługiwanie się aparaturą naukową oraz klasycznym eksperymentem jest wykluczone i dlatego wykorzystanie reguły „ceteris paribus conditionibus" (metody izolacji) jest utrudnione. Właśnie statystyka, w jakimś stopniu, lukę tę zapełnia. Warto w tym miejscu powołać się na R. Lewina, który w jednym z numerów „Science" (z sierpnia 1983 r.) zwrócił uwagę, iż w ekologii wiedzę o zjawiskach uzyskiwano, prawie do ostatnich czasów, głównie na podstawie teorii oraz modeli (nawet

bardzo sformalizowanych), przy założeniu występowania współzawodnictwa między osobnikami. Od pewnego jednak czasu zaczęto przeprowadzać także badania z wykorzystaniem eksperymentów terenowych, połączonych z uważniejszą analizą danych uzyskanych poprzez obserwację konkretnych zjawisk. Efektem tego był wyraźny przełom w interpretacji uzyskanych wyników oraz większe docenianie rzeczywistości. Moim zdaniem, podobna sytuacja występuje na terenie ekonometrii, szczególnie w jej amerykańskim wydaniu, która wciąż poświęca główną uwagę modelowaniu nieraz atrakcyjnych hipotez, z przyjęciem zbyt arbitralnych założeń, bez odpowiedniego odwołania się do faktów. W naukach empirycznych zaś ostatecznym kryterium prawdy jest zawsze zagadnienie zgodności wypowiedzianych sądów z rzeczywistością.

Być może, że wciąż rozwijająca się metoda reprezentacyjna otworzy przed ekonometrią nowe obszary analiz, w większym stopniu uwzględniających realny świat zjawisk, korygujący idealne konstrukcje formalne, będące wytworem li tylko rozumu. Również dzięki tej metodzie poszczególni badacze, jak i instytucje samorządowe czy lokalne stowarzyszenia o niewielkich nawet możliwościach finansowych czy technicznych mogą — obok profesjonalnych jednostek — podejmować niezależne badania bardziej zgodne z własną koncepcją podejścia do rozważanego zagadnienia.

W związku z tym nasuwa się szereg dalszych myśli. W każdym reprezentacyjnym badaniu występują jakby dwie jego strony, tzn. sposób zbierania danych wraz z ich pomiarem oraz procedura wnioskowania statystycznego. Na obecnym etapie rozwoju występuje wyraźna skłonność do intensywnego rozbudowywania numerycznych technik, służących szacowaniu parametrów za pomocą nieraz bardzo wymyślnych estymatorów. Wymagają one nieraz spełnienia zbyt wielu i zbyt mocnych założeń, co nie pozwala na szersze ich zastosowanie w praktyce. Z drugiej zaś strony odczuwa się brak propozycji, pozwalających na tworzenie specyficznych schematów losowania dla potrzeb nauk społeczno-ekonomicznych tak, jak ma to miejsce np. w biometrii, doświadczalnictwie agro- i zootechnicznym, czy w statystycznej kontroli jakości. Mało kto z wymienionego kręgu ludzi próbował, na podobnej drodze, rozwiązywać niektóre zagadnienia używając przy tym mniej skomplikowanych, ale za to bardziej użytecznych narzędzi, których dostarcza matematyka. Zdaję sobie sprawę, że takie postawienie sprawy nie może satysfakcjonować przedstawicieli nauk matematycznych, ale postulat ten chyba znalazłby aprobatę u użytkowników metod ilościowych.

Powróćmy do stwierdzenia, iż w dyscyplinach społeczno-ekonomicznych aparatura naukowa, w jakiejś mierze jest zastępowana przez metody statystyczne. Wypowiadając tę myśl wypada uświadomić sobie fakt, iż istnieją dwa sposoby stabilizacji eliminowanych z badań przyczyn. Jeden zakłada, że jednostki badania są losowo przydzielane poszczególnym poziomom, czy kategoriom i wtedy ma się do czynienia ze schematem w pełni ulosowionym (zrandomizowanym). Drugi powołuje się na postulat

celowego upodobnienia wszystkich jednostek ze względu na eliminowane czynniki (jednorodność zbiorowości ze względu na cechy kwalifikujące). Już w 1920 r. Międzynarodowy Instytut Statystyczny wyłonił komisję dla zbadania mocnych i słabych stron celowego i losowego doboru jednostek do próby. Raport sporządzony w tej sprawie wyraźnie potwierdził głębokie przekonanie wielu ludzi, iż najbardziej racjonalne jest posługiwanie się doborem celowym (A. Jensen: *Report on the Representative Method in Statistics.* Bulletin of the International Statistical Institute 22, 1926). W efekcie konfrontacji tych przeciwstawnych punktów widzenia, szczególnie na polu nauk społeczno-ekonomicznych i w związku z metodą reprezentacyjną, wyłoniło się rozwiązanie kompromisowe, polegające na zastosowaniu badań częściowo ulosowionych (nie w pełni zrandomizowanych). W tym przypadku niejednorodną zbiorowość dzieli się na bardziej jednolite warstwy, by z kolei losować z nich odpowiednio liczne próby, jak ma to np. miejsce w schemacie losowania warstwowego. Tak więc postulat losowości zostaje uwzględniony wtedy, gdy nie ma już możliwości dalszej sensownej stratyfikacji zbiorowości.

Poruszony problem wiąże się z posiadaniem tzw. „dodatkowej informacji". Dlatego trzeba dodać, że w zależności od rodzaju tej informacji i stopnia jej przejrzystości konstruowane są także różne iloczynowe lub regresyjne, proste lub złożone estymatory, (Cz. Bracha: *Wykorzystanie informacji o cechach dodatkowych w badaniach reprezentacyjnych,* Zakład Badań Statystyczno-Ekonomicznych GUS oraz PAN, Warszawa 1987).

Najbardziej konsekwentnym stanowiskiem w tej mierze charakteryzuje się jednak tzw. podejście bayesowskie (M. Boyer, R. E. Kihlstrom (ed.) *Bayesian Models in Economic Theory,* North-Holland, Amsterdam 1984). Także i bliskie im rozważania mają szczególne znaczenie dla dokładności, precyzji i efektywności badań przeprowadzonych na terenie, na którym użycie aparatury naukowej jest wykluczone, lub zbyt kosztowne.

Stwierdzono więc, że ludzie mają większą skłonność do posługiwania się doborem celowym. Co by jednak nie sądzono o doborze losowym człowiek nie jest w stanie, w skali ziemskiej, poznać „prawdziwej wartości parametru, opisującego obserwowane lub wywołane zjawisko, gdyż — jak wiemy — działają nieuniknione, liczne oraz wielokierunkowe odchylenia powstałe pod wpływem przyczyn o charakterze losowym, a których nie kontrolowane źródła tkwią w samym badaniu, narzędziach pomiarowych i nieuchwytnych zmianach w warunkach otoczenia, zachodzących w trakcie przeprowadzanych badań. A. Einstein (*Quoted in Born* 1949) mógł krytykować probabilistyczne podejście w fizyce i stać na gruncie pełnego determinizmu, ale przy całym szacunku dla tego wielkiego umysłu, determinizm w koncepcjach społeczno-ekonomicznych jest nie tylko nie do utrzymania, ale prowadzi do katastrofalnych skutków praktycznych, a przede wszystkim do totalizmu. Nie wnikając głębiej w tę problematykę trzeba przyznać, że większość interesujących nas zjawisk podlega przemianom w obrębie granic,

wyznaczonych przez dwa krańcowe stany modelowe, których nigdy w pełni nie osiągają. Są to sytuacje, w których występuje:

— maksymalna organizacja, a entropia równa się zeru,
— maksymalna entropia, gdy organizacja równa się zeru.

W procesie kreacji i rozpadu realizują się wyłącznie stany pośrednie, leżące pomiędzy tymi dwoma (zresztą mało interesującymi z poznawczego punktu widzenia) biegunami: idealnego determinizmu oraz pełnego chaosu. Stąd można tylko orzekać o większym lub mniejszym prawdopodobieństwie pojawienia się jakiegoś stanu pośredniego. Wynika z tego pewna zasada ogólna, którą D. Hume ujął w zdaniu: „Nie ma nic ważniejszego w nauce i w życiu, jak rozróżnianie tego, co istotne od tego, co przypadkowe".

Czemu służą tego rodzaju rozważania? Trzeba mianowicie zdać sobie jasno sprawę, że w zakresie omawianych zjawisk „zmienne losowe" mogą być generowane wyłącznie dzięki właściwie zastosowanej metodzie reprezentacyjnej. Wszystkie inne źródła danych społeczno-ekonomicznych mają charakter wtórny i nigdy z pełnym przekonaniem (nie mówiąc już o pełnym uzasadnieniu) nie można przyjąć, iż ma się wtedy rzeczywiście do czynienia ze zmiennymi losowymi. Stosowanie zaś nawet najbardziej wyrafinowanych metod, do wątpliwych danych, nie potrafi zmniejszyć ryzyka popełnienia zasadniczego błędu systematycznego. Dzisiaj jest to szczególnie ważne, gdy dysponujemy szerokim wachlarzem użytecznych metod ilościowych oraz coraz efektywniejszymi programami komputerowymi, służącymi przetwarzaniu danych i obliczeniom numerycznym, a jednocześnie wyraźnie brak rzetelnej, pełnej i porównywalnej informacji. W tej chwili „wąskim gardłem" analiz społeczno-gospodarczych są właściwe i dostępne bazy danych. Właśnie metoda reprezentacyjna jest jednym z ważniejszych sposobów ich tworzenia. W związku z tym, co przed chwilą zostało stwierdzone, pozwolę sobie na kolejną dygresję. Mianowicie należałoby się zastanowić nad powołaniem w uczelniach ekonomicznych, na kierunku „Cybernetyka ekonomiczna", specjalizacji pod nazwą „Pozyskiwanie i przetwarzanie danych społeczno-gospodarczych", która byłaby umiejętnie zestawionym amalgamatem takich dyscyplin, jak: teoria, traktująca o podstawowych kategoriach ekonomicznych, informatyka, metodologia i metodyka badań, teoria pomiaru, teoria błędu, metody taksonomiczne i inne techniki grupowania danych, prezentacja tabelaryczna i graficzna danych, ich przekształcanie itd. Chodziłoby o kształcenie specjalistów, którzy zaspokajaliby potrzeby dużych korporacji gospodarczych, instytucji, zajmujących się zbieraniem i analizą danych, jak też wszelkich instytucji i placówek o charakterze studialnym czy też ośrodków koordynujących życie gospodarcze. W tej chwili Główny Urząd Statystyczny w tym zakresie wykonuje ogromną pracę, ale nie jest już w stanie podołać wszystkim potrzebom, które stale będą rosnąć w miarę realizacji przewidywanych i nieuchronnych zmian w gospodarce narodowej. Za przestrogę można przy tym przyjąć to, co napisali już dawno F. Yates i M. J. R. Healy:

„Komputery są dobrymi sługami, ale złymi panami. Wiele statystycznych nonsensów spłodzono przy biurkach, ale nie da się ich porównać z zalewem, który grozi w tym zakresie ze strony komputerów, jeżeli będą one używane bezmyślnie i bez należytej kontroli". („*How Should We Reform the Teaching of Statistics*", Journal of the Royal Statistical Society, A 127, 1964). Nie jest właściwe przechodzenie do porządku dziennego nad przykładami działalności naukowej, polegającej na „pakowaniu" do komputera byle jakich danych, bez żadnej przewodniej myśli i „mielenie" ich w nim przy pomocy „ad hoc" dobranych metod, by na koniec w sposób naciągany interpretować uzyskane wyniki.

Już z tych ogólnych uwag, które poczyniłem wynika, iż metoda reprezentacyjna jest działem statystyki, legitymującym się wieloma interesującymi problemami teoretycznymi, jak też posiada duże znaczenie, jako narzędzie pozyskiwania danych oraz ich analizy, przy czym jej przydatność jest szczególnie cenna dla dyscyplin społeczno-ekonomicznych. Nic tedy dziwnego, iż w świecie naukowym, przede wszystkim anglosaskim, cieszy się ona dużym zainteresowaniem, o czym świadczy bogata i wciąż rosnąca literatura przedmiotu. Najwybitniejsi statystycy poświęcają jej sporo uwagi. Mimo tego, zakres problematyki tej dziedziny jest tak szeroki i wiele zagadnień zostało dotąd ledwie dotkniętych. W zasadzie jest ona dość wyczerpująco opracowana dla próby prostej, myślę przy tym o tych osiągnięciach, które powszechnie zostały zaakceptowane przez badaczy.

Inne rodzaje prób ograniczają się do sposobów szacowania najbardziej podstawowych charakterystyk statystycznych. W tym zakresie zupełnym ugorem leży weryfikacja hipotez statystycznych, może poza analizą sekwencyjną, która po dwudziestu latach milczenia przeżywa obecnie swoją drugą młodość. Nie znaczy to, że nie pojawia się wiele przyczynków i prac teoretycznych, pozostają one jednak wciąż jeszcze w fazie dyskusji teoretycznej. Dyskusja ta i uzgadnianie sądów trwa bez przerwy na forum międzynarodowym, ukazując coraz to nowsze, pasjonujące nawet aspekty poruszanych spraw. Ostatnio ukazało się np. obszerne dzieło pod redakcją P.R. Krishnaiah'a i C.R. Rao („Sampling", North-Holland, Amsterdam 1988), gdzie wielu znanych specjalistów pisze o takich sprawach, jak: pobieranie próby w czasie, próby powtarzalne, estymacja ilorazowa i regresyjna, estymatory odporne, estymacja wieloparametrowa, oszacowania optymalne, podejście bayesowskie do szacowania parametrów itd. Wypada więc wyrazić zdziwienie stosunkowo małym zainteresowaniem, jakim się cieszy metoda reprezentacyjna w naszym kraju. Można zaryzykować stwierdzenie, że prawie wyłącznie związana jest z Głównym Urzędem Statystycznym. Gdyby nie profesor R. Zasępa, to przez bez mała 40 lat nie doczekalibyśmy się wyczerpującego polskiego podręcznika, traktującego o metodzie reprezentacyjnej. Z tłumaczeń z tego zakresu wymienić można tylko jedną pozycję V. Barnetta (*Elementy teorii pobierania próby*, PWN, Warszawa 1982) i to nie najszczęśliwiej dobraną. Na ten temat mało

publikuje się artykułów, mało jest też adeptów, którzy bez reszty chcieliby się poświęcić tej dziedzinie wiedzy statystycznej. Na uczelniach ekonomicznych metoda reprezentacyjna także nie cieszy się zbytnią estymą, dowodzi tego choćby wymiar godzin dla zajęć z tego przedmiotu. Ukazuje się niewiele skryptów, traktujących o tej metodzie, dotąd — jak wiem — wydała takowe Szkoła Główna Planowania i Stytystyki oraz Akademia Ekonomiczna w Krakowie.

Jest to zaskakujące, Polska posiada bowiem duże tradycje na tym polu działania naukowego, jak i w praktycznym zastosowaniu tej metody do badań zjawisk społeczno-gospodarczych. Można tu wymienić choćby tylko prof J. Spławę-Neymana, który wraz z E. S. Pearson'em sformułował podstawy klasycznej teorii estymacji, dziś ogólnie akceptowanej, poza może radykalnymi zwolennikami tzw. estymacji bayesowskiej.

Podobnie, jeszcze przed badaniami podjętymi przez „ojca" metody reprezentacyjnej A.N. Kiaera, który w 1895 r. poddał rozpoznaniu robotników norweskich („Sur les méthodes représentatives on typologiques appliqués à la statistique". Bulletin de l'Institute International de Statistique 11, 1915), rozwinął swą działalność prof Wydziału Medycznego Uniwersytetu Jagiellońskiego N. Cybulski. Zorganizował poważne i zakrojone na szeroką skalę badania nad sposobem odżywiania się ludności wiejskiej w ówczesnej Galicji. Podjęte zostały pod auspicjami Wydziału Towarzystwa Opieki Zdrowia w Krakowie w 1891 r. na wniosek N. Cybulskiego oraz H. Jordana. Akcję zainicjowano pod wrażeniem, jakie zrobiła książka S. Szczepanowskiego (*Nędza Galicji w cyfrach i program energicznego rozwoju gospodarstwa krajowego*, Lwów 1888). Odpowiedni kwestionariusz, zawierający 38 pytań, odnoszących się do 3 grup ludności: najzamożniejszych, średnio zamożnych i ubogich został rozesłany do nauczycieli ludowych — ankieterów, mieszkających we wszystkich powiatach administracyjnych. Z każdego powiatu wybranych zostało 10 miejscowości, zwrócono 566 wypełnionych formularzy, co stanowiło 76,5% ogółu rozesłanych. Zebrany materiał został częściowo opracowany dopiero w 1983 r. przez dr K. Baranek z Akademii Ekonomicznej w Krakowie (praca doktorska pt. *Odżywianie się ludności wiejskiej na tle sytuacji społeczno-gospodarczej Galicji w końcu XIX wieku. Analiza statystyczna*, AE, Kraków 1983). Tak więc N. Cybulski może być uznany za jednego z prekursorów metody reprezentacyjnej, tak jak inny Galicjanin C. Ciompa może być zaliczony do grona jednego z pierwszych ekonometryków, a to dzięki książce: *Zarys teorii ekonometryi i teorii buchalteryi* (Towarzystwo Szkoły Handlowej, Lwów 1910). Od tych czasów metody szacowania parametrów rozwijają się w ramach zarówno metody reprezentacyjnej, jak i ekonometrii, będąc ciągle przedmiotem owocnych dociekań teoretycznej i praktycznej natury.

Wybrane problemy badań reprezentacyjnych w światowym piśmiennictwie statystycznym ostatnich lat

doc dr hab. Andrzej Balicki
Uniwersytet Gdański

Licząca już przeszło 90 lat metoda reprezentacyjna (Kiër, 1895) jest ciągle żywą dyscypliną statystyczną. Jej rozwój jest nieustanny i odpowiada stale rosnącemu zapotrzebowaniu na informacje o zjawiskach społecznych, gospodarczych i przyrodniczych. Wyrazem rosnącego zainteresowania badaniami reprezentacyjnymi było powstanie International Association of Survey Statisticians (IASS) oraz odpowiednich sekcji wielu narodowych stowarzyszeń statystyków. Periodycznie odbywają się konferencje i sympozja, które są terenem konfrontacji poglądów statystyków, teoretyków i praktyków oraz forum prezentacji nowych idei i zastosowań praktycznych. Licznie ukazują się publikacje artykułowe i opracowania książkowe oraz sprawozdania z prowadzonych badań. Od 1975 r. wydawany jest przez Statistics Canada, w porozumieniu z IASS, periodyk „The Survey Methodology Journal", poświęcony teoretycznym i praktycznym aspektom badań statystycznych.

Celem niniejszego referatu jest wskazanie na pewne zagadnienia badań reprezentacyjnych, które dominują w piśmiennictwie statystycznym ostatnich lat. Integralną jego częścią jest obszerna bibliografia — przewodnik po literaturze angielskojęzycznej (gł. amerykańskiej, kanadyjskiej i brytyjskiej).

ZAGADNIENIA BADAŃ REPREZENTACYJNYCH WYMAGAJĄCE ROZWOJU

Oceniając postęp w rozwoju teorii, metod i praktyki badań reprezentacyjnych w latach 1968—1978 Horvitz [40] wymienia te aspekty, które szczególnie interesowały statystyków-praktyków:

— modele błędu całkowitego, jako podstawa projektowania badań,
— przybliżone metody obliczania błędów losowych i składowych wariancji, z uwzględnieniem składowych, związanych z błędami nielosowymi,
— analiza danych i wnioskowanie w złożonych projektach badawczych,
— wielokrotne operaty losowania (multiple sampling frames),
— kontrola jakości procedur badawczych,
— projekty badań w czasie (longitudinal) i analiza ich wyników,
— techniki łączenia danych, pochodzących z różnych źródeł dla małych dziedzin (small area).

Mimo stałego udoskonalania metodologii, badania reprezentacyjne pozostawiały i pozostawiają nadal — w opinii praktyków — wiele do życzenia, ze względu na niską jakość projektów czy niedoskonałe sposoby analizowania badanych populacji ludzkich.

Jako dalekosiężny środek poprawy jakości badań reprezentacyjnych Horvitz proponuje informacyjny system projektów badań (sample survey design information system). Miałby on służyć zasadniczo koncepcji tzw. total survey design — TSD, którą można określić, jako metodę wstępnej analizy alternatywnych procedur badawczych, a której wykorzystanie w projektowaniu umożliwi zrównoważoną allokację środków finansowych dostępnych na badanie, minimalizującą ogólny błąd oszacowania (o systemie analizy TSD pisze też Dalenius [14], odnosząc go również do projektowania badań zintegrowanych). Aby stosować TSD potrzebne są modele błędu (total survey error models) [zob. też Lessler, 53]. W ogólności, wobec rozproszonych informacji o składnikach błędu i ich rozmiarach w danych pochodzących z badań, o metodach pomiaru zmiennych czy też kosztach w odniesieniu do różnych populacji i różnych projektów badań reprezentacyjnych, niezbędny jest dla użycia TSD skomputeryzowany system wymiany informacji.

Z kolei Murthy [57], w oparciu o długoletnie doświadczenia, zdobyte w krajach rozwijających się, podaje swoją listę 12 problemów, których rozwiązanie warunkuje efektywne planowanie i prowadzenie badań reprezentacyjnych oraz praktyczną użyteczność wyników. Wymienimy je w kolejności i brzmieniu za oryginałem:

— szacowanie rozkładów częstości, zwłaszcza ogonów,
— skrócone (short-cut) metody szacowania błędu i ich skuteczność,
— pojęciowe problemy w interpretacji błędu szacunków (głównie w przypadku małych prób),
— rozłożenie nakładów na badania w czasie,
— tworzenie szeregów statystycznych (np. czasowych) w oparciu o badania metodą reprezentacyjną,
— problemy integracji badań,
— wykorzystanie badań reprezentacyjnych w inwentaryzacji zasobów naturalnych,
— szeroko zakrojone studia empiryczne nad efektywnością projektów badawczych,
— badanie „time functions" wraz z funkcjami wariancji,
— projektowanie badań analitycznych,
— analiza danych z próby, uwzględniająca błędy,
— wykorzystanie danych z badań reprezentacyjnych do opracowania map statystycznych.

Podano dwie listy problemów, szczególnie ważnych dla praktyki badań reprezentacyjnych i użytkowników ich wyników. Nie są one pełne.

Ciekawe wyliczenie problemów badań reprezentacyjnych prezentuje Kish [49]. Otóż, z punktu widzenia przyszłości, stara się on sformułować problemy, których rozwiązanie jest bardzo potrzebne (needed-N) dla skutecznej realizacji badań reprezentacyjnych. Co do niektórych problemów, można mieć nadzieję, że zostaną rozwiązane mniej lub bardziej zadowalająco w rozsądnie krótkim czasie, np. do 2001 r. (expected-E). Na rozwiązanie innych — z uwagi na stopień trudności — trzeba jeszcze poczekać (waiting-W). Są też problemy, które zapewne zostaną rozwiązane wkrótce (E) lub w dalszej przyszłości (W), a które są zbyteczne, choć nie pozbawione wartości (superfluous-S). Na taką czterosektorową mapę wpisał Kish następujące zagadnienia:

NE: — pospisowe oszacowania dla małych dziedzin (szersza dostępność większych i lepszych prób, teoretyczne wsparcie, wykorzystanie metod bayesowskich, estymator Steina-Jamesa),
— poszerzenie bazy wnioskowania dla subklas (cross-classes) i zwiększenie precyzji oszacowań (rozwój teorii, metody bayesowskie),
— łączenie subiektywnych i obiektywnych oszacowań,
— kryteria poststratyfikacji (adiustacji ważonej),
— zastępowanie spisów badaniami metodą reprezentacyjną,
— unifikowanie teorii badań reprezentacyjnych i eksperymentalnych (problem skończoności realnych populacji, ograniczenia prostych addytywnych modeli liniowych, randomizacja, terminologia),
— metody wnioskowania w badaniach reprezentacyjnych (zbliżenie teorii metody reprezentacyjnej i klasycznej teorii statystyki),
— błędy losowe: obliczanie, przedstawianie, parametry,
— projekty badań wielocelowych,
— selekcja kontrolowana,
— błędy odpowiedzi i braku odpowiedzi,
— rozszerzenie zastosowań badań reprezentacyjnych na nowe obszary.

NW: — teoria rozkładu analitycznych statystyk (np. współczynnika regresji) z prób nieprostych (bez założenia niezależności),
— ogólna teoria dla błędów odpowiedzi,
— ogólne rozwiązanie problemu, rzadkich jednostek (rare el.).

SE: — teoria metody reprezentacyjnej bez zastosowań.

SW: — podstawy (foundations) teorii metody reprezentacyjnej z populacji skończonych.

W przedstawionych listach problemów dominuje podejście pragmatyczne, choć wiadomo, że praktyczne wykorzystanie metody reprezentacyjnej musi opierać się na dobrych podstawach teoretycznych. Różny jest poziom ogólności tych problemów, ale nie było celem referatu ani ich porządkowanie, ani grupowanie. Mamy więc sprzed dziesięciu lat wskazania kierunków dla działań statystyków w zakresie metody reprezentacyjnej.

Każdy sam, według swej najlepszej wiedzy, może wykreślić lub dopisać do rejestru pewne pozycje; może ocenić rzeczywisty postęp w tej dziedzinie.

W dalszej części opracowania wskażemy na pewne tylko osiągnięcia, głównie z zakresu badań nad błędami nielosowymi.

NIELOSOWE BŁĘDY W BADANIU STATYSTYCZNYM

To, że błędy nielosowe mogą mieć dużo większy wpływ na wyniki badań niż błędy losowe jest wiadome od dawna. Uzbierała się już obszerna literatura, traktująca o tych błędach: jak one powstają, jak je kontrolować i jak je mierzyć. Dalenius [13] opracował bibliografię, zasadniczo angielskojęzyczną, na temat nielosowych błędów w badaniach, przyjmując szerokie znaczenie pojęcia badanie, włączając w nie zarówno badania reprezentacyjne, jak i pełne i uwzględniając szeroką różnorodność typów populacji oraz metod gromadzenia danych, jak również nadając szeroki sens pojęciu błędu nielosowego, choć w większości literatura traktuje o błędach pomiaru.

Podana w trzech częściach bibliografia zawiera łącznie 1412 pozycji, które ukazały się do roku 1971. Trudno jest określić, o ile wzbogaciło się piśmiennictwo z tego zakresu przez kolejne 17 lat, ale postęp w badaniach błędów nielosowych oraz estymacji w warunkach występowania błędów jest ogromny (bibliografia późniejszych lat została zamieszczona w książce [42], vol. 2 — poza zasięgiem).

Wśród błędów nielosowych *coraz większe znaczenie zyskuje problem niekompletnych danych.* Systematyczne badania w tym węższym zakresie podjęto nie tak dawno i duża część publikacji o błędach nielosowych, lub w związku z nimi, dotyczy właśnie niekompletnych danych.

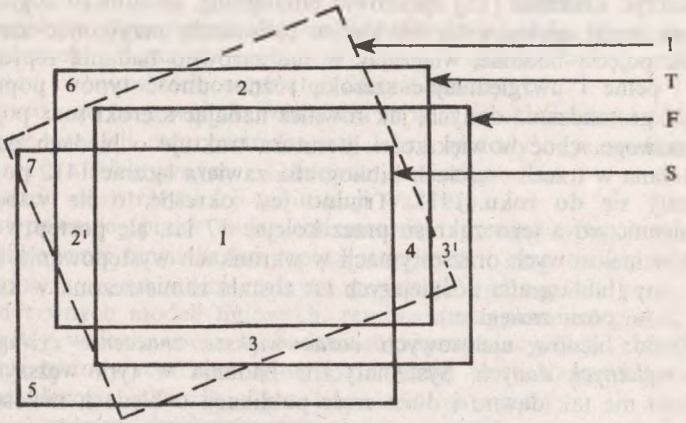
Samo pojęcie błędu, jego kompozycja, źródła i mechanizmy powstawania, kontrola oraz metody badania dokładności danych, jak również sposoby usuwania błędów z danych, były niejednokrotnie już przedmiotem dyskusji. Najpełniejszym przeglądem tych problemów w polskim piśmiennictwie jest książka J. Kordosa [51].

W badaniach nad błędami braku odpowiedzi pewną rolę mogą spełnić schematy, uwypuklające „miejsce" powstawania tych błędów i ich możliwy wpływ na realizację badania. W konstruowaniu takich schematów ważne jest rozróżnienie między „poziomami" populacji. Cochran [9] jasno odróżnił populację celu (target population), która odpowiada idealnej koncepcji badawczej i którą chcielibyśmy badać metodą reprezentacyjną od populacji poddanej losowaniu (sampled population), do której mamy obserwacyjny dostęp za pomocą operatu losowania [zob. też Dalenius, 14]. Dalej poszedł Kish [50], odróżniając 4 poziomy populacji, odpowiadające czterem etapom badania reprezentacyjnego (koncepcja, wybór próby, gromadzenie danych, opracowanie i analiza): populację celu, populację operatu (frame population = = sampled pop.), populację badania (survey population) oraz populację wnioskowania (inference population). Trzecią kategorię populacji, spośród

wymienionych, stanowi zbiór jednostek, które po wylosowaniu mogą być zidentyfikowane i zbadane. Natomiast populacja wnioskowania jest konceptualną populacją jednostek, dla których dane mogłyby stać się osiągalne na etapie tabulacji po zbadaniu i przetworzeniu, jeśliby zostały wybrane lub po adiustacji, jeśli takowa byłaby potrzebna.

Te cztery „poziomy" populacji oddzielone są od siebie różnymi kategoriami nielosowych błędów kompletności (pokrycia i nieuzyskania danych). Murthy [58] przedstawia pewien ideowy wykres 1, który odzwierciedla te różnice i uwydatnia ich przyczyny.

WYKRES 1. RÓŻNICE MIĘDZY POPULACJAMI (WEDŁUG MURTHY)

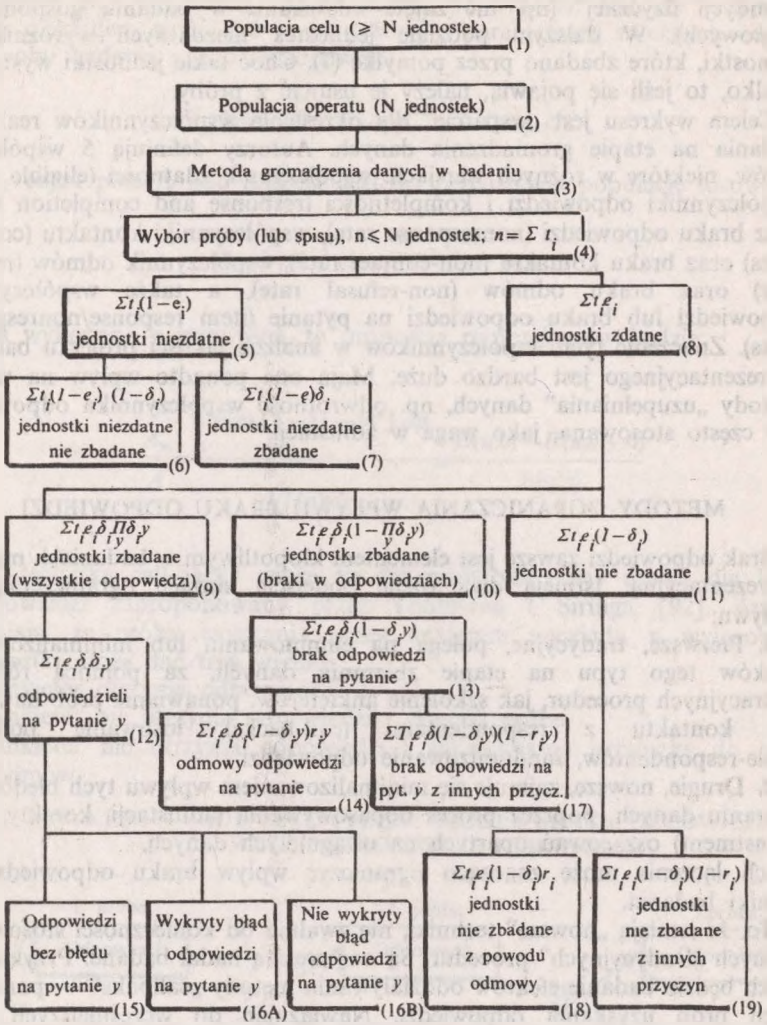


Objaśnienia do wykresu 1.

Populacja celu (*T*), operatu (*F*), badania (*S*), wnioskowania (*I*).
1 — Jednostki w populacjach celu (T) i operatu (F), które mogłyby być zbadane, gdyby zostały wylosowane w konkretnych warunkach badania (*TFS*), 2 — Brakujące jednostki lub niedostępność operatu — brak pokrycia (*TF'S'*). 2 — Niedostępność operatu — lecz jednostki te mogłyby być zbadane dzięki uzupełniającemu operatowi lub specjalnej procedurze badawczej (*TF'S*), 3 — Jednostki uboczne lub powtórzone w operacie, lecz zbadane (*T'FS*), 3 — Jednostki uboczne lub powtórzone w operacie, lecz nie zbadane (*T'FS'*), 4 — Zdatne jednostki w operacie, nie zbadane z różnych przyczyn (nonresponse) (*TFS'*), 5 — Uboczne jednostki poza operatem, lecz zbadane (*T'F'S*), 6 — Brak pokrycia z powodu odrzucenia, imputacji, braku lub niewłaściwego ważenia (*TI'*), 7 — Nadmierne pokrycie, wynikające z niewłaściwego włączenia, imputacji lub nieodpowiedniego ważenia (*T'I*), 2+2' — Zdatne jednostki poza operatem (*TF'*), 3+3 — Uboczne lub dublowane jednostki w operacie (*T'F*), 1+2 — Zdatne jednostki, zbadane (właściwi respondenci) (*TS*), 2+4 — Brak pokrycia i brak kontaktu (nie-respondenci) (TS'), 3+5 — Niewłaściwi respondenci (T'S).

Platek i Gray [65] rozdzielili populację celu na różne kategorie jednostek według typów odpowiedzi lub braku odpowiedzi, opracowując następujący wykres 2. Wyjaśnijmy krótko pewne kwestie — otóż jednostki, które zgodnie z wyborem powinny być zbadane zostały podzielone na dwie klasy: jednostki zdatne, tj. nadające się (ang. eligible, fr. admissible) (8) oraz niezdatne, tj. nie nadające się dla celów badania (5), przy czym kryterium

WYKRES 2. ROZKŁAD POPULACJI WEDŁUG ODPOWIEDZI LUB ICH BRAKU (WEDŁUG PLATKA I GRAYA)



Objaśnienia:

$t_i = 1{,}0$ (jednostka wybrana/nie wybrana do próby),

$e_i = 1{,}0$ (jednostka zdatna/nie zdatna do badania),

$\delta_i = 1{,}0$ (jednostka zbadana/nie zbadana),

$\delta_i y = 1{,}0$ (odpowiedź/brak odpowiedzi na pytanie y u jednostki i),

$r_i = 1{,}0$ (odmowa/inna przyczyna, gł. brak kontaktu).

„przydatności" nie może być określone, jeśli się nie ma kontaktu z jednostką, podczas gdy kiedy indziej jest ono oczywiste z powodu „kondycji fizycznej" (np. nie zajęte mieszkanie w badaniu gospodarstw domowych). W dalszym podziale jednostek niezdatnych wyróżnia się jednostki, które zbadano przez pomyłkę (7). Choć takie jednostki występują rzadko, to jeśli się pojawią, należy je usunąć z próby.

Celem wykresu jest „wsparcie" dla określenia współczynników realizacji badania na etapie gromadzenia danych. Autorzy definiują 5 współczynników, niektóre w różnych wersjach: współczynnik zdatności (eligible rate), współczynniki odpowiedzi i kompletności (response and completion rates) oraz braku odpowiedzi (nonresponse rate), współczynniki kontaktu (contact rates) oraz braku kontaktu (non-contact rate), współczynnik odmów (refusal rate) oraz braku odmów (non-refusal rate), a także współczynniki odpowiedzi lub braku odpowiedzi na pytanie (item response/nonresponse rates). Znaczenie tych współczynników w analizie jakości projektu badania reprezentacyjnego jest bardzo duże. Mają one ponadto wpływ na wybór metody „uzupełniania" danych, np. odwrotność współczynnika odpowiedzi jest często stosowana, jako waga w adiustacji.

METODY OGRANICZANIA WPŁYWU BRAKU ODPOWIEDZI

Brak odpowiedzi zawsze jest elementem kłopotliwym w badaniach metodą reprezentacyjną. Istnieją dwa różne podejścia służące ograniczeniu ich wpływu:

1. Pierwsze, tradycyjne, polega na eliminowaniu lub minimalizowaniu braków tego typu na etapie zbierania danych, za pomocą różnych operacyjnych procedur, jak szkolenie ankieterów, ponawianie prób nawiązania kontaktu z respondentami (call-backs), losowanie podprób z nie-respondentów, randomizowanie odpowiedzi.

2. Drugie, nowsze, zajmuje się minimalizowaniem wpływu tych błędów po zebraniu danych, poprzez proces dopasowywania (adiustacji, korekty, ang. adjustment) oszacowań opartych na osiągniętych danych.

Ich łączenie może znacząco ograniczyć wpływ braku odpowiedzi na wyniki badania.

To, że istnieją „nowsze" techniki, nie zwalnia od konieczności stosowania różnych „tradycyjnych" procedur. Są one zresztą nadal badane. Przykładem niech będzie badanie efektów oddziaływania metody „callbaks" — ponawianych prób uzyskania odpowiedzi. Nawiązując do wcześniejszych prac Politza, Hartleya, Simmonsa i Deminga (1940—1953), Frankela i Dutka [24] opracowali probabilistyczny model oparty na rozkładzie *beta* dla liczby ponowień, które powinny być podjęte, aby uzyskać optymalnie wysoki wskaźnik odpowiedzi i zweryfikowali jego przydatność. W modelu tym zakłada się, że dla danej populacji ludzkiej istnieje pewna utajona funkcja odpowiedzi (ludzie cechują się różnym prawdopodobieństwem podej-

mowania określonych działań — tu: uczestniczenia w badaniu), która jest gęstością rozkładu prawdopodobieństwa odpowiedzi. Autorzy przyjęli, że tym rozkładem może być rozkład *beta*, o parametrach zależnych od celu, sposobu badania i jego wykonawcy:

$$f(p)=p^{u-1}(1-p)^{v-1}, \quad 0<p<1$$

przy czym pole A pod funkcją gęstości przedstawia populację losowania

$$A=\int_0^1 p^{u-1}(1-p)^{v-1}\,dp=B(u,v)$$

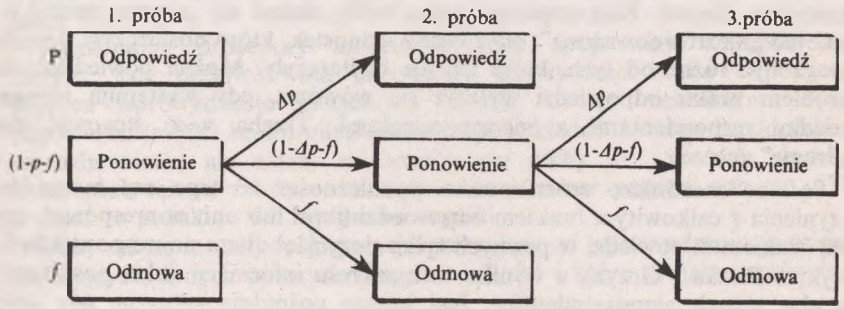
Jeśli wykonano n ponowień, to proporcja respondentów będzie:

$$\frac{A_n}{A}=\frac{\int_0^1 f(p)\,[1-(1-p)^n]\,dp}{\int_0^1 f(p)\,dp}=\frac{B(u,v)-B(u,v+n)}{B(u,v)}$$

Z techniką call-backs wiąże się też probabilistyczny model dla braku odpowiedzi zaproponowany przez Thomsena i Siringa [92]. Autorzy ustalają, że próba uzyskania przez ankietera wywiadu u wylosowanej jednostki może dać trzy wyniki:

1) ankieter otrzyma odpowiedź,
2) ankieter nie otrzyma odpowiedzi i ponowi próbę,
3) ankieter nie otrzyma odpowiedzi i zakwalifikuje jednostkę do grupy odmów.

WYKRES 3. PROBABILISTYCZNY MODEL BRAKU ODPOWIEDZI (WEDŁUG THMOSENA I SIRINGA)



Niech p i f oznaczają prawdopodobieństwa wyniku odpowiednio (1) i (3) w pierwszym podejściu, przy czym przyjmuje się, że w kolejnych próbach f jest stałe, zaś wynik (1) pojawia się z prawdopodobieństwem Δp w próbie drugiej i następnych (Δ może być większe od 1, jeśli ankieter będzie dociekliwy). Można to zilustrować graficznie (wykres 3).

Niech C oznacza zdarzenie, że ankieter otrzyma odpowiedź od wybranej jednostki w C-tej próbie, a zatem:

$$P(C=1)=p,$$
$$P(C=2)=(1-p-f)\,\Delta p,$$
$$P(C=3)=(1-p-\mathrm{f})\,(1-\Delta p-f)\,\Delta p,$$

i ogólnie

$$P(C=\mathrm{c})=\begin{cases} p & \text{jeśli } c=1 \\ (1-p-f)\,(1-\Delta p-f)^{c-2}\,\Delta p & \text{jeśli } c\geqslant 2 \end{cases}$$

Parametry modelu: p, Δ oraz f można oszacować z próby metodą największej wiarygodności (o ile dostępne są informacje o liczbie ponowionych prób w odniesieniu do każdej jednostki).

Model może być uogólniony poprzez zróżnicowanie p między jednostkami; możliwe jest przyjęcie, że p jest generowane przez rozkład *beta* lub też przyjęcie p jako stałej dla pewnych podgrup w populacji itd.

Model służy oczywiście głębszym celom — może być wykorzystany w adiustacji w związku z obciążeniem brakiem odpowiedzi, do badania związku między *MSE* (średni błąd kwadratowy) a liczbą ponowień, a dalej do oceny allokacji środków między pierwszą a kolejnymi próbami uzyskania odpowiedzi (zob. *TSD*).

Mimo efektywnego wykorzystania „tradycyjnych" procedur ograniczania wpływu braku odpowiedzi, zgromadzone dane są tym defektem obciążone. W praktyce bowiem badacz ma ograniczoną możliwość kontrolowania mechanizmu, który determinuje, które jednostki wylosowane do próby dostarczą danych. Niekompletna próba może okazać się „niereprezentatywna" lub „niezrównoważona", gdyż cechy jednostek, które dostarczyły danych mogą być różne od tych, które ich nie dostarczyły. Można powiedzieć, że problem braku odpowiedzi wyłania się wówczas, gdy występują różnice między respondentami a nierespondentami. Trzeba więc stosować też „drugie" metody.

Są one zasadniczo zróżnicowane w zależności od tego czy mamy do czynienia z całkowitym brakiem odpowiedzi (total lub unit nonresponse), czy też brakiem odpowiedzi w pewnych tylko pozycjach (item nonresponse) (zob. wykres Platka i Graya), a wynika to z zakresu informacji, jakie posiadamy o obu typach nierespondentów. Jest jeszcze pośrednia sytuacja, gdy ilość

brakujących informacji, w odniesieniu do jednostki, jest duża (partial nonresponse) i wówczas wybór metody nie jest oczywisty.

Redukowanie obciążenia powodowanego brakiem odpowiedzi (nonresponse bias) objęte jest terminem „adiustacji", a konkretne metody stosowane w tym celu powinny uwzględniać różnice cech respondentów i nie-respondentów. I choć termin ten jest ogólny, to w literaturze stosowany bywa często w związku z całkowitym brakiem odpowiedzi (nonresponse adjustment, adjustment for nonresponse), zaś dla metod traktowania braku odpowiedzi dla pewnych pozycji zarezerwowano termin „imputacji" (imputation).

Metody adiustacji są omawiane i analizowane w wielu opracowaniach. Ogólny przegląd procedur dają prace [3, 7, 8, 43, 44, 47, 51]. W teorii i praktyce dominują adiustacyjne metody ważenia, których istota polega na zwiększeniu wag określonych grup respondentów tak, żeby reprezentowali oni także nierespondentów. Wymagają one dodatkowych informacji bądź to o nierespondentach, bądź o całej populacji. Stosuje się najczęściej następujące procedury:

— ważenie według populacji (poststratyfikacja),
— ważenie według próby [3, 8],
— (interacyjna) metoda ilorazów (raking ratio adjustment) [59,60,62],
— ważenie według prawdopodobieństw odpowiedzi (metody typu Politza-Simmonsa [16, 17]).

Różne problemy związane z metodami ważenia wymagają jeszcze opracowania. Statystycy starają się odpowiedzieć na wiele pytań: jak wybór zmiennych do zdefiniowania klas ważenia (weighting classes lub adjustment cells) wpływa na błąd całkowity, a w tym na obciążenie brakiem odpowiedzi i wariancję losową i jaką zasadą należy się kierować wybierając te zmienne adiustacyjne?, jak optymalizować wielkość klas ważenia? i wiele innych [por. Kish, 49].

Punktem wyjścia dla rozważań nad tym zagadnieniem było określenie wzorów na obciążenie i wariancję estymatora (sumy), przy założeniu adiustacji [Platek, Singh, Tremblay, 97]. Próby zaś odpowiedzi na postawione pytania są do odnalezienia w publikacjach [3, 62, 7] oraz z punktu widzenia praktyki w [42] vol. 1. Kontynuację tych badań stanowi praca [53,93].

Wymienione procedury adiustacyjne nie „poprawiają" bezpośrednio danych, lecz przystosowują estymatory do sytuacji braku odpowiedzi, poprzez wprowadzenie specjalnych założeń co do mechanizmu braku odpowiedzi (model braku odpowiedzi może zakładać, że prawdopodobieństwo odpowiedzi nie zależy od szacowanej cechy — ignorable response pattern lub też, że zależy — nonignorable response pattern; zob. np. [30], a przez to redukują ich wpływ na wyniki.

Służą zaś temu metody imputacji, polegające na wprowadzeniu w miejsce brakujących danych pewnych innych wielkości i które — jak wspomniano

— stosowane są w sytuacji braku odpowiedzi tylko w niektórych pozycjach. Posługując się predykcyjnym ujęciem Rubina [78] można ideę imputacji opisać następująco:

Określony jest łączny rozkład $f(X_1, ..., X_N)$, wyrażający statystyczne zachowanie populacji pełnych rekordów (ciągów informacji, dotyczących jednej jednostki) zarówno w odniesieniu do pojedynczych cech ilościowych, jak i jakościowych. Bez uszczerbku dla ogólności, dla rekordu, który wymaga imputacji, N zmiennych można podzielić następująco: $X_1, ..., X_{m_i}$, które wymagają uzupełnienia oraz $X_{m_{i+1}}, ..., X_N$, dla których nie jest ona potrzebna. Można więc otrzymać rozkład warunkowy $f(X_1, ..., X_{m_i} | x_{m_{i+1}}, ..., x_N)$. Imputowane wartości $y_1, ..., y_{m_i}$ są wybierane dla $X_1, ..., X_{m_i}$ ze zbioru:

$$\{y_1, ..., y_{m_i} : f(y_1, ..., y_{m_i} | x_{m_{i+1}}, ..., x_N) > 0\}$$

Wyboru można dokonać stosując różne procedury.

Imputacja jest zagadnieniem dość często dyskutowanym w ostatnim czasie, o czym zaświadcza niemała liczba publikacji [1, 2, 3, 7, 8, 10, 15, 16, 35, 44, 45, 46 i in.], tym niemniej jest to ciągle problem otwarty; niewiele bowiem osiągnięto w zakresie krytycznej oceny procedur i ich porównania, a jeszcze mniej — z powodu matematycznych komplikacji — w dziedzinie analitycznych ocen alternatywnych procedur (wpływ na obciążenie i wariancję estymatorów).

Ostatnio wiele miejsca w literaturze zajmują procedury: „hot-deck" oraz imputacji wielokrotnej. Imputacyjne procedury „hot-deck" są pewną klasą procedur, polegających na tym, że niepełna odpowiedź jest kompletowana przy wykorzystywaniu wartości z jednej lub większej liczby kompletnych zapisów w tym samym zbiorze (tj. z tego samego badania — the hot deck; „hot" podkreśla użycie danych z tej samej próby, w odróżnieniu od informacji z innych źródeł — cold deck), przy czym wybór tych zapisów zmienia się wraz z rekordem, wymagającym imputacji. Wprowadzenie np. średniej z osiągniętych danych na „wakujące" pole nie jest procedurą hot-deck, gdyż wybór użytych rekordów do obliczenia średniej jest niezależny od rekordu kompletowanego). Powszechniej stosowane lub próbowane procedury, to: sekwencyjna, hierarchiczna, wyboru losowego oraz najbliższego sąsiada, a wybór metody (także spośród innych) uzależniony jest od typu zmiennej, dla której imputację się przeprowadza.

Nie ma niestety precyzyjnej definicji tej procedury (zob. [41]). Nieliczne są też teoretyczne badania nad procedurami hot-deck, np. wpływu na *MSE* [2, 3, 12, 16, 70].

Pomimo wielu korzyści, pojedyncza lub deterministyczna imputacja (zastąpienie brakującej informacji jedną inną) ma poważną wadę, polegającą

na tym, że zastosowanie do tak poprawionych danych standardowych metod analizy powoduje, że brakujące dane traktuje się tak, jakby były znane, a przez to dostarczone wyniki są zbyt „ostre", gdyż nie uwzględniają zmienności, pochodzącej z nieznanych brakujących wartości.

Tak rozumując, Rubin [79] zaproponował imputację wielokrotną (multiple imputation), polegającą na tworzeniu dwu lub więcej ciągów uzupełnionych danych, tak że można przeprowadzić dwie lub więcej analiz, po jednej dla każdego ciągu. Jeżeli przy danym modelu błędu (dotyczący możliwego zachowania brakującej wartości), wykonano imputacje wielokrotnie, to wnioski, które prawidłowo oświetlą zmienność powodowaną przez nieznane wartości, można wyciągnąć poprzez łączenie oddzielnych wniosków ze skompletowanych danych w sposób, który uwzględni zarówno zmienność wewnątrzimputacyjną, jak i międzyimputacyjną.

Załóżmy, że Θ jest parametrem, który chcemy szacować oraz że w próbie, na skutek braku odpowiedzi, obserwowano jedynie n_1 wartości spośród możliwych n, $Y=(Y_1, ..., Y_n)$, $(n_1 < n)$. Powtarzamy m-krotnie właściwą imputację (tzn. oddającą losową zmienność zgodnie z modelem) dla każdej brakującej wartości, otrzymując w ten sposób m skompletowanych ciągów danych, a dalej m oszacowań $\hat{\Theta}_i$ i odpowiadających im wariancji U_i $(i=1, ..., m)$. Ostatecznym estymatorem parametru Θ jest $\bar{\Theta}_m = \frac{1}{m}\sum_{i=1}^{m} \hat{\Theta}_i$, estymatorem zaś jego wariancji $T = \bar{U}_m + \frac{m+1}{m} B_m$,

gdzie: $\bar{U}_m = \frac{1}{m}\sum_{i=1}^{m} U_i$ jest średnią wariancją wewnątrzimputacyjną oraz

$B_m = \frac{1}{m+1}\sum_{i=1}^{m} (\hat{\Theta}_i - \bar{\Theta}_m)^2$ jest wariancją międzyimputacyjną.

Imputacja wielokrotna nie jest kolejną procedurą, konkurencyjną lub uzupełniającą w stosunku do technik imputacji pojedynczej, lecz stanowi jakościowo nowe podejście traktowania braków odpowiedzi w badaniach, gdyż pozwala na odzwierciedlenie niepewności tkwiącej w brakujących danych.

Bayesowskie uzasadnienie dla imputacji wielokrotnej, dalsze szczegóły teorii i wyniki analitycznych badań można znaleźć w nielicznych jeszcze publikacjach [35, 79, 80, 82, 83, 84, 85].

Pewne sposoby rozwiązania problemu częściowych odpowiedzi (co się zdarza w przypadku bardzo obszernych kwestionariuszy) podali Hocking i Smith [37], uogólnione ostatnio przez Hockinga [36].

MODELOWE PODEJŚCIE W METODZIE REPREZENTACYJNEJ

W ostatnich kilkunastu latach ukazało się wiele artykułów, dotyczących podstaw badania metodą reprezentacyjną, sugerujących zmiany w statystycznym modelu próbkowania probabilistycznego, które można nazwać podejściem randomizacyjnym (randomization lub design-based approach). Przypomnijmy, że polega ono na tym, że próby są wybierane z populacji skończonych według procedur probabilistycznych, wprowadzanych jako część projektu badawczego i w których estymatory odzwierciedlają prawdopodobieństwa wyboru w taki sposób, że można otrzymać i praktycznie stosować zgodne estymatory średniego błędu kwadratowego. Tak więc tradycyjna teoria metody reprezentacyjnej opiera się na rozkładzie prawdopodobieństwa tworzonym przez losowy mechanizm wyboru jednostek; dla danej skończonej populacji i ustalonej liczebności próby, przy określonym planie próbkowania, obciążenie i wariancja estymatora są definiowane w kategoriach przeciętnych ze wszystkich możliwych prób i są stałe, tzn. nie zależą od aktualnie obserwowanej próby. Wartości populacyjne są w tym podejściu traktowane jako stałe. Podejście to nie wymaga zatem wprowadzenia żadnego założonego modelu, poza wykorzystaniem prawdopodobieństw przyjętych w procedurze wyboru, a jedynym założeniem jest duża liczebność próby, przy której oszacowania z próby miałyby rozkład w przybliżeniu normalny.

W nowym podejściu, które można nazwać modelowym (model-based approach) wartości populacyjne traktowane są jako realizacje zmiennych losowych, które rozkładają się zgodnie z pewnym modelem. Taki model rozkładu tworzy podstawę wnioskowania, a procedura wyboru próby gra podrzędną rolę, głównie dla uniknięcia obciążenia wyboru.

Uwypuklijmy to rozróżnienie za Littlem [54] i przyjmijmy, że realizujemy badanie z wykorzystaniem schematu losowania prostego bez zwracania (simple random sampling). Niech N oznacza wielkość populacji, x_i — wartość zmiennej x dla jednostki i $(i=1, ..., N)$ i niech $\bar{X}=\sum_{i=1}^{N} x_i/N$ oznacza średnią w populacji. Zdefiniujmy funkcję wskaźnikową δ, gdzie $\delta_i=1$ jeśli jednostka i trafi do próby oraz $\delta_i=0$ w przeciwnym razie. Niech w końcu $n=\sum_{i=1}^{N} \delta_i$ będzie wielkością próby.

Wnioskowanie randomizacyjne opiera się na średniej próbkowej: $\bar{x}=\sum_{i=1}^{N} \delta_i x_i/n$, traktowanej jako losowa funkcja wskaźników $\delta=(\delta_1, ..., \delta_N)$ ze stałymi wartościami populacyjnymi $x_1, ..., x_N$. Dla umiarkowanie dużych

prób, $\bar{x}=\bar{x}(\delta)$ ma rozkład normalny ze średnią $\bar{X}$ i wariancją $(1-n/N)S^2/n$, gdzie S^2 jest populacyjną wariancją x_i. Za pomocą symboli zapiszemy:

$$\bar{x}(\delta)\,|\,(x_1, \ldots, x_N) \sim N\left(\bar{X}, \left(1-\frac{n}{N}\right)\frac{S^2}{n}\right)$$

Ekwiwalentne wnioskowanie oparte na modelu, dla tego problemu, traktuje wartości x_i jako niezależne zmienne losowe o jednakowym rozkładzie normalnym ze średnią μ i wariancją σ^2, tj. $x_i \sim N(\mu, \sigma^2)$ $(i=1, \ldots, N)$, gdzie μ oraz σ^2są parametrami superpopulacji, w przeciwieństwie do $\bar{X}$ i S^2, które są wielkościami populacji (parametrami badanej, realnej populacji). Dalsze ujęcie jest dwuwariantowe:

— w podejściu Royalla [71] wnioskowanie opiera się na własnościach estymatora średniej $\bar{X}$ (np. $\bar{x}$) w powtarzalnym losowaniu z rozkładu superpopulacji,

— w bayesowskim podejściu Ericsona [78] określa się rozkłady a priori dla μ oraz σ^2 i wnioskowanie opiera się na posteriorycznym rozkładzie $\bar{X}$ przy danej próbie. W podejściu tym zakładając, że próby są umiarkowanie duże, a rozkłady a priori dla μ i σ^2 rozproszone (nieinformacyjne), uzyskany rozkład a posteriori dla μ jest rozkładem normalnym $\mu\,|\,\text{próba} \sim N(\bar{x}, \frac{s^2}{n})$ oraz podobnie dla $\bar{X}$:

$\bar{X}\,|\,\text{próba} \sim N\left[\bar{x}\left(1-\frac{n}{N}\right)\frac{s^2}{n}\right]$ (w podejściu Royalla s^2 należy zastąpić przez *MSE*).

Dla *lpbz* podejście tradycyjne i modelowe w oparciu o rozkład normalny dają ten sam wynik, ale tak nie jest przy bardziej złożonych planach próbkowania.

Przyjrzyjmy się zresztą bliżej podejściu Royalla [71, 72], które nazwiemy krótko: predykcyjnym (the linear least squares prediction approach), w którym centralną rolę odgrywają estymatory złożone, główne ilorazowe, na przykładzie szacowania sumy za pomocą zwykłego estymatora ilorazowego w schemacie *LPBZ*.

Niech y będzie badaną zmienną, której wartości są w przybliżeniu proporcjonalne do wartości dodatnio określonej, pomocniczej zmiennej x. Przyjmijmy, że zmienność y wzrasta wraz ze wzrostem x. Wartości zmiennej x dla wszystkich jednostek populacji są znane (taką zmienną nazywamy często zmienną dodatkową), zaś wartości y — tylko dla n elementów w próbie. Wartości $y_1, \ldots, y_N$ traktujemy jako realizacje zmiennych losowych $Y_1, \ldots, Y_N$. Zapiszmy populacyjną sumę jako: $T=\Sigma_s y_i+\Sigma_{\bar{s}}\, y_i$, czyli jako znaną sumę dla jednostek w próbie (s) plus nieznaną sumę jednostek poza próbą ($\bar{s}$). Szacowanie T jest więc równoważne predykcji sumy nie obserwowanych zmiennych losowych $\{y\colon i\in\bar{s}\}$.

Probabilistyczny model, oddający podstawową strukturę badanej populacji, możemy formalnie zapisać:

$$EY_i = \beta x_i, \quad VarY_i = \sigma^2 x_i, \quad Cov(Y_i, Y_j) = 0 \quad (i \neq j)$$

Najlepszym, przy tym modelu, nie obciążonym liniowym estymatorem sumy T jest estymator ilorazowy: $\hat{T}_s = N\bar{x}\bar{y}_s/\bar{x}_s$, gdzie $\bar{x}$ jest średnią populacyjną, a $\bar{x}_s$ i $\bar{y}_s$ są średnimi w próbie n-elementowej. Zmienność oceny $\hat{T}_s$ mierzona wariancją błędu (lub MSE) jest:

$$E(\hat{T}_s - T)^2 = \frac{N}{n}(N-n)\sigma^2 \bar{x}_{\bar{s}}\bar{x}/\bar{x}_s$$

gdzie: $\bar{x}_{\bar{s}}$ jest średnią jednostek poza próbą i maleje ze wzrostem $\bar{x}_s$, co oznacza, że optymalną (tj. minimalizującą MSE) jest taka próba, która się składa z n jednostek, których wartości x są największe.

Jeśli założony model nie jest spełniony, tj. $EY_i = h_i$, to estymator ilorazowy ma obciążenie:

$$B = E(\hat{T}_s - T) = (N-n)\left[\frac{\bar{x}_{\bar{s}}}{\bar{x}_s}\bar{h}_s - \bar{h}_{\bar{s}}\right]$$

Na przykład, jeśli w rzeczywistości $EY_i = \beta_0 + \beta x_i$, to obciążenie wyniesie:

$$B = (N-n)\beta_0\left[\frac{\bar{x}_{\bar{s}}}{\bar{x}_s} - 1\right] = N\beta_0\frac{(\bar{x} - \bar{x}_s)}{\bar{x}_s}$$

i znika, jeśli próba jest zrównoważona ze względu na x ($\bar{x}_s = \bar{x}$), lecz może być duże w przypadku próby źle zrównoważonej.

Dalsze szczegóły badania własności estymatora można znaleźć w późniejszych pracach [39, 74, 75, 76, 67].

Teoria randomizacji jest atrakcyjna przez to, że pozwala uniknąć określenia modelu. Rozkład prawdopodobieństwa, który jest jej podstawą jest znany, podczas gdy model wprowadza element subiektywny. Niemniej traci ona swą obiektywność, gdy pojawią się odstępstwa od probabilistycznego modelu wyboru próby, na skutek błędów pokrycia i braku odpowiedzi i jedynie założenie o losowym charakterze tego typu błędów może obronić tę obiektywność.

Istnieją więc i rozwijają się dwa konkurencyjne podejścia: tradycyjne i modelowe, a w ramach tego drugiego znowu dwa: bayesowskie

i niebayesowskie. Brewer i Särndal [6] wyliczają nawet sześć różnych podejść.

Pierwszy artykuł na temat „podstaw" badania reprezentacyjnego napisał Godambe [26]. Spośród wielu innych wymienić należy: Basu [4], Royalla [72] oraz Smitha [89], do których artykułów najczęściej odnoszą się statystycy. Zasadniczą ideą tych artykułów było ponowne sformułowanie statystycznego modelu badań opartych na wyborze statystycznym tak, aby był dostosowany do ogólnego modelu wnioskowania statystycznego. Wymienione artykuły zainicjowały szybki rozwój teorii i metod opartych na zmienionych podstawach, ale też wywołały nie kończącą się dyskusję między zwolennikami różnych podejść. Dalsze pozycje bibliografii traktujące o „podstawach" — teorii oraz metodach i zastosowaniach, to [31, 32, 5, 21, 28, 27, 38, 43, 90]. Nowe podejście objęło też takie aspekty badań, jak: błędy nielosowe [20, 29, 35, 55, 68, 88] czy szacowanie dla małych dziedzin [68].

BIBLIOGRAFIA

[1] Bailar B.A., Bailar III J.C. (1983): *Comparision of the biases of the hot-deck imputation procedure with an „equal-weights" imputation procedure*, w: Incomplete Data in Sample Surveys, vol. 3, op. cit.: 299—311

[2] Bailar III J.C., Bailar B.A. (1978): *Comparision of two procedures for imputing missing survey values*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 462—467

[3] Bailar B.A., Bailey L., Corby C. (1978): *A comparision of some adjustment and weighting procedures for survey data*, w: Survey Sampling and Measurement, op. cit.: 175—198

[4] Basu D. (1971): *An essay on the logical foundations of survey sampling*. Part I, w: Foundations of Statistical Inference (eds. V.P. Godambe and D.A. Sprott), New York, Holt: 203—242

[5] Basu D. (1978): *Relevance of randomization in data analysis*, w: Survey Sampling and Measurement, op. cit.: 267—292

[6] Brewer K.R.W., Särndal C.E. (1983): *Six approaches to enumerative survey sampling (with discussion)*, w: Incomplete Data in Sample Surveys, vol. 3, op.cit.: 363—405 (zawiera obszerną bibliografię)

[7] Chapman D.W. (1983): *An investigation of nonresponse imputation procedure for the health and nutrition examination survey*, w: Incomplete Data in Sample Survey, vol. 1, op.cit.: 435—483

[8] Chapman D.W., Bailey L., Kasprzyk D. (1986): *Nonresponse adjustment procedures at the US Census Bureau*, Survey Methodology, 12(2)

[9] Cochran W.G. (1963): *Sampling Techniques* (2rd ed.), New York, Wiley

[10] Colledge M.J., Johnson J.H., Paré R., Sande I.G. (1978): *Large scale imputation of survey data*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 431—436

[11] Cox B.G., Cohen S.B. (1985): *Methodological Issues for Health Care Survey*, New York, Marcel Dekker

[12] Cox B.G., Folsom R.E. (1978): *An empirical investigation of alternate item nonresponse adjustments*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 219—223
[13] Dalenius T. (1977): *Bibliography on non-sampling errors in surveys*, Int. Stat. Rev., 45: I (A—G); 71—89, II (H—G): 181—197, III (R—Z): 303—317
[14] Dalenius T., (1986). *Elements of Survey Sampling*, Stockholm, SAREC
[15] David M., Little R.J.A., Samuhel M.E., Triest R K. (1986): *Alternative methods for CPS income imputation*, J. Am. Stat. Assoc., 81: 29—41
[16] Drew J.H., Fuller W.A. (1980): *Modelling nonresponse in surveys with callbacs*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc., 639—642
[17] Drew J.H., Fuller W.A. (1981): *Nonresponse in complex multiphase surveys*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 623—628
[18] Ericson W.A. (1969): *Subjective bayesian models in sampling finite populations I*, J. Roy. Stat. Soc. B, 31: 195—234
[19] Ernst L.F. (1978): *Weighting to adjust for partial nonresponse*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 468—473
[20] Fienberg S.E., Tanur J.M. (1986): *The design and analysis of longitudinal surveys: controversies and issues of coast and continuity*, w: Survey Research Design (eds. R. Boruch and R. Pearson), New York, Springer—Verlag: 60—93
[21] Fienberg S.E., Tanur J.M. (1987): *Experimental and sampling structures: parallels diverging and meeting*, Int. Stat. Rev., 55(1): 75—96 (zawiera obszerną bibliografię)
[22] Ford B.L. (1976): *Missing data procedures: a comparative study*, Proc. Soc. Stat. Sec. Am. Stat. Assoc.
[23] Ford B.L. (1983): *An overview of hot-deck procedures*, w: Incomplete Data in Sample Surveys, vol. 2, op. cit.: 185—207
[24] Frankel L.R., Dutka S. (1983): *Survey design in anticipation of nonresponse and imputation*, w: Incomplete Data in Sample Surveys, vol. 3, op. cit.: 69—83
[25] Giles P., Patrick C. (1986): *Imputation options in a generalized edit and imputation system*, Survey Methodology, 12(1): 49—60
[26] Godambe V.P. (1955): *A unified theory of sampling from finite populations*, J. Roy. Stat. Soc. B, 17: 269—278
[27] Godambe V.P. (1966): *A new approach to sampling from finite populations*, J. Roy. Stat. Soc. B 28: 310—328
[28] Godambe V.P. (1982): *Estimation in survey sampling: robustness and optimality*, J. Am. Stat. Assoc., 77: 393—406
[29] Godambe V.P., Thompson M.E. (1986): *Some optimality results in the presence of nonresponse*, Survey Methodology, 12(1): 29—36
[30] Greenlees W.S., Reece J.S., Zieschang K.G. (1982): *Imputation of missing values when the probability of response depends on the variable being imputed*, J.Am. Stat. Assoc., 77: 251—261
[31] Hansen M.H., Madow W.G. (1978): *Estimation and inferences from sample surveys: Some comments on recent developments*, w: Survey Sampling and Measurement, op. cit.: 341—357
[32] Hansen M.H., Madow W.G., Tepping B.J. (1983): *An evaluation of model-dependent and probability-sampling inferences in sample surveys (with discussion)*, J. Am. Stat. Assoc. 78: 776—793
[33] Hauck M. (1974): *Planning field operations*, w: Handbook of Marketing Research (ed. P. Ferber), New York, McGraw-Hill: 147—159

[34] Herson J. (1976): *An investigation of relative efficiency of least square prediction to conventional probability sampling plans*, J. Am. Stat. Assoc., 71: 700—703
[35] Herzog T.N., Rubin D.B. (1983): *Using multiple imputation to handle nonresponse in sample surveys*, w: Incomplete Data in Sample Surveys, vol. 2, op.cit.: 209—245
[36] Hocking R.R. (1983): *The design and analysis of sample surveys with incomplete data. Reduction of respondent burden*, w: Incomplete Data in Sample Surveys, op. cit.: 107—124
[37] Hocking R.R., Smith W.B. (1972): *Optimum incomplete multinormal samples*, Technometrics, 14(2): 299—308
[38] Hoem J.M. (*1985*): *Weighting, misclassification, and other issues in the analysis of survey samples of life histories*, w: Longitudinal Analysis of Labor Market Data (eds. J.J. Heckman and B. Singer), Cambridge Univ. Press: 249—293
[39] Horn S.D., Horn R.A., Duncan D.B. (1976): *Estimating heteroscedastic variances in linear models*, J. Am. Stat. Assoc., 70: 380—385
[40] Horvitz D.G. (1978): *Some design issues in sample surveys*, w: Survey Sampling ang Measurement, op. cit.: 3—11
[41] *Hot-deck procedures: Introduction and Discussion*, w: Incomplete Data in Sample Surveys, vol. 3, op. cit.: 351—360
[42] *Incomplete Data in Sample Surveys* (1983). Eds. W.G. Madow and J. Olkin: vol. 1. *Report and Case Studies*, vol. 2. *Theory and Bibliographies;* vol. 3. *Proceedings of the Symposium*, New York, Academic Press
[43] Kalton G. (1983a): *Models in the practice of survey sampling*, Int. Stat. Rev., 51: 175—188
[44] Kalton G. (1983b): *Compensating for Missing Survey Data, Ann Arbor, Survey Research Center*, Univ. of Michigan
[45] Kalton G. (1986): *Handling wave nonresponse in panel surveys*, Journal of Official Stat., 2:
[46] Kalton G., Kasprzyk D. (1982): *Imputing for missing survey response*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 22—31
[47] Kalton G., Kasprzyk D. (1986): *The treatment of missing survey data*, Survey Methodology, 12(1): 1—16
[48] Kalton G., Kish L. (1984): *Some efficient random imputation methods*, Communications in Statistics — Theory and Methods, 13(16): 1919—1939
[49] Kish L. (1978): *On the future if survey sampling*, w: Survey Sampling and Measurement, op.cit.: 13—21
[50] Kish L. (1979): *Populations for survey sampling*, Survey Statistician, 1: 14—15
[51] Kordos J. (1987): *Dokładność danych w badaniach społecznych*, Biblioteka Wiadomości Statystycznych Tom 35, Warszawa 1987 r.
[52] Kvitz F.J. (1977): *Toward a standard definition of response rate*, Public Opinion Quarterly, Summer: 265—267
[53] Lessler J.T. (1983): *An expanded survey error model*, w: Incomplete Data in Sample Surveys, vol. 3, op.cit.: 259—270
[54] Little R.J.A. (1982): *Models for nonresponse in sample surveys*, J. Am. Stat. Assoc., 77: 237—250
[55] Little R.J.A. (1983): *Superpopulation models for nonresponse*, w: Incomplete Data in Sample Surveys, vol. 2, op. cit., part VI
[56] Little R.J.A. (1986): *Survey nonresponse adjustments for estimates of means*, Int. Stat. Rev., 54: 139—157

[57.] Murthy M.N. (1978). *Use of sample surveyes in national planning in developing countries*, w: Survey Sampling and Measurement, opt. cit.: 231—253
[58.] Murthy M.N. (1983). *A framework for studying incomplete data*, w: Incomplete Data in Sample Surveys, vol. 3, op.cit.: 7—24
[59.] Oh H.L., Scheuren F. (1978a): *Multivariate raking ratio estimation in the 1973 exact match study*, Proc. Survey Res. Meth. Sec. Am. Stat. Asscoc.: 716—722 (zawiera obszerną bibliografię)
[60.] Oh H.L., Scheuren F. (1978b): *Some unresolved application issues in raking ratio estimation*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 723—728
[61.) Oh H.L.,Scheuren F. (1980): *Estimating the variance impact of missing CPS income data*, Proc.Survey Res. Meth. Sec. Am. Stat. Assoc.: 408—415
[62.] Oh H.L., Scheuren F. (1983): *Weighting adjustment for unit nonresponse*, w: Incomplete Data in Sample Surveys, vol.2, op.cit.: 143—184
[63.] Oh H.L., Scheuren F., Nisselson H. (1980): *Differential bias impacts of alternative Census Bureau hot deck procedures for imputing missing CPS income data*, Proc. Survey Res. Meth. Sec. Am. Stat. Assoc.: 416—420
[64.] Platek R., Gray G.B. (1978): *Non-response and imputation*, Survey Methodology, 4(2)
[65.] Platek R., Gray G.B. (1985): *Some aspect of nonresponse adjustment*, Survey Methodology, 11: 1—14
[66.] Platek R., Gray G.B. (1986). *On the definitions of response rate*, Survey Methodology, 12(1): 17—27
[67.] Rao J.N.K. (1978): *Comments on paper by Basu and Royall and Cumberland*, w: Surveys Sampling and Measurement, op.cit.: 323—329
[68.] Rao J.N.K., Hughes E. (1983): *Comparison of domains in the presence of nonresponse*, w: Incomplete Data in Sample Surveys, vol.3, op.cit.: 215—226
[69.] Report (part I) (1983). w: Incomplete Data in Sample Surveys, vol. 1, op.cit: 3—103
[70.] Rizvi M.H. (1983): *An empirical investigation on some item nonresponse adjustment procedures*, w: Incomplete Data in Sample Surveys. vol.1, op.cit., part II, chpt. 8
[71.] Royall R.M. (1970): *On finite population sampling theory under certain linear regression modeles*, Biometrics, 57: 377—387
[72.] Royall R.M. (1971): *Linear regression models in finite population sampling theory*, w: Foundations of Statistical Inference (eds V.P. Godambe and A. Sprott), New York, Holt
[73.] Royall R.M. (1976): *Current advances in sampling theory: Implications for human observational studies*, American Journal of Epidemiology, 104(4): 463—473
[74.] Royall R.M., Cumberland W.G. (1978a): *An empirical study of prediction theory in finite population sampling: Simple random sampling and the ratio estimator*, w: Survey Sampling and Measurement, op.cit.:
[75.] Royall R.M., Cumberland W.G. (1978b): *Variance estimation in finite populations*, J. Am. Stat. Assoc., 73:
[76.] Royall R.M., Eberhardt K.R. (1975): *Variance estimates for the ratio estimator*, Sankhya C, 37: 43—52
[77.] Royall R.M., Herson J. (1973): *Robust estimation in finite population*, J.Am.Stat.Assoc., 68: 880—889
[78.] Rubin D.B. (1976): *Inference and missing data*, Biometrika, 63: 581—592

[79.] Rubin D.B. (1978): *Multiple imputations in sample survey. A phenomenological bayesian approach to nonresponse*, Proc.Survey Res.Meth.Sec. Am.Stat.Assoc.: 20—34
[80.] Rubin D.B. (1979): *Ilustrating the use of multiple imputations to handle nonresponse in sample surveys*, Bull.Int.Stat.Inst., 48(2): 517—532
[81.] Rubin D.B. (1983): *Conceptual issues in the presence of nonresponse*, w: Incomplete Data in Wample Surveys, vol.2, op.cit.: part IV/12
[82.] Rubin D.B. (1986a): *Basic ideas of multiple imputation for nonresponse*, Survey Methodology, 12(1): 37—47
[83.] Rubin D.B. (1986b): *Statistical matching using file concatenation with adjusted weight and multiple imputations*, Journal of Business and Econ. Stat.,: 87—94
[84.] Rubin D.B. (1986c): *Multiple Imputaion for Nonresponse in Surveys*, New York, Wiley
[85.] Rubin D.B., Schenker N. (1986): *Multiple imputation for interval estimation from simple random samples with ignorable nonresponse*, J.Am.Stat.Assoc., 81: 366—374
[86.] Sande I.G. (1983): *Hot-deck imputation procedures*, w: Incomplete Data in Sample Surveys, vol.3, op.cit.: 339—349
[87.] Santos R.L. (1981): *Effects of imputation on regression coefficients*, Proc.Survey Res.Meth.Sec. Am.Stat.Assoc.: 140—145
[88.] Singh B., Sedransk J. (1983): *Bayesian procedures for survey design when there is nonresponse*, w: Incomplete Data in Sample Surveys, vol.3, op.cit.: 227—247
[89.] Smith T.M.F. (1976): *The foundations of survey sampling: A review*, J.Roy.Stat.Soc. A 139: 183—204
[90.] Smith T.M.F. (1983): *On the validity of inferences from non-random samples*, J.Roy.Stat.Soc. A 146: 394—403
[91.] *Survey Sampling and Measurement* (1978). Ed. N.K.Namboodiri, New York, Academic Press
[92.] Thomsen I., Siring E. (1983): *On the causes and effects of nonresponse*, w: Incomplete Data in Sample Surveys, vol.3, op.cit.: 25—59
[93.] Tremblay V. (1986): *Practical criteria for definition of weighting classes*, Survey Methodology, 12(1): 85—97
[94.] Vacek P.M., Ashikaga T. (1980): *An examination of the nearest neighbor rule for imputing missing values*, Proc.Stat.Comp. Sec. Am.Stat.Assoc.: 326—331
[95.] Welniak E.J., Coder J.F. (1980): *A measure of the bias in the March CPS earning imputation system*, Proc.Survey.Res.Meth. Sec. Am.Stat.Assoc.: 421—425
[96.] Wiseman F., McDonald Ph. (1980): *Toward the development of industry standards for response and nonresponse rates*. Report no. 80—101, Marketing Science Institute, Cambridge, Mass.
[97.] Platek R., Singh M.P.,Tremblay V. (1978): *Adjustment for nonresponse in surveys*, w: Survey Sampling and Measurement, op.cit.: 157—174

Zagadnienie struktury badań reprezentacyjnych w Zintegrowanym Systemie Badań Gospodarstw Domowych (ZSBGD)

dr Józef Bielecki
Uniwersytet Gdański

Pierwsze badania gospodarstw domowych w zintegrowanej formie datują się od 1940 r., kiedy to w St. Zjedn. Ameryki zainicjowano ciągłe badanie reprezentacyjne bezrobocia (the Sample Survey of Unemployment), przekształcone w 1943 r. w ciągłe badanie ludnościowe (the Current Population Survey). To ostatnie prowadzone jest do dziś i ma postać wielotematycznego badania zintegrowanego.

Z metodologicznego punktu widzenia duże znaczenie praktyczne miała koncepcja „próby-matki" (ang. master sample). Po raz pierwszy „próbę-matkę" zastosowano w ciągłych badaniach rolnych, prowadzonych przez Departament Rolnictwa St. Zjedn. Ameryki.

W 1980 r. Biuro Statystyczne ONZ (Organizacja Narodów Zjednoczonych) wprowadziło w życie ambitny progrym statystyczny o nazwie *National Household Survey Capability Programme* (NHSCP)[1]), którego głównym celem jest pomoc organizacyjno-metodologiczna, jak i finansowa ONZ w budowie i rozwoju zintegrowanych systemów badań gospodarstw domowych w krajach rozwijających się. W ramach NHSCP opracowano założenia metodologiczno-organizacyjne badań w ramach zintegrowanych systemów.

Kraje rozwinięte, dysponujące infrastrukturą badań statystycznych, mogą rozwijać swoje systemy badań zgodnie z ogólnoświatowym programem (NHSCP). W październiku 1982 r. prezes GUS powołał zespół specjalistów do opracowania założeń organizacyjno-metodologicznych (ZSBGD) w Polsce. Całokształt problemów dotyczących tego systemu przedstawiono w publikacji [12, 26].

Rozwój ZSBGD (Zintegregowanego Systemu Badań Gospodarstw Domowych) w krajach o różnych systemach statystyki państwowej wiąże się nierozłącznie z doskonaleniem rozwiązań metodologicznych w zakresie systemowych badań reprezentacyjnych. Metodologia tych badań musi być adekwatna do specyfiki różnotematycznych badań przewidzianych ogólnokrajowym programem. W konsekwencji oznacza to stosowanie strategii odpowiednich kompromisów metodologicznych.

[1]) W języku polskim: Krajowy Program Doskonalenia Badań Gospodarstw Domowych.

Wybór struktury badań reprezentacyjnych w ZSBGD jest podporządkowany podstawowym wymogom jego sprawnego funkcjonowania. Wymogi te obejmują:

— ekonomiczną efektywność badań,
— elastyczność organizacyjną,
— użyteczność danych otrzymanych z badań.

Ekonomiczna efektywność badań oznacza, że w ramach ZSBGD powinny być stosowane takie rozwiązania administracyjne, organizacyjne i metodologiczne, których rezultatem jest maksymalna obniżka kosztów, przypadających na dane badanie przewidziane programem.

Elastyczność organizacyjna oznacza możliwość modyfikacji programów badań w okresie ich realizacji w zależności od pilności potrzeb użytkowników danych. Innym aspektem elastyczności organizacyjnej jest możliwość szybkiej adaptacji zespołów pracy terenowej do różnorodnych technik zbierania danych, wynikających ze specyfiki nowych badań.

Użyteczność danych otrzymanych z badań
to podstawowy wymóg stawiany przez użytkowników danych. Spełnienie tego wymogu jest możliwe dzięki stosowaniu takich strategii organizacyjno-metodologicznych, które gwarantują:

a) minimalizację błędów losowych i nielosowych,
b) wysoki stopień integracji danych,
c) aktualność danych (minimalizacja okresu czasowego między zakończeniem zbierania informacji i dostarczeniem danych przetworzonych dla użytkowników).

Z punktu widzenia metodologii badań reprezentacyjnych, właściwe spełnienie pierwszego i trzeciego wymogu prowadzi do konfliktu metodologicznego na skutek różnorodności tematycznej oraz specyfiki badań. Chcąc na przykład zapewnić wysoką jakość danych, poprzez minimalizację błędów losowych, należałoby dla każdego badania przedmiotowego stosować indywidualny schemat losowania, co z kolei wpłynęłoby na niezrealizowanie wymogu wysokiej efektywności ekonomicznej w całym systemie.

Biorąc pod uwagę ten fakt dochodzimy do wniosku, że podstawowe wymogi sprawnego funkcjonowania ZSBGD mogą być spełnione w dużym zakresie poprzez **strategię zróżnicowanych kompromisów metodologicznych.** Stopień zróżnicowania w kompromisach metodologicznych zależy od ogólnego celu badań w systemie (cel podstawowy) oraz od celów szczegółowych stawianych poszczególnym badaniom.

Podstawowym celem badań gospodarstw domowych w ramach ZSBGD jest gromadzenie kompleksowej i zintegrowanej informacji statystycznej

o stanach i procesach społeczno-ekonomicznych, zachodzących w populacji gospodarstw domowych. Tego rodzaju informacja jest użyteczna dla decydentów gospodarczych i politycznych oraz do celów badawczych. Szczegółowe cele badań dotyczą struktury pożądanej informacji przedmiotowej.

Warunek kompleksowości i zintegrowania danych statystycznych oraz względy kosztowe decydują w bezpośredni sposób o zakresie kompromisów metodogicznych.

Wybór systemu losowania prób w ZSBGD odnosi się z reguły do konkretnego krajowego programu badań realizowanego w określonym przedziale czasowym (np. w okresie 5 lat). Realizacja kolejnego programu badań może wymagać modyfikacji lub zmiany dotychczasowego schematu losowania. W praktyce badań zintegrowanych modyfikacja schematu losowania sprowadza się do całkowitej wymiany „próby matki", zmiany liczebności próby JLOS[2]) względnie zmian w estymacji parametrów populacji.

W procedurze decyzyjnej, związanej z wyborem systemu losowania prób w ZSBGD, można wyodrębnić dwie fazy:

Faza I. Wybór schematu losowania „próby matki" wspólnego dla wszystkich badań objętych krajowym programem,

Faza II. Wybór schematu losowania próby JLOS dla poszczególnych badań programowych.

Pierwsza faza decyzyjna opiera się na zasadzie **pełnego kompromisu** w wyborze schematu losowania „próby matki", tj. niezależnie od przedmiotu i specyfiki w technice zbierania danych w każdym badaniu wykorzystywane będą te same JLPS[3]) (próba jednostopniowa) względnie JLPS i JLDS[4]) (próba dwustopniowa), przy zastosowaniu jednego schematu losowania. Kompromis ten gwarantuje spełnienie wymogu efektywności ekonomicznej badań oraz elastyczności organizacyjnej bowiem z „próby matki" można zawsze wylosować podpróbę do realizacji pożądanego badania, bez dodatkowych nakładów [20].

W drugiej fazie decyzyjnej stosowana jest **strategia ograniczonych kompromisów,** według której niektóre badania mają ten sam schemat losowania, a niektóre zróżnicowany. Zagadnienie to omawiane jest w dalszej części referatu.

Biorąc pod uwagę wymóg ekonomicznej efektywności badań oraz elastyczności organizacyjnej powinno się dążyć do maksymalizacji rozwiązań

[2]) JLOS — jednostki losowania ostatniego stopnia.
[3]) JLPS — jednostki losowania pierwszego stopnia.
[4]) JLDS — jednostki losowania drugiego stopnia.

kompromisowych. Z drugiej strony rozwiązania takie nie gwarantują dobrej jakości danych (większe błędy losowe i nielosowe), zmniejszając tym samym użyteczność danych.

WYBÓR SCHEMATU LOSOWANIA „PRÓBY MATKI"

Jak już wykazano pierwsza faza decyzyjna związana z wyborem metody próbkowania w ZSBGD dotyczy wyboru schematu losowania „próby matki". Przez „próbę matkę" rozumieć będziemy dużą próbę jednostek, z której można losować podpróby, zależne lub niezależne, do poszczególnych badań jednorazowych lub do kolejnych cykli badań powtarzalnych [20, 4]. Tak zdefiniowana „próba matka" jest losowo pobraną podzbiorowością konkretnej populacji generalnej wykorzystywaną do badań na ustalony okres, np. 5 lat.

W przypadku badań gospodarstw domowych „próba matka" składa się z jednostek przestrzennych, wylosowanych z określonej populacji (ang. Sampled population) [18]. Populacja ta jest zbiorowością jednostek, dla których jest praktycznie możliwe zbudowanie operatu losowania. Populacja taka zawiera wszystkie ograniczenia podmiotowego zakresu badania np. wyłączenie jednostek na obszarach niedostępnych.

W praktyce zintegrowanych badań gospodarstw domowych, próba matka" losowana jest dla danego zbioru badań przewidzianych programem, którego okres realizacji pokrywa się z reguły z okresem „żywotności" wylosowanej „próby matki" [7].

Decyzja o wykorzystaniu do badań programowych „próby matki" uwarunkowana jest następującymi względami:

a) znaczną redukcją czasu i kosztu opracowania operatu losowania JLOS,
b) możliwością opracowania operatu JLOS w sytuacji kompletnego braku informacji o tych jednostkach,
c) możliwością usprawnienia organizacji zbierania danych w terenie (tworzenie punktów nadzoru i kontroli w JLPS pozostających przez kilka lat w próbie),
d) możliwością integracji danych indywidualnych lub zagregowanych, pochodzących z różnych badań przedmiotowych.

Wadą „próby matki" jest jej statyczność w czasie. Ma to szczególne znaczenie wówczas, gdy w jednostkach przestrzennych znajdujących się poza próbą zachodzą istotne zmiany. Taki stan rzeczy wpływać będzie na obciążenia szacunków. Zagadnienie to jest szczegółowo omówione w [5].

Strukturę procedury decyzyjnej, dotyczącej wyboru schematu losowania „próby matki", przedstawiono w tabl. 1.

Decydując się na losowanie i wykorzystanie „próby matki" do różnotematycznych badań gospodarstw domowych należy ustalić, czy próba taka będzie składać się wyłącznie z jednostek losowania pierwszego stopnia (JLPS) czy będzie to próba dwustopniowa, którą tworzyć będą JLPS

TABL. 1. STRUKTURA DECYZYJNA W WYBORZE SCHEMATU LOSOWANIA „PRÓBY MATKI" W ZSBGD

Etapy decyzyjne	Warianty decyzyjne w wyborze
1. Określenie składu „próby matki"	1.1. Próba jednostopniowa (złożona z JLPS) 1.2. Próba wielostopniowa (złożona zwykle z JLPS i JLDS)
2. Definicja jednostek losowania do „próby matki"	2.1. Terytorialne jednostki administracyjne (jednostki tego samego szczebla) 2.2. Obwody spisowe lub połączenia obwodów spisowych w większe jednostki terytorialne 2.3. Sektory przestrzenne (mniejsze jednostki specjalnie utworzone do losowania)
3. Warstwowanie populacji jednostek losowania do „próby matki"	3.1. Warstwy — dziedziny studiów 3.2. Warstwy według przestrzennego rozkładu cech (cechy-klasyfikatory) 3.3. Pseudowarstwy
4. Ustalenie liczebności „próby matki"	4.1. Liczebność ustalana arbitralnie
5. Lokalizacja jednostek „próby matki" w warstwach	5.1. Lokalizacja proporcjonalna 5.2. Lokalizacja arbitralna
6. Określenie prawdopodobieństw losowania jednostek do „próby matki"	6.1. Losowanie z prawdopodobieństwem proporcjonalnym do wielkości (PPW) 6.2. Losowanie z prawdopodobieństwem stałym (PS)
7. Wybór techniki losowania jednostek do „próby matki"	7.1. Losowanie bezzwrotne 7.2. Losowanie zwrotne
8. Określenie sposobu wykorzystania „próby matki" do badań programowych	8.1. Wszystkie jednostki „próby matki" uczestniczą w kolejnych badaniach lub cyklach badania okresowego 8.2. Z „próby matki" losowane są podpróby do kolejnych badań jednorazowych lub cykli badań okresowych 8.2.a. Niezależna podpróba dla każdego badania lub cyklu badania okresowego 8.2.b. Rotacja podprób

i JLDS. Dysponowanie „próbą matką" dwustopniową wpływa na znaczną redukcję kosztu badań oraz na znaczne ułatwienie organizacji obserwacji statystycznej w kolejnych badaniach.

Próba taka umożliwia również szeroki zakres integracji danych (integracja danych zagregowanych na poziomie JLPS i JLDS, jak i danych indywidualnych jeżeli np. JLTS[5]) podlegają częściowej rotacji). Należy jednak podkreślić, że „próba matka" dwustopniowa, ze względu na

[5]) JLTS — jednostki losowania trzeciego stopnia.

statyczność jednostek, zwłaszcza JLDS, wpływa na zwiększenie błędów losowych i nielosowych [20]. W praktyce zintegrowanych badań gospodarstw domowych najczęściej stosowana jest „próba matka" jednostopniowa. Przykłady stosowania „próby matki" dwustopniowej można znaleźć w publikacji [20].

Ważnym problemem decyzyjnym jest **określenie jednostki losowania** do „próby matki". W tabl. 1 wymieniono trzy rodzaje jednostek przestrzennych, które stosowane są w badaniach gospodarstw domowych. Określenie odpowiedniej jednostki przestrzennej do „próby matki" jest przedsięwzięciem o największym stopniu kompromisu.

Z teorii badań reprezentacyjnych wiadomo, że dobór jednostek zespołowych w badaniach wielostopniowych uwarunkowany jest w dużym stopniu specyfiką przedmiotową badania, a ściślej stopniem niejednorodności zespołów, ze względu na interesujące nas cechy statystyczne [5, 11]. Im większa jednostka zespołowa (przestrzenna) tym większa niejednorodność tej jednostki, ze względu na daną cechę, przy czym stopień jednorodności dla poszczególnych cech jest zróżnicowany. Można zatem wnioskować, że duże jednostki przestrzenne stosowane w badaniach wielostopniowych gwarantują większą redukcję błędu losowego, jednakże czynnikiem ograniczającym w tym przypadku jest koszt badania.

W praktyce zintegrowanych badań gospodarstw domowych stosuje się rozwiązanie kompromisowe, przyjmując za jednostkę losowania do „próby matki" najbardziej uniwersalną jednostkę terytorialną o ustalonych granicach i charakteryzującą się możliwie dużą stabilnością czasową swej wielkości.

W przypadku dwustopniowej „próby matki" jednostkami pierwszego stopnia losowania (JLPS) są z reguły większe jednostki administracyjne, takie jak: stany, obwody lub okręgi administracyjne, miasta i gminy wiejskie. Wybór jednostki terytorialnej wyższego lub niższego szczebla administracyjnego, tzn. większej lub mniejszej warunkowany jest dwoma podstawowymi czynnikami:

— nieaktualnością lub brakiem operatu losowania dla jednostek mniejszych,
— ograniczeniami finansowo-organizacyjnymi w realizacji programu badań.

Wylosowanie większych jednostek terytorialnych, co do których istnieje informacja statystyczna (rejestry danych o wielkości jednostek), umożliwia opracowanie operatu losowania dla jednostek drugiego stopnia losowania.

Ograniczenia finansowo-organizacyjne mają szczególne znaczenie w krajach o dużej powierzchni. Poprzez losowanie większych jednostek terytorialnych możliwa jest lepsza organizacja prac związanych z opracowaniem operatu losowania (np. instalacja punktów organizacji pracy terenowej w wylosowanych jednostkach) oraz oszczędności finansowe poprzez redukcję kosztów transportu (mniejsze przestrzenne rozproszenie jednostek).

Jednostkami drugiego stopnia losowania, w dwustopniowej „próbie matce", są zwykle obwody spisowe, dzielnice lub sektory miast, ad-

ministracyjnie wyodrębnione wsie lub specjalnie utworzone do celów badań, sektory przestrzenne.

W praktyce najczęściej stosuje się próbę jednostopniową, w której jednostkami losowania są mniejsze jednostki terytorialno-administracyjne lub połączenia obwodów spisowych (w krajach o małej powierzchni — obwody spisowe). W Polsce, w ramach ZSBGD, jednostkami losowania do „próby matki" jednostopniowej są tzw. terenowe punkty badań, będące rejonami statystycznymi lub zespołami rejonów, liczących minimum 150 mieszkań [13, 10]. W zintegrowanym systemie badań gospodarstw domowych Indii jednostopniową „próbę matkę" tworzą wsie w strefie wiejskiej i sektory miejskie w strefie miejskiej. W Maroku „próba matka" jest dwustopniowa; w strefie miejskiej próbę JLPS stanowią zespoły okręgów spisowych, zaś próba JLDS składa się z okręgów spisowych, z kolei w strefie wiejskiej w próbie JLPS znajdują się zespoły wsi, a w próbie JLDS wiejskie okręgi spisowe.

Biorąc pod uwagę względy estymacji parametrów i ich precyzji pożądane jest, aby „próba matka" jednostek losowania drugiego stopnia (JLDS) miała możliwie minimalną dyspersję wielkości tych jednostek, zwłaszcza gdy JLPS losowane są z prawdopodobieństwem proporcjonalnym do wielkości (PPW). Warunkom tym odpowiadają z reguły okręgi spisowe.

Warstwowanie populacji jednostek losowania do „próby matki" ma trzy podstawowe cele:
— wyodrębnienie warstw, obszarów zwanych dziedzinami studiów (ang. domains of study), dla których szacowane są parametry z akceptowalną precyzją[6]),
— wyodrębnienie w miarę jednorodnych podzbiorów, w celu zabezpieczenia reprezentatywności populacji generalnej,
— polepszenie precyzji szacunku parametrów w poszczególnych badaniach programowych.

W zintegrowanych badaniach gospodarstw domowych, jako dziedziny studiów, wyodrębnia się następujące podpopulacje terytorialne:
— strefa miejska i strefa wiejska kraju,
— regiony administracyjne kraju z ewentualnym podziałem na strefę miejską i wiejską,
— ważne podregiony administracyjne oraz duże aglomeracje miejskie,
— inne ważne obszary kraju o wyraźnie ustalonej delimitacji przestrzennej.

Publikacja danych dla poszczególnych dziedzin studiów wynika z potrzeb użytkowników, zwłaszcza decydentów. Bardzo często decydenci skłonni są zgłaszać zapotrzebowanie na dane dla dziedzin mniejszych, jak okręgi administracyjne, gminy czy miasta. Zasadniczym ograniczeniem jest wymóg precyzji oszacowań parametrów dla poszczególnych podpopulacji. Stosowa

[6]) W literaturze algielskojęzycznej dziedziny studiów określa się również jako dziedziny publikacji danych (domains of data publication).

ne powszechnie metody estymacji bazują na dostatecznie licznej próbie JLDS, co w przypadku małych dziedzin studiów nie jest spełnione [2, 1]. W ostatnich latach w metodologii badań reprezentacyjnych opracowano nowe metody szacowania parametrów dla małych dziedzin studiów, a wdrożenie ich w praktyce będzie niezwykle użyteczne.

W Polsce w planach badań ZSBGD przewiduje się 49 dziedzin studiów odpowiadających podziałowi kraju na województwa. W Indii wyodrębniono 31 dziedzin studiów odpowiadających poszczególnym stanom. W Maroku dziedzinami studiów jest 7 regionów kraju z podziałem na strefy: miejską i wiejską.

W każdej z wyodrębnionych dziedzin można dokonać warstwowania jednostek losowania do „próby matki" w celu zabezpieczenia lepszej reprezentatywności próby oraz zwiększenia precyzji szacunku parametrów. Jak podano w tablicy 1, warstwowanie takie można przeprowadzić dwoma metodami: warstwowanie w oparciu o cechy — klasyfikatory (warianty cech jakościowych i wartości liczbowe cech ilościowych) oraz tworzenie tzw. pseudowarstw.

Jako przykłady cech — klasyfikatorów można podać wskaźnik: urbanizacji, produkcji przemysłowej, produkcji rolnej, proporcji etnicznej, dynamiki ludnościowej, jak i inne wskaźniki, charakterystyczne dla przestrzennej struktury kraju. Najprostszym warstwowaniem jest podział dziedziny studiów na strefę miejską i wiejską, jeżeli strefy te nie są dziedzinami studiów. Dla przykładu w St. Zjedn. Ameryki w systemie badań gospodarstw domowych, prowadzonych przez Biuro Spisów, JLPS, którymi są okręgi administracyjne zwane „counties"[7]), warstwowane są w poszczególnych dziedzinach studiów według 5 cech klasyfikatorów [25].

Pseudowarstwy tworzy się poprzez losowanie systematyczne. Przykładem takiego losowania jest wybór systematyczny, po uprzednim uporządkowaniu, jednostek według określonego kryterium (np. losowanie JLPS w strefie miejskiej po uprzednim uporządkowaniu miast według liczby ludności). Zagadnienia te przedstawione są szczegółowo u Kisha [11]. Należy podkreślić, że warstwowanie według cech — klasyfikatorów jest możliwe jeśli posiadamy pełną i aktualną informację o klasyfikatorach.

Kolejnym etapem decyzyjnym jest **ustalenie liczebności „próby matki".** W praktyce zintegrowanych badań gospodarstw domowych liczebność „próby matki" ustalana jest arbitralnie, w oparciu o następujące uwarunkowania:

— wielkość populacji jednostek losowania do „próby matki" i jej rozkład przestrzenny,
— liczba dziedzin studiów,
— sposób wykorzystania „próby matki" do badań programowych,

---

[7]) „Counties" rzadko występują pojedyńczo, łączone są w grupy.

— liczba badań przewidzianych programem,
— okres wykorzystania „próby matki" (okres „żywotności" próby).

Wielkość populacji jednostek losowania do „próby matki" i jej rozkład przestrzenny zależy od obszaru kraju i gęstości zaludnienia. W krajach o dużym obszarze i dużej gęstości zaludnienia liczebność „próby matki" będzie większa w porównaniu do krajów mniejszych.

Z kolei im większa liczba wyodrębnionych dziedzin studiów, tym większa liczebność „próby matki". Wynika to z faktu, że „próba matka" musi gwarantować odpowiednią reprezentatywność każdej dziedziny studiów [4].

Jeśli chodzi o sposób wykorzystania „próby matki", to w przypadku wariantu 8.2 (patrz tabl. 1) liczebność próby powinna być znacznie większa, aby zabezpieczyć większą zmienność jednostek w podpróbach. Z kolei im większa liczba badań przewidzianych programem, dla których należy losować podpróby i im dłuższy okres wykorzystywania „próby matki" tym większa powinna być jej liczebność.

Dla przykładu w Indii jednostopniowa „próba matka" liczy 430 JLPS, w Maroku dwustopniowa „próba matka" liczy 76 JLPS i 145 JLDS [20]. W Polsce w ramach ZSBGD przewiduje się, że „próba matka" będzie liczyć 1800 JLPS [16].

Lokalizacja jednostek „próby matki" w dziedzinach studiów, jak i w warstwach utworzonych według cech — klasyfikatorów, może być proporcjonalna lub arbitralna. W praktyce rozdysponowanie jednostek próby następuje zgodnie z proporcją liczby ludności, w poszczególnych dziedzinach studiów, jak i w warstwach. Względy specjalnej ważności niektórych dziedzin lub warstw mogą być powodem lokalizacji arbitralnej.

Wybór prawdopodobieństwa losowania jednostek do „próby matki" zależy od przewidywanych schematów losowania próby JLDS dla poszczególnych badań programowych. W pryktyce badań gospodarstw domowych z reguły zakłada się samoważenie próby JLDS. Zastosowanie w takich przypadkach stałego prawdopodobieństwa (PS) do losowania JLPS, które prawie zawsze są zróżnicowanej wielkości, byłoby mało efektywne i stwarzałoby konieczność użycia systemu wag w procedurze estymacji. Powszechne zastosowanie ma zatem wybór JLPS z prawdopodobieństwem proporcjonalnym do wielkości (PPW). W przypadku „próby matki" dwustopniowej JLDS mogą być losowane z PS, jeśli ich wielkość jest minimalnie zróżnicowana (współczynnik zmienności wielkości JLDS nie większy od 5%).

W wyborze techniki losowania jednostek do „próby matki" bierze się pod uwagę dwa względy: efektywność losowania oraz łatwość losowania. Spełnienie tych dwóch wymogów osiąga się m.in. poprzez zastosowanie losowania systematycznego z losowym początkiem wyboru (losowanie bezzwrotne).

W strukturze decyzyjnej bardzo istotną sprawą jest **określenie sposobu wykorzystania „próby matki" do badań programowych.** Zasadniczo możliwe tu są dwa rozwiązania: „próba matka" jest panelem i z każdej JLPS losuje

się, w sposób zależny lub niezależny, próbę JLDS dla każdego badania (lub cyklu badania okresowego) względnie z „próby matki" losowane są podpróby do kolejnych badań lub cykli badań.

Zaletami pierwszego rozwiązania jest możliwość zwiększenia efektywności ekonomicznej badań oraz całkowitej integracji danych badań, przynajmniej na poziomie JLPS (jeżeli próba jest dwustopniowa). Zwiększona efektywność ekonomiczna badań jest wynikiem redukcji kosztów przygotowania operatu losowania oraz obserwacji statystycznej, ze względu na korzystanie z tych samych jednostek przestrzennych w próbie. Wadą takiego rozwiązania jest zbyt duża stagnacja jednostek w próbie, dająca w efekcie małą zmienność jednostek przestrzennych do różnych badań. Może to powodować nieco ograniczoną reprezentatywność „próby matki" dla wszystkich badań programowych.

Drugie rozwiązanie może występować w dwóch wariantach (w tabl. 1 poz. 8,2a i poz. 8,2b). W pierwszym wariancie — do każdego badania lub cyklu badania okresowego losuje się z „próby matki" niezależną podpróbę. Rozwiązanie takie umożliwia prowadzenie różnych badań przedmiotowych w tym samym czasie, co w efekcie daje większą elastyczność organizacyjną badań. Wadą takiego postępowania jest wyższy koszt badań oraz niemożliwość integracji danych, gdy „próba matka" jest jednostopniowa. Integracja danych zagregowanych możliwa jest tylko w przypadku „próby matki" dwustopniowej na poziomie JLPS. W drugim wariancie dokonuje się rotacji podprób do poszczególnych badań lub cykli badań okresowych. Jest to zwykle rotacja częściowa. Rozwiązanie takie daje obniżkę kosztu badania w porównaniu do wariantu pierwszego, ponadto umożliwia integrację danych z różnych badań zarówno zagregowanych, jak i indywidualnych, zwłaszcza gdy próby jednostek wyższych stopni są panelami.

Możliwe są również rozwiązania będące kombinacją wariantu 8,2a i wariantu 8,2b, np. niezależne podpróby dla pewnych badań jednorazowych i rotacja częściowa lub zerowa dla cykli badania okresowego.

WYBÓR SCHEMATÓW LOSOWANIA PRÓBY DLA BADAŃ PRZEDMIOTOWYCH WYKORZYSTUJĄCYCH „PRÓBĘ MATKĘ"

Wybór schematów losowania próby dla poszczególnych badań przedmiotowych opiera się na **strategii ograniczonych kompromisów.** Jeżeli głównym celem ZSBGD jest integracja danych indywidualnych z różnych badań przedmiotowych, wówczas należałoby stosować całkowity kompromis, polegający na stosowaniu tego samego schematu losowania próby JLDS do wszystkich badań. W efekcie ta sama próba JLDS (próba gospodarstw domowych) wykorzystywana byłaby do obserwacji cech każdego badania oraz każdego cyklu badania jednorazowego. Rozwiązanie takie stosowano w początkowym okresie funkcjonowania ZSBGD w Indii,

jednakże zostało ono zaniechane ze względu na przeciążenie gospodarstw domowych (efekt zmęczenia), którego skutkiem był wzrost błędów nielosowych.

W tej fazie decyzyjnej istotne znaczenie ma kompromis między wymogiem integracji danych (stopień integracji i zakres integracji) oraz wymogiem elastyczności metodologicznej, w zależności od specyfiki badań. W tabl. 2 przedstawiono podstawowe etapy decyzyjne w wyborze schematu losowania próby JLDS oraz możliwe warianty decyzyjne. Ze względu na fakt, że większość problemów związanych z losowaniem próby JLDS jest szeroko omawiana w literaturze ograniczono się jedynie do problemów wagi zasadniczej.

TABL. 2. STRUKTURA DECYZYJNA W WYBORZE SCHEMATU LOSOWANIA JEDNOSTEK DLA POSZCZEGÓLNYCH BADAŃ PRZEDMIOTOWYCH W ZSBGD

Etapy decyzyjne	Możliwe warianty decyzyjne
1. Określenie liczby faz losowania	1.1. Badanie jednofazowe 1.2. Badanie dwufazowe w przypadku badań specjalnych
2. Ustalenie liczby kolejnych stopni losowania próby	2.1. Jednakowa liczba stopni losowania dla wszystkich badań 2.2. Zróżnicowana liczba stopni losowania dla poszczególnych badań
3. Ustalenie liczebności podpróby jednostek losowania z „próby matki”	3.1. Arbitralne ustalenie liczebności 3.2. Ustalenie próby optymalnej 3.3. Jednakowa liczebność podpróby dla wszystkich badań programowych 3.4. Zróżnicowana liczebność podpróby dla poszczególnych badań
4. Definicja jednostek kolejnych stopni losowania	4.1. Jednakowo zdefiniowane jednostki losowania kolejnych stopni dla wszystkich badań 4.2. Jednostki losowania kolejnych stopni inne dla niektórych badań
5. Decyzja o warstwowaniu jednostek kolejnych stopni losowania	5.1. Warstwowanie według cech — klasyfikatorów 5.2. Pseudowarstwowanie 5.3. Jednakowe warstwowanie dla wszystkich badań 5.4. Zróżnicowane warstwowanie dla niektórych badań
6. Ustalenie liczebności próby jednostek kolejnych stopni losowania	6.1. Liczebność ustalona arbitralnie 6.2. Liczebność ustalona według formuł teoretycznych

TABL. 2. STRUKTURA DECYZYJNA W WYBORZE SCHEMATU LOSOWANIA JEDNOSTEK DLA POSZCZEGÓLNYCH BADAŃ PRZEDMIOTOWYCH W ZSBGD (dok.)

Etapy decyzyjne	Możliwe warianty decyzyjne
6. Ustalenie liczebności próby jednostek kolejnych stopni losowania (dok.)	6.3. Jednakowa liczebność próby JLOS dla wszystkich badań
	6.4. Zróżnicowana liczebność próby JLOS dla poszczególnych badań
7. Lokalizacja jednostek w warstwach	7.1. Lokalizacja arbitralna
	7.2. Lokalizacja proporcjonalna
	7.3. Lokalizacja optymalna
	7.4. Jednakowa lokalizacja dla wszystkich badań
	7.5. Zróżnicowana lokalizacja dla niektórych badań
8. Określenie prawdopodobieństw losowania i technik losowania do próby jednostek kolejnych stopni	8.1. Losowanie z PPW
	8.2. Losowanie z PS
	8.3. Losowanie zwrotne
	8.4. Losowanie bezzwrotne
	8.5. Jednakowa procedura losowania dla wszystkich badań
	8.6. Zróżnicowana procedura dla niektórych badań
9. Określenie sposobu wykorzystania próby JLOS do poszczególnych badań jednorazowych i cykli badań okresowych	9.1. Jedna próba JLOS panelem dla wszystkich badań i cykli badania okresowego
	9.2. Częściowa rotacja próby JLOS do badań i cykli badań
	9.3. Całkowita rotacja próby JLOS
	9.4. Stosowanie podprób przenikających
	9.5. Próby JLOS niezależne
10. Wybór metody estymacji parametrów populacji i oceny precyzji estymacji	10.1. Jednakowa metoda dla wszystkich badań
	10.2. Zróżnicowana metoda dla niektórych badań

W niektórych sytuacjach badania gospodarstw domowych w ZSBGD mogą być wielofazowe; dwu lub trzyfazowe, Ma to miejsce wówczas, kiedy zachodzi konieczność prowadzenia badania specjalistycznego (specjalna obserwacja) na ograniczonej liczebnie próbie JLOS, wylosowanej z większej próby, wykorzystywanej uprzednio lub jednocześnie do innego badania. Przykład trzyfazowego badania gospodarstw domowych podano w [3].

Badania wielofazowe gwarantują efektywność ekonomiczną i pożądaną integrację danych. Ich wadą jest przeciążenie tych samych jednostek obserwacją statystyczną, co wpływa na zwiększenie błędów nielosowych [20].

Ustalenie liczby kolejnych stopni losowania warunkowane jest dwoma głównymi czynnikami:
— specyfiką danego badania,
— ograniczeniami finansowo-organizacyjnymi w realizacji badań.

Przykładowo, ogólne badania demograficzne są zwykle badaniami wykorzystującymi próbę jednostek zespołowych (badanie jednostopniowe). W takim przypadku obserwacja demograficzna może być prowadzona na wszystkich elementach jednostek „próby matki" lub na elementach jednostek jej podpróby. W badaniach siły roboczej, gdzie próba JLDS jest znacznie mniejsza niż w badaniu demograficznym, wskazane jest warstwowanie jednostek obserwacji, konieczne staje się stosowanie losowania dwustopniowego, a nawet i trzystopniowego w celu opracowania powarstwowanego operatu losowania [17, 6] w taki sam sposób można postępować przy planowaniu badania dochodów i wydatków gospodarstw domowych.

Biorąc pod uwagę aspekt finansowo-organizacyjny będziemy unikać różnicowania stopni losowania dla poszczególnych badań. Z punktu widzenia precyzji oszacowań parametrów wskazane jest ograniczanie, w miarę możliwości, liczby stopni losowania.

W zintegrowanych badaniach gospodarstw domowych, wykorzystujących jednostopniową „próbę matkę", elementy próby są zawsze jednostkami losowania pierwszego stopnia (JLPS). Powstaje zatem problem — ile takich jednostek należy wylosować do podpróby przeznaczonej do danego badania. Z teoretycznego punktu widzenia istnieje możliwość ustalenia optymalnej liczebności takiej podpróby. W praktyce występują ograniczenia na skutek braku informacji o kosztach oraz o wariancji podstawowych cech. W takich warunkach liczebność podpróby JLPS ustalana jest arbitralnie, w oparciu o kryterium kosztowe i wymogi precyzji oszacowań parametrów.

Innym ważnym etapem decyzyjnym jest ustalenie liczebności próby jednostek ostatniego stopnia losowania (JLDS). W przypadku losowania dwustopniowego są to JLDS, które w praktyce są gospodarstwami domowymi lub mieszkaniami. W warunkach braku informacji o kosztach badania i zmienności cech próbę taką ustala się arbitralnie, tak aby spełnione były powyższe wymogi. Liczebność próby JLDS może być zróżnicowana w zależności od specyfiki danego badania programowego, np. inna liczebność próby dla badania demograficznego, inna dla badania dochodów i wydatków, inna dla badania siły roboczej, inna dla badania wyżywienia.

Kryterium kosztowe i wymogi precyzji decydują o wzajemnej relacji między liczebnością JLPS i JLDS (jeżeli JLDS są gospodarstwami

domowymi), bowiem inna jest konsekwencja w aspekcie nakładów i precyzji oszacowań w dysponowaniu próbą 10000 gospodarstw losowanych z 500 JLPS, a inna gdy ta sama próba losowana jest z 200 JLPS. Te właśnie wymogi mogą decydować o zróżnicowaniu wielkości próby JLDS przy takiej samej liczebności podpróby JLPS, wylosowanej z „próby matki".

Kompromisowe rozwiązanie, polegające na losowaniu próby o tej samej liczebności JLDS, może być podyktowane ograniczeniami w zakresie środków finansowych przeznaczonych na badania oraz trudnościami w organizacji obserwacji statystycznej.

Kolejnym ważnym zagadnieniem decyzyjnym jest sposób wykorzystania próby JLOS zarówno w badaniach jednorazowych, jak i w cyklach badań okresowych. Jeżeli do każdego badania lub cyklu badania losuje się niezależną podpróbę JLPS z „próby matki" względnie próbę zależną z całkowitą rotacją, to próba taka służy wyłącznie jednemu badaniu a indywidualne dane z badań nie mogą być ze sobą zintegrowane.

Częściowy kompromis osiąga się stosując próbę JLOS rotowaną dla cykli badań okresowych lub dla więcej niż jednego badania jednorazowego. Zalety metody rotacyjnej omówione są szczegółowo w literaturze [14]. Metoda ta ułatwia integrację danych indywidualnych.

W celu ułatwienia estymacji parametrów można stosować podpróby przenikające (minimum dwie podpróby JLDS) [11].

W drugiej fazie decyzyjnej, dotyczącej badań reprezentacyjnych należy ustalić metodę estymacji parametrów i ich wariancji. W praktyce badań zintegrowanych stosuje się zwykle pełny kompromis metodologiczny, przyjmując taką samą technikę estymacji dla wszystkich badań programowych. Powszechnie stosowany jest estymator ilorazowy (ang. ratio estimator) o postaci:

$$R = \frac{Y}{X}$$

Estymator ten jest obciążony, ale obciążenie to jest nieistotne, kiedy liczebność próby jest wystarczająco duża. Przy pomocy estymatora R można szacować zarówno wartość średnią, jak i proporcje oraz wskaźnik natężenia. Dodatkowym atutem jest pełne oprogramowanie całej procedury estymacyjnej.

UWAGI KOŃCOWE

- Wybór struktury badań reprezentacyjnych w Zintegrowanym Systemie Badań Gospodarstw Domowych oparty jest na strategii zróżnicowanych kompromisów.

- Konieczność kompromisów metodologicznych wynika z natury sytemu, który obejmuje programy badań wielotematycznych, okresowych i jednorazowych.
- Zakres kompromisów metodologicznych uwarunkowany jest z jednej strony ograniczeniami finansowo-organizacyjnymi w realizacji badań, z drugiej zaś wymogami jakości danych (minimalizacja błędów losowych i nielosowych).
- W systemie badań podstawowym rozwiązaniem organizacyjno-metodologicznym jest wykorzystanie do badań „próby matki" przestrzennych jednostek losowania.
- Stosowanie „próby matki" do realizacji różnotematycznych badań gospodarstw domowych narzuca konieczność elastycznej strategii ograniczonych kompromisów w wyborze schematów losowania prób dla poszczególnych badań programowych.
- Szczegółowa dokumentacja realizacji programów badań w ramach ZSBGD daje możliwość ciągłego doskonalenia rozwiązań z zakresu metodologii systemu badań reprezentacyjnych.

BIBLIOGRAFIA

[1] J. Bielecki (1981): *Etapas y organizaciones de investigaciones muestrales.* INEC. Quito. 1981.

[2] J. Bielecki (1986): *Prace nad integracją badań gospodarstw domowych w krajach Trzeciego Świata.* Wiadomości Statystyczne nr 6.

[3] J. Bielecki (1987): *Zagadnienie integracji badań gospodarstw domowych w krajach Trzeciego Świata,* w „Problemy integracji statystycznych badań gospodarstw domowych" GUS, Warszawa.

[4] Chris Scott (1987): *Problèmes de l'echantillon maître.* Séminaire sur les sondages. Maroc. Janvier.

[5] Chris Scott, Trudy Harham: *Major Issues of Survey and Sample Design.* World Fertility Survey 1972—1984. Symposium 24—27 April 1984. Paper 142/1.

[6] Foreman, E. K. (1983): *Integrated Programmes of Household Surveys: Design Aspects* „Bulletin of the International Statistical Institute 50/2/ s. 1344—1362.

[7] *Handbook of Household Surveys.* Studies in Methods S.F. No 31. United Nations, New York 1984.

[8] *Household Surveys in Asia:* Organisation and Methods. UN ESCAP. Bangkok, 1981.

[9] L. Kish (1980): *Design and estimation for domains.* „The Statisticians" 29.

[10] L. Kish (1987): *Statistical Design for Research.* John Wiley and Sons.

[11] L. Kish (1965): *Survey Sampling.* 1965. New York. John Wiley and Sons.

[12] J. Kordos (1983): *Możliwości zbudowania zintegrowanego systemu badań gospodarstw domowych w Polsce.* „Wiadomości Statystyczne" nr 9.

[13] J. Kordos, S. Mierzejewski (1987): *Perspektywy Zintegrowanego Systemu Badań Gospodarstw Domowych,* „Problemy integracji statystycznych badań gospodarstw domowych" GUS, Warszawa, 1987.

[14] J. Kordos (1972): *Wybrane problemy metody rotacyjnej w badaniach budżetów rodzinnych,* w „Eksperymentalne badania budżetów rodzinnych metodą rotacyjną". Biblioteka Wiadomości Statystycznych GUS, Tom 18.

[15] *Las encuestas de hogares en America Latina,* NU, CEPAL, Santiago de Chile 1983.

[16] B. Lednicki (1987): *Schemat losowania „próby matki" do Zintegrowanego Systemu Badań Gospodarstw Domowych w Polsce,* w „Problemy integracji statystycznych badań gospodarstw domowych" GUS, Warszawa.

[17] *Manuel des enquêtes démographiques par sondage en Afrique.* CEA. Addis-Abeba, Septembre 1974.

[18] M.N. Murthy (1974): *Evolution of multi-subject sample survey systems.* „International Statistical Review" 42.

[19] M.N. Murthy (1968): *On designing and conducting multi-subject household enquiries with reference to permanent survey organisation.* „Sankhya" 30/B/.

[20] *Sampling Frames and Sample Designs for Integrated Household Surveys Programmes.* NHSCP UN. New York. 1986.

[21] *Small Area Statistics. An International Symposium.* J. Wiley and Sons. 1987.

[22] J. Steczkowski (1988): *Zastosowanie metody reprezentacyjnej w badaniach społeczno-ekonomicznych.* PWN, Warszawa.

[23] M. Tin, G. Mandishona (1987): *Master sample used in national household surveys of Zimbabwe.* ISI, Tokyo September, 8—16, 1987.

[24] R. Torene, L.G. Torene, M.Z. Hoque (1987): *The practical side of using master samples.* ISI, Tokyo, September 8—16, 1987.

[25] *The Current Population Survey Design and Methodology.* Bureau of the Census. Technical Paper 40. Washington 1978.

[26] *Założenia Zintegrowanego Systemu Badań Gospodarstw Domowych w Polsce.* w „Problemy integracji statystycznych badań gospodarstw domowych. GUS, Warszawa 1987.

[27] R. Zasępa (1972): *Metoda reprezentacyjna.* PWE, Warszawa 1972.

[28] R. Zasępa (1981): *Efektywność badania rotacyjnego w porównaniu z badaniem ciągłym w przypadku dwustopniowego losowania próby.* Metoda reprezentacyjna w masowych badaniach statystycznych. „Teoria i praktyka". ZBSE, Zeszyt 122.

[29] R. Zasępa (1987): *O pewnych zagadnieniach planowania badania reprezentacyjnego,* „Zastosowanie metody reprezentacyjnej w badaniach statystycznych GUS (1981—1986) ZBSE, Zeszyt 166.

Stosowanie wybranych testów statystycznych do danych z prób nieprostych

dr hab. Czesław Bracha
Szkoła Główna Planowania i Statystyki

Teoria statystyki matematycznej dostarcza reguł wnioskowania statystycznego (estymacji nieznanych parametrów i weryfikacji hipotez) na podstawie prób prostych tzn. takich, w których obserwacje są stochastycznie niezależne i mają ten sam rozkład prawdopodobieństwa. W praktyce badań reprezentacyjnych, prowadzonych na dużą skalę (np. przez GUS, ośrodki badania opinii publicznej itp.), próby proste nie występują, gdyż względy ekonomiczne i organizacyjne zmuszają statystyków do stosowania takich schematów (i ich kombinacji), jak: (a) losowanie bez zwracania; (b) losowanie warstwowe; (c) losowanie wielostopniowe; (d) losowanie z różnymi prawdopodobieństwami. To powoduje, że w próbie znajdują się obserwacje zależne i o różnych rozkładach. Próby losowane według wymienionych schematów i ich kombinacji są próbami nieprostymi (złożonymi, ang. complex samples).

Klasyczna teoria badań reprezentacyjnych rozwiązała problem estymacji takich parametrów, jak średnia czy wartość globalna cechy. Rozwiązała również problem szacowania wariancji estymatorów wcześniej wymienionych parametrów. Gorzej rzecz miała się z szacowaniem takich parametrów, jak współczynniki regresji oraz z weryfikacją hipotez statystycznych. Jednakże i w tych dziedzinach można zauważyć, na przestrzeni ostatniego ćwierćwiecza, olbrzymi postęp. Wywołany jest on zarówno prowadzonymi przez wielu statystyków pracami czysto teoretycznymi, jak i pracami, w których symulacja komputerowa odgrywa podstawową rolę.

Za prekursora nieklasycznego podejścia do danych z prób nieprostych należy uznać, jak sądzę, H.S. Konijna [13], który podał sposoby szacowania współczynników regresji, na podstawie prób losowanych dwustopniowo. Następne prace, to: J. Sedransk [22,23], P. Armitage [1], H.L. Jones [8] oraz G. Nathan [15]. Duże znaczenie, dla dalszego rozwoju badań nad wnioskowaniem statystycznym na podstawie prób nieprostych, miały prace: P.J. Mc Carthy'ego [14], L. Kish i M.R. Frankel [11] oraz M.R. Frankel [4], w których została podana dość prosta metoda szacowania wariancji złożonych estymatorów. Metoda ta wykorzystuje wielokrotne powtórzenia zrównoważone (balanced repeated replication — BRR).

Lata siedemdziesiąte i osiemdziesiąte przyniosły całą lawinę prac poświęconych omawianym tu zagadnieniom. Do najważniejszych z nich zaliczyć można: D. Holt, A.J. Scott i P.D. Ewings [6], D. Holt, T.M.F.

Smith i P.D. Winter [7], G. Nathan i D. Holt [16], J.N.K. Rao i A.J. Scott [19], A.J. Scott i D. Holt [21] oraz J.N.K. Rao i A.J. Scott [20].

W wymienionych wyżej pracach dominowały dwie tematyki: analiza regresji oraz analiza danych jakościowych (różne odmiany testu χ^2 Pearsona). Ponieważ dopuszczalne rozmiary referatu uniemożliwiają zajęcie się (nawet pobieżne) tak obszerną tematyką, dalej zastaną omówione tylko niektóre problemy związane ze stosowaniem testów χ^2 do danych pochodzących z prób nieprostych.

TESTY χ^2 PEARSONA

Testy χ^2 Pearsona dla prób prostych są testami ilorazu wiarogodności. Zasada ich konstrukcji jest zawsze taka sama, choć różnią się wzorami statystyk testowych i liczbami stopni swobody. Do testów χ^2 Pearsona należą:

1. Test zgodności rozkładu empirycznego z rozkładem hipotetycznym.
2. Test zgodności dwóch lub więcej rozkładów empirycznych (test jednorodności).
3. Test niezależności dwóch lub więcej zmiennych losowych dyskretnych.

Dla próby prostej w każdym z trzech wymienionych przypadków statystyka testowa ma (gdy H_0 jest prawdziwa) asymptotyczny rozkład χ^2 o określonej liczbie stopni swobody, która zależy wyłącznie od liczby szacowanych na podstawie próby parametrów. Jeżeli próba nie jest prosta, to statystyki określone klasycznymi wzorami nie mają już asymptotycznego rozkładu χ^2 i ich stosowanie, co potwierdziły liczne badania empiryczne, prowadzi do znacznego zwiększenia rozmiaru testu (prawdopodobieństwa błędu I rodzaju), który w skrajnych przypadkach może sięgać nawet 80%. W takich sytuacjach test staje się bezużyteczny. Dlatego też liczni statystycy podjęli próby modyfikacji statystyk testowych i podania ich własności.

Prekursorem w tej dziedzinie był P. Armitage [1], który pokazał, że w przypadku próby nieprostej statystyka χ^2 w teście zgodności ma rozkład o innej liczbie stopni swobody niż w przypadku próby prostej. Dalsze prace poświęcone testom χ^2 dla prób nieprostych prowadzili: G. Nathan [15], J.E. Cohen [2] i J.P. Fellegi [3]. Przełomową, moim zdaniem, pracą poświęconą tym zagadnieniom był artykuł D. Holta, A.J. Scotta i P.D. Ewingsa [6]. Cytowana ostatnio trójka autorska zaproponowała prostą i zarazem skuteczną modyfikację testów χ^2 Pearsona. Dalszymi pracami z tej dziedziny są: J.N.K. Rao i A.J. Scott [19, 20].

Test zgodności

Niech cecha X przyjmuje wartości należące do $k\,(k \geqslant 2)$ rozłącznych przedziałów klasowych (kategorii) i niech p_i oznacza prawdopodobieństwo tego, że X przyjmie wartość z i-tego przedziału, przy czym $p_i > 0$ dla

$i=1,...,k$ oraz $\sum_{i=1}^{k} p_i=1$. Załóżmy dalej, że interesuje nas zweryfikowanie hipotezy $H_0 : \mathbf{p}=\mathbf{p}_0$, gdzie $\mathbf{p}=[p_i]_{(k-1)x1}$ oraz $\mathbf{p}_0=[p_{i0}]_{(k-1)x1}$ jest wektorem hipotetycznych prawdopodobieństw związanych z $\mathbf{p}$.

Powszechnie znaną i stosowaną statystyką testową jest statystyka:

$$\chi_P^2=n\sum_{i=1}^{k}\frac{(\hat{p}_i-p_{i0})^2}{p_{i0}} \tag{1}$$

gdzie $\hat{p}_i$ jest estymatorem p_i. Jeżeli próba jest prosta i n jest dostatecznie duże, to w przypadku prawdziwości H_0 statystyka χ_P^2 ma w przybliżeniu rozkład $\chi^2(k-1)$.

Statystykę (1) można przedstawić w następującej, równoważnej postaci:

$$\chi_P^2=n(\hat{\mathbf{p}}-\mathbf{p}_0)^T\,\mathbf{P}_0^{-1}\,(\hat{\mathbf{p}}-\mathbf{p}_0) \tag{2}$$

gdzie: $\mathbf{P}_0=diag\,(\mathbf{p}_0)-\mathbf{p}_0\mathbf{p}_0^T$ oraz $\hat{\mathbf{p}}=[\hat{p}_i]_{(k-1)x1}$.

W przypadku próby prostej mamy:

$$\sqrt{n}\,(\hat{\mathbf{p}}-\mathbf{p}_0)\xrightarrow{\;L\;}N_{k-1}\,(\mathbf{0},\mathbf{P}_0)^{1)} \tag{3}$$

Jeżeli próba nie jest prosta, a z taką najczęściej ma się do czynienia w praktyce badań reprezentacyjnych, to wtedy

$$\sqrt{n}\,(\hat{\mathbf{p}}-\mathbf{p}_0)\xrightarrow{\;L\;}N_{k-1}(\mathbf{0},\tau) \tag{4}$$

gdzie: τ jest macierzą kowariancji wektora $\sqrt{n}(\hat{\mathbf{p}}-\mathbf{p}_0)$.

W takiej sytuacji statystyka χ_P^2 nie posiada już asymptotycznego rozkładu $\chi^2(k-1)$ i zamiast niej należy stosować statystykę Walda, postaci:

$$\chi_W^2=n(\hat{\mathbf{p}}-\mathbf{p}_0)^T\,\hat{\mathbf{V}}^{-1}\,(\hat{\mathbf{p}}-\mathbf{p}_0) \tag{5}$$

gdzie: $\hat{\mathbf{V}}=[\hat{v}_{ij}]_{(k-1)x(k-1)}$ jest zgodnym estymatorem macierzy τ.

A. Wald udowodnił, że jeżeli H_0 jest prawdziwa, to

$$\chi_W^2\xrightarrow{\;L\;}\chi^2(k-1) \tag{6}$$

Ponieważ dla masowych badań statystycznych nie szacuje się na ogół pełnej macierzy $\mathbf{V}$, to nie można stosować statystyki χ_W^2 i wielu statystyków

---

[1]) Zapis $N_k(\mathbf{m},\sum)$, przy $k\geq 2$ oznacza k-wymiarowy rozkład normalny o podanych w nawiasach parametrach.

zamiast niej stosuje χ_P^2, co może doprowadzić do znacznego (w stosunku do założonego P) zwiększenia rozmiaru testu.

Poniżej zostanie przedstawiona prosta modyfikacja statystyki (2). Niech $\mathbf{X}=\sqrt{n}(\hat{\mathbf{p}}-\mathbf{p}_0)$ i wtedy (2) przyjmuje postać

$$\chi_P^2=\mathbf{X}^T\mathbf{P}_0^{-1}\mathbf{X} \tag{7}$$

Ponieważ v jest macierzą symetryczną i dodatnio określoną, to istnieje taka macierz trójkątna dolna L, że $\mathbf{V}=\mathbf{L}\mathbf{L}^T$. Stąd dla wektora losowego $\mathbf{Y}=\mathbf{L}^{-1}\mathbf{X}$ będzie spełniony warunek

$$\mathbf{Y}\xrightarrow{L} N_{k-1}\ (\mathbf{0},\mathbf{I}) \tag{8}$$

i (7) można zapisać w postaci:

$$\chi_P^2=\mathbf{Y}^T\mathbf{L}^T\mathbf{P}_0^{-1}\mathbf{L}\mathbf{Y} \tag{9}$$

Ponieważ $\mathbf{L}^T\mathbf{P}_0^{-1}\mathbf{L}$ jest macierzą symetryczną i dodatnio określoną, to istnieje taka macierz ortgonalna A, że

$$\Lambda=\mathbf{A}^T\mathbf{L}^T\mathbf{P}_0^{-1}\mathbf{L}\mathbf{A} \tag{10}$$

przy czym $\Lambda=\mathrm{diag}\,(\lambda_{1,0},\ldots,\lambda_{k-1,0})$ oraz $\lambda_{1,0}\geqslant\ldots\geqslant\lambda_{k-1,0}$ ($\lambda_{i,0}$ są wartościami własnymi macierzy $\mathbf{L}^T\mathbf{P}_0^{-1}\mathbf{L}$ lub na co jedno wychodzi macierzy $\mathbf{P}_0^{-1}\mathbf{V}$). Wykorzystując transformację $\mathbf{Z}=\mathbf{A}^T\mathbf{Y}\,(\mathbf{Z}\xrightarrow{L} N_{k-1}(\mathbf{0},\mathbf{I}))$ statystykę (9) można zapisać w postaci:

$$\chi_P^2=\mathbf{Z}^T\mathbf{Z}=\sum_{i=1}^{k-1}\lambda_{i,0}Z_i^2\xrightarrow{L}\sum_{i=1}^{k-1}\lambda_{i,0}\chi_i^2(1)=U \tag{11}$$

Z powyższego wzoru wynika, że przy prawidłowości H_0 statystyka χ_P^2 ma asymptotyczny rozkład $\chi^2(k-1)$ tylko wówczas, gdy wszystkie wartości własne $\lambda_{i,0}$ są równe jedności. Ma to miejsce tylko wtedy, gdy $\mathbf{V}=\mathbf{P}_0$, czyli wtedy, gdy próba jest prosta.

Jeżeli $k=2$, to macierze $\mathbf{P}_0$ i V redukują się do skalarów i wtedy:

$$\lambda_{1,0}=\frac{nD^2(\hat{p}_1)}{p_{1,0}(1-p_{1,0})}=deff(\hat{p}_1) \tag{12}$$

gdzie $deff(\cdot)$ jest wprowadzonym przez L. Kisha [9] wyrażeniem zwanym efektem schematu losowania (design effect), jest to iloraz wariancji estymatora, przy zastosowanym schemacie losowania do wariancji estymatora dla próby prostej przy założeniu, że w obu przypadkach liczebności próby są takie same.

Dla $k>2$ macierz $\mathbf{D}_0=\mathbf{P}_0^{-1}\mathbf{V}$ nazywana jest przez analogię macierzą efektu schematu losowania (design effect matrix).

Zależność (11) można by wykorzystać przy weryfikacji H_0, gdyby znane były $\lambda_{i,0}$ lub ich estymatory $\hat{\lambda}_{i,0}$, wtedy można by skorzystać z rozkładu ważonych sum zmiennych losowych $\chi_i^2(1)$ podanego przez H. Solomona i M.A. Stephensa [24]. Jednakże do wyznaczenia $\hat{\lambda}_{i,0}$ potrzebna jest znajomość macierzy $\mathbf{P}_0^{-1}\hat{\mathbf{V}}$. Łatwo zauważyć, że w przypadku znajomości macierzy $\hat{\mathbf{V}}$ wygodniej posłużyć się statystyką χ_W^2, gdyż takie rozwiązanie wymaga tylko odwrócenia macierzy stopnia $k-1$ bez konieczności szukania wartości własnych. W przypadku, gdy nie jest znana cała macierz $\hat{\mathbf{V}}$, a tylko jej elementy diagonalne, tzn. $\hat{v}_{ii}=n\hat{D}^2(\hat{p}_i)$, można skorzystać z następującego przybliżenia. Gdyby wszystkie wartości własne $\lambda_{1,0}, ..., \lambda_{k-1,0}$ niewiele się różniły między sobą (tym samym od średniej $\bar{\lambda}=\frac{1}{k-1}\sum_{i=1}^{k-1}\lambda_{i,0}$), to wtedy można by napisać:

$$U=\sum_{i=1}^{k-1}\lambda_{i,0}\chi_i^2(1)\approx\bar{\lambda}\sum_{i=1}^{k-1}\chi_i^2(1)=\bar{\lambda}\chi^2(k-1) \tag{13}$$

Ponieważ:

$$\bar{\lambda}=\frac{1}{k-1}tr(\mathbf{P}_0^{-1}\mathbf{V})=\frac{1}{k-1}\sum_{i=1}^{k}\frac{v_{ii}}{p_{i,0}}, \tag{14}$$

to jako estymatora $\bar{\lambda}$ można użyć statystyki:

$$\hat{\lambda}=\frac{n}{k-1}\sum_{i=1}^{k}\frac{\hat{D}^2(\hat{p}_i)}{p_{i,0}} \tag{15}$$

i tym samym dokonać następującej modyfikacji statystyki χ_P^2:

$$\chi_m^2=\chi_p^2/\hat{\lambda}=\frac{(k-1)\sum_{i=1}^{k}\frac{(\hat{p}_i-p_{i,0})^2}{p_{i,0}}}{\sum_{i=1}^{k}\frac{\hat{D}^2(\hat{p}_i)}{p_{i0}}} \tag{16}$$

Statystyka χ_m^2 w przypadku dużych n i prawdziwości H_0 ma rozkład zbliżony do $\chi^2(k-1)$.

Badania symulacyjne, przeprowadzone przez cytowaną, już trójkę autorską oraz przez: J.N.K. Rao i A.J. Scotta [19], J.N.K. Rao i M.A. Hidiroglou [18] wykazały wysoką zgodność rozkładu χ_m^2 z rozkładem $\chi^2(k-1)$. O walorach modyfikacji (16) świadczy to, że w przypadku

zastosowania χ_P^2 rozmiar testu przy zadanym $\alpha=0{,}05$ sięgał nawet 0,80, podczas gdy dla χ_m^2 wahał się w granicach $\langle 0{,}05;\ 0{,}06\rangle$.

Należy tu dodać, że automatyczne wyważenie próby nie upoważnia do traktowania jej jako prostej (w badaniach symulacyjnych próby były automatycznie wyważone). Potwierdza to również teoretyczny przykład zamieszczony w pracy J.N.K. Rao i A.J. Scott [19].

Test jednorodności

Załóżmy obecnie, że danych jest r $(r\geqslant 2)$ populacji. Niech $\mathbf{P}_i=[p_{ij}]_{(k-1)\times 1}$, gdzie p_{ij} jest prawdopodobieństwem tego, że cecha X_i przyjmie wartość należącą do j-tej kategorii $(j=1, ..., k-1;\ i=1, ..., r)$. Chcemy zweryfikować $H_0: \mathbf{p}_1=...=\mathbf{p}_r=\mathbf{p}$, wobec alternatywy $H_1: \bigvee_i \mathbf{p}_i\neq\mathbf{p}$. W tym celu z i-tej populacji losowana jest według wybranego schematu n_i — elementowa próba (nie wymaga się, aby dla wszystkich populacji zastosowano identyczne schematy losowania). Niech $\hat{\mathbf{p}}_i=[\hat{p}_{ij}]$ będzie estymatorem wektora $\mathbf{p}_i$, natomiast $\hat{\mathbf{p}}=\sum_{i=1}^{r} f_i\hat{\mathbf{p}}_i$ — wektora $\mathbf{p}$, gdzie $n=\sum_{i=1}^{r} n_i$ oraz $f_i=n_i/n$.

Jeżeli wszystkie próby byłyby proste, to statystyka:

$$\chi_P^2=\sum_{i=1}^{r} n_i \sum_{j=1}^{k}\frac{(\hat{p}_{ij}-\hat{p}_j)^2}{\hat{p}_j}=n\hat{\mathbf{q}}^T(\mathbf{F}\otimes\hat{\mathbf{P}}^{-1})\hat{\mathbf{q}} \tag{17}$$

gdzie $\hat{\mathbf{q}}^T=[(\hat{\mathbf{p}}_1-\hat{\mathbf{p}}_r)^T ... (\hat{\mathbf{p}}_{r-1}-\hat{\mathbf{p}}_r)^T]$, $\mathbf{f}=[f_i]$
$\mathbf{F}=diag\,(\mathbf{f})-\mathbf{f}\mathbf{f}^T$ oraz $\hat{\mathbf{P}}=diag\,(\hat{\mathbf{p}})-\hat{\mathbf{p}}\hat{\mathbf{p}}^T$, miałaby w przypadku prawdziwości H_0 asymptotyczny rozkład $X^2[(k-1)(r-1)]$. W przypadku prób nieprostych należy użyć do weryfikacji H_0 statystyki Walda w postaci:

$$X_W^2=\hat{\mathbf{q}}^T\hat{\Lambda}^{-1}\hat{\mathbf{q}} \tag{18}$$

gdzie $\hat{\Lambda}$ jest estymatorem macuerzy kowariancji wektora $\mathbf{q}$. Jeżeli H_0 jest prawdziwa, to χ_W^2 określona wzorem (18) ma asymptotyczny rozkład $\chi^2[(k-1)(r-1)]$. W przypadku nieznajomości pełnej macierzy $\hat{\Lambda}$, a tylko jej elementów diagonalnych należy zastosować statystykę:

$$\chi_m^2=X_P^2/\hat{\lambda} \tag{19}$$

gdzie χ_P^2 określona jest wzorem (17), zaś:

$$\hat{\lambda}=\frac{1}{(r-1)(k-1)}\sum_{i=1}^{r}(1-f_i)n_i\sum_{j=1}^{k}\frac{\hat{D}^2(\hat{p}_{ij})}{\hat{p}_j} \tag{20}$$

Okazuje się (potwierdziły to wspomniane już badania symulacyjne), że i statystyka (19) ma w przybliżeniu rozkład $\chi^2[(k-1)(r-1)]$, gdy H_0 jest prawdziwa.

Test niezależności

Załóżmy, że interesują nas cechy X i Y, przy czym X przyjmuje wartości należące do r ($r \geqslant 2$) kategorii, natomiast Y — do c ($c \geqslant 2$) kategorii. Niech dalej p_{ij} oznacza prawdopodobieństwo tego, że X przyjmie wartość należącą do i-tej kategorii, natomiast Y — do j-tej ($i=1, ..., r$; $j=1, ..., c$) oraz $\boldsymbol{P}=[p_{11}p_{12} \dots p_{rc-1}]^T$. Celem jest zweryfikowanie hipotezy, że X i Y są stochastycznie niezależne lub, co na jedno wychodzi, $H_0: p_{ij}=p_{i.}p_{.j}$ dla $i=1, ..., r$; $j=1, ..., c$, gdzie $p_{i.}=\sum_{j=1}^{c} p_{ij}$ $(i=1, ..., r)$ oraz $p_{.j}=\sum_{i=1}^{r} p_{ij}$ $(j=1, ..., c)$.

W przypadku n-elementowej próby prostej zachodzi:

$$\sqrt{n}(\hat{\boldsymbol{p}}-\boldsymbol{p}) \xrightarrow{L} N_{rc-1}(\boldsymbol{0},\boldsymbol{P}) \tag{21}$$

gdzie $\hat{p}_{ij}$ oraz $\hat{\boldsymbol{p}}$ są estymatorami parametrów p_{ij} oraz $\boldsymbol{p}$, odpowiednio, natomiast $\boldsymbol{P}=diag(\boldsymbol{p})-\boldsymbol{p}\boldsymbol{p}^T$. Statystyka:

$$X_P^2=n\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(\hat{p}_{ij}-\hat{p}_{i.}\,\hat{p}_{.j})^2}{\hat{p}_{i.}\,\hat{p}_{.j}} \tag{22}$$

gdzie $\hat{p}_{i.}=\sum_{j=1}^{c}\hat{p}_{ij}$ oraz $\hat{p}_{.j}=\sum_{i=1}^{r}\hat{p}_{ij}$

ma w przybliżeniu rozkład $X^2[(r-1)(c-1)]$ (jeżeli H_0 jest prawdziwa i n jest dostatecznie duże). Statystykę (22) można zapisać w następującej, równoważnej postaci:

$$\chi_p^2=n[\boldsymbol{h}(\hat{\boldsymbol{p}})]^T(\hat{\boldsymbol{P}}_r^{-1}\otimes\hat{\boldsymbol{P}}_c^{-1})\boldsymbol{h}(\hat{\boldsymbol{p}}) \tag{23}$$

gdzie: $\hat{\boldsymbol{P}}=(\hat{p}_{11}, \hat{p}_{12}, ..., \hat{p}_{rc-1})^T$, $\boldsymbol{h}(\hat{\boldsymbol{p}})=[h_{ij}(\hat{\boldsymbol{p}})]_{(r-1)(c-1)\times 1}$, przy czym $h_{ij}(\hat{\boldsymbol{p}})=\hat{p}_{ij}-\hat{p}_{i.}\,\hat{p}_{.j}$, $\hat{\boldsymbol{P}}_r=(\hat{p}_{1.} \dots \hat{p}_{r-1,.})^T$, $\hat{\boldsymbol{P}}_c=(\hat{p}_{.1} \dots \hat{p}_{.,c-1})^T$, $\hat{\boldsymbol{P}}_r=diag(\hat{\boldsymbol{p}}_r)-\hat{\boldsymbol{p}}_r\hat{\boldsymbol{p}}_r^T$, $\hat{\boldsymbol{P}}_c=diag(\hat{\boldsymbol{p}}_c)-\hat{\boldsymbol{p}}_c\hat{\boldsymbol{p}}_c^T$.

Jeżeli próba nie jest prosta, to:

$$\sqrt{n}(\hat{\mathbf{p}}-\mathbf{p}) \xrightarrow{L} N_{rc-1}(\mathbf{0},\mathbf{V}) \tag{24}$$

przy czym $\mathbf{V} \neq \mathbf{P}$ i wtedy należy stosować statystykę w postaci:

$$\chi_W^2 = n[\boldsymbol{h}(\hat{\mathbf{p}})]^T \{\hat{\mathbf{V}}[\mathbf{h}(\hat{\mathbf{p}})]\}^{-1} \mathbf{h}(\hat{\mathbf{p}}) \tag{25}$$

gdzie: $\hat{\mathbf{V}}[\mathbf{h}(\hat{\mathbf{p}})]/n$ jest estymatorem macierzy kowariancji $\mathbf{V}[\mathbf{h}(\hat{\mathbf{p}})]/n$. W przypadku prawdziwości H_0 rozkład χ_W^2 dąży do rozkładu $\chi^2[(r-1)(c-1)]$, przy $n \to \infty$. Ponieważ $\hat{\mathbf{V}}[\mathbf{h}(\hat{\mathbf{p}})]$ jest macierzą stopnia $(r-1)(c-1)$, to należy wątpić w to, aby przy większych r i c, macierz ta była dostępna. Bardziej realne, chociaż w przypadku wielu badań i tak zbyt optymistyczne, jest założenie, że dostępne są tylko oceny wariancji estymatorów $h_{ij}(\hat{\mathbf{p}})$. Oznaczmy te oceny przez $\hat{D}^2[h_{ij}(\hat{\mathbf{p}})]$ (są to elementy diagonalne macierzy $\hat{\mathbf{V}}[\mathbf{h}(\hat{\mathbf{p}})]/n$). Wykorzystując je można podać modyfikację statystyki (22). Zmodyfikowana statystyka jest postaci:

$$\chi_{m(\delta)}^2 = \chi_P^2/\hat{\delta}, \text{ gdzie } \hat{\delta} = \frac{n}{(r-1)(c-1)} \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\hat{D}^2[h_{ij}(\hat{p})]}{\hat{p}_{i.}\,\hat{p}_{.j}} \tag{26}$$

Ponieważ znacznie łatwiejsze (a tym samym i realniejsze) jest otrzymanie ocen elementów diagonalnych macierzy kowariancji $\mathbf{V}/n$, czyli $\hat{D}^2(\hat{p}_{ij})$, to cytowani autorzy rozpatrywali również następującą modyfikację:

$$\chi_{m(d)}^2 = \chi_P^2/\hat{d}, \text{ gdzie } \hat{d} = \frac{n}{rc-1} \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\hat{D}^2(\hat{p}_{ij})}{\hat{p}_{ij}} \tag{27}$$

Jednakże i w tym przypadku należy wątpić, aby były dostępne oceny wariancji dla ocen prawdopodobieństw dla każdej celi tablicy. Statystyk najczęściej może liczyć tylko na dostępność $\hat{D}^2(\hat{p}_{i.})$, dla $i=1, ..., r-1$ oraz $\hat{D}^2(\hat{p}_{.j})$, dla $j=1, ..., c-1$. W takich przypadkach dostępne są dwie modyfikacje statystyki χ_P^2, określone parą wzorów:

$$\chi_{m(\hat{d}_r)}^2 = \chi_P^2/\hat{d}_r, \text{ gdzie } \hat{d}_r = \frac{n}{r-1} \sum_{i=1}^{r} \frac{\hat{D}^2(\hat{p}_{i.})}{\hat{p}_{i.}} \tag{28}$$

oraz

$$\chi_{m(\hat{d}_c)}^2 = \chi_P^2/\hat{d}_c, \text{ gdzie } \hat{d}_c = \frac{n}{c-1} \sum_{j=1}^{c} \frac{\hat{D}^2(\hat{p}_{.j})}{\hat{p}_{.j}}. \tag{29}$$

Badania symulacyjne przeprowadzone przez D. Holta, A. J. Scotta i P. D. Ewingsa [6] oraz J. N. K. Rao i A. J. Scotta [19] potwierdziły wysoką zgodność rozkładu statystyki $\chi_{m(\delta)}^2$ z rozkładem $\chi^2[(r-1)(c-1)]$. Należy tu zwrócić uwagę na fakt, że w przeciwieństwie do testów zgodności i jednorodności χ^2, zastosowanie statystyki χ_P^2 do weryfikacji hipotezy o niezależności cech w dwuwymiarowej tablicy kontygencyjnej nie zawsze prowadziło do zwiększenia rozmiaru testu w stosunku do założonego.

A. J. Scott i J. N. K. Rao stwierdzili, że zawsze $\hat{\delta}$ jest mniejsze od $\hat{d}$ oraz od $min(\hat{d}_r, \hat{d}_c)$. Tym samym stosowanie modyfikacji $\chi^2_{m(\hat{d})}$, $\chi^2_{m(\hat{d}_r)}$ oraz $\chi^2_{m(\hat{d}_c)}$ prowadzi zawsze do zaniżenia rozmiaru testu w stosunku do założonego.

Badania empiryczne przeprowadzone przez cytowaną wyżej trójkę autorską pokazały, że przy założonym poziomie istotności $\alpha = 0{,}05$, rzeczywisty może wahać się od 0,00 do 0,05. Na ogół, choć nie zawsze, lepsze przybliżenie daje użycie $\hat{d}_m = min(\hat{d}_r, \hat{d}_c)$ zamiast $\hat{d}$.

WNIOSKI KOŃCOWE

W praktyce badań reprezentacyjnych najczęściej używane są schematy wielostopniowe, z warstwowaniem jednostek losowania na różnych stopniach oraz różnicujące prawdopodobieństwa wyboru. Dodatkową komplikację stanowi zastosowanie schematów wielofazowych i uogólnianie wyników badania nie na całą populację, lecz na jej podzbiory, zwane dziedzinami studiów, których nie można wyodrębnić przed przeprowadzeniem badania (przed wylosowaniem próby). To wszystko powoduje, że efekt schematu losowania znacznie przekracza jedność (głównie w wyniku występowania korelacji wewnątrzzespołowej) i stosowanie w takich przypadkach klasycznych testów χ^2 Pearsona prowadzi do tak znacznego zwiększenia ich rozmiaru (w stosunku do założonego), że czyni je nieprzydatnymi.

Dlatego też statystyki χ^2_P należałoby zastępować w takich przypadkach statystykami Walda. Wymagałoby to jednak znajomości ocen pełnych macierzy kowariancji. Wydaje się rzeczą niemożliwą, aby w takich badaniach reprezentacyjnych prowadzonych przez GUS, jak np. mikrospis, badanie dzietności kobiet itp. dla każdej zestawionej tablicy korelacyjnej była szacowana macierz kowariancji, gdyż koszt takiego postępowania byłby kilkadziesięciokrotnie wyższy niż przy tradycyjnym postępowaniu (przy tablicy o wymiarach 5×5 macierz kowariancji ma wymiary 24×24). Zwiększyłyby się też odpowiednio rozmiary publikacji zawierających wyniki badań.

Stosowanie opisanych modyfikacji statystyk χ^2_P znacznie zmniejsza pracochłonność obliczeń w stosunku do χ^2_W, ale mimo to pozostaje oszacowanie wariancji dla estymatorów odpowiednich częstości. Dlatego też istotną rolę odgrywają tu metody szacowania wariancji estymatorów złożonych. Wydaje się, że największą przydatność mogą mieć metody *BRR* czy wykorzystujące rozwinięcie w szereg Taylora.

Z drugiej strony wiadomo, że niektóre tablice są zestawiane na wszelki przypadek i najczęściej nie są przez nikogo wykorzystywane. Dlatego też wydaje się, że rozsądnym byłoby publikowanie dokładnego opisu schematu losowania oraz tablic z najważniejszymi danymi i udostępnianie zain-

teresowanym instytucjom kopii taśm (dysków) z interesującymi je danymi. Przy takim rozwiązaniu instytucje mogłyby same szacować wariancje (macierze kowariancji), w zależności od potrzeb przeprowadzanych przez siebie analiz.

BIBLIOGRAFIA

[1] P.Armitage (1966): *The chi-square test for heterogeneity of proportions after adjustment for stratification.* JRSS s. B 28.1, str. 150—163.

[2] J. E. Cohen (1976): *The distribution of the chi-squared staistic under clustered sampling from contingency tables.* JASA 71, str. 665—670.

[3] I. P. Fellegi (1980): *Approximate tests of independence and goodness of fit based on stratified multistage samples.* JASA 75, str. 261—268.

[4] M. R. Frankel (1971): *Inference from survey sampling*: ann *empirical investigation.* University of Michigan.Am Arbor, Michigan.

[5] D. H. Freeman (Jr), J. L. Freeman i D. B. Brock (1977): *Modularization for the analysis of complex sample survey data.* Bulletin of the ISI, 47.3, str. 3—20.

[6] D. Holt, A. J. Scott i P. D. Ewings (1980): *Chi-squared test with survey data.* JRSS s.A 143.3, str. 303—320.

[7] D. Holt, T. M. F. Smith i P. D. Winter (1980): *Regression analysis of data from complex surveys.* JRSS s.A 143.4, str. 474—487.

[8] H. L. Jones (1968): *The analysis of variance of data from stratified subsamples.* JASA 63, str. 64—86.

[9] L. Kish (1965): *Survey sampling.* JW&S. New York-London-Sydney.

[10] L. Kish i M. R. Frankel (1970): *Balanced repeated replication for standard errors.* JASA 65, str. 1071—1094.

[11] L. Kish i M. R. Frankel (1974): *Inference from complex samples.* JRSS s.B 36.1, str. 1—37.

[12] G. G. Koch, D. H. Freeman (Jr) i J. L. Freeman (1975): *Strategies in the multivariate analysis of data from complex surveys.* Review of the ISI 43.1, str. 59—78.

[13] H. S. Konijn (1962): *Regression analysis in sample surveys.* JASA 57, str. 590—606.

[14] P. J. Mc Carthy (1969): *Pseudo replication: half samples.* Review of the ISI 37.3, str. 239—364.

[15] G. Nathan (1969): *Tests of independence in contingency tables from stratified samples.* Artykuł w publikacji wydanej przez N. L. Johnsona i H. Smitha pt. New Developments in Survey Sampling. New York, JW&S, Inc. str. 578—600.

[16] G. Nathan i D. Holt (1980): *The effect of survey design on regression analysis.* JRSS s.B 42.3, str. 377—386.

[17] D. Pfefferman i G. Nathan (1977): *Regression analysis of data from complex samples.* Bulletin of the ISI, 47.3, str. 21—42.

[18] J. N. K. Rao i M. A. Hidiroglou (1981): *Chisquare tests for the analysis of categorial data from the Canada health survey.* Bulletin of the ISI 49, str. 699—718.

[19] J. N. K. Rao i A. J. Scott (1981): *The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables.* JASA 76, str. 221—230.

[20] J. N. K. Rao i A. J. Scott (1984): *On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data.* The Annals of Statistics 12.1, str. 46—60.

[21] A. J. Scott i D. Holt (1982): *The effect of two-stage sampling on ordinary least squares methods.* JASA 77, str. 848—854.

[22] J. Sedransk (1965): *Analytical surveys with cluster sampling.* JRSS s.B 27.2, str. 264—278.

[23] J. Sedransk (1967): *Designing some multifactor analytical studies.* JASA 62, str. 1121—1139.

[24] H. Solomon i M. A. Stephens (1977): *Distribution of sum of weighted chi-square variables.* JASA 72, str. 881—885.

PROBLEMY ESTYMACJI W BADANIACH REPREZENTACYJNYCH

Niekompletne dane, ich korekta i efekty[1])

Richard Platek

W ostatnich kilku latach w większości państw znacznie wzrosła liczba badań statystycznych, jako źródła szacunków, dotyczących najróżniejszych dziedzin. Na dokładność tych szacunków wpływa wiele czynników. Jednym z nich jest brak odpowiedzi oraz niekompletne i nielogiczne dane. W każdym badaniu, bez względu na rodzaj i zastosowaną metodę zbierania danych występuje zjawisko braku odpowiedzi, z następujących powodów:

a) nie wszystkie jednostki danej populacji zostały wzięte pod uwagę w operacie badania (losowania);
b) nie odszukano jednostek figurujących w tym operacie, a wchodzących w skład badania;
c) jednostki odmówiły udziału w badaniu.

Oprócz zjawiska braku odpowiedzi w badaniu mogą wystąpić pozycje, które są wypełnione częściowo lub zawierają niepoprawne odpowiedzi.

Rozważano często kwestię, czy w przypadku pomijania czynnika braku odpowiedzi i niekompletnych danych, szacunki oparte na informacjach otrzymanych spełniają cele danego badania. Na przykład, dokonując szacunku przy założeniu, że dana pozycja badania obejmuje 15% populacji, zastanawiano się jaki wywrze wpływ na szacunek wskaźnik braku odpowiedzi, który wynosiłby 15 lub 20%? W jakim stopniu potencjalny błąd systematyczny (obciążenie), spowodowany brakiem odpowiedzi mógłby przeważyć błąd, wynikający ze schematu losowania próby?

---

[1]) Referat przekazany został przez Autora w języku angielskim. Tytuł oryginału: *INCOMPLETE DATA, ADJUSTMENTS AND EFECTS*. Tłumaczyła Beata Jakubczak.
Richard Platek — uprzednio dyrektor wykonawczy Biura Statystycznego Kanady, obecnie doradca w Biurze Statystycznym i Banku Światowym.

Większość statystyków-praktyków oraz analityków danych uważa stopień braku odpowiedzi i niekompletność danych za wskazówkę co do jakości danych, ponieważ wpływa on na szacunki przez wprowadzenie zarówno możliwego błędu systematycznego, jak i wzrostu wariancji spowodowanej zmniejszeniem efektywnej wielkości próby. Związek między błędem systematycznym a częstością braku odpowiedzi jest mniej istotny, ponieważ wynika on zarówno z częstości braku odpowiedzi, jak i z różnic w cechach respondentów i nierespondentów.

Uważa się, że w wielu badaniach statystycznych prowadzonych na szeroką skalę, gdyby zmierzono błędy spowodowane brakiem odpowiedzi lub niekompletnością danych, znacznie przekroczyłyby one błędy losowe, przynajmniej w zakresie szczegółowych dezagregacji. Jednakże, w większości badań identyfikowane są jedynie błędy losowe, a inne składniki błędów nie są odpowiednio zapisywane i analizowane. Ponadto występują potencjalne źródła błędów systematycznych i zlekceważenie ich wpływu na szacunki doprowadzić może do wyników o jakości nie do zaakceptowania. Dlatego też niwelowanie wpływu braku odpowiedzi oraz odpowiedzi błędnych i niekompletnych jest bardzo ważne i powinno być dokonywane na różnych etapach, włącznie z planowaniem badania, zbieraniem danych, przetwarzaniem i dokonywaniem szacunków.

Jednym ze sposobów rozwiązywania problemu braku odpowiedzi, po zebraniu danych, jest zastosowanie metody imputacji lub korekty wag na etapie przetwarzania i dokonywania szacunków. Podczas gdy wyrównania w wyniku braku odpowiedzi mogą mniej lub więcej skutecznie zredukować błąd systematyczny, to dobrze przygotowane zbieranie danych utrzymywać będzie brak odpowiedzi na poziomie możliwym do przyjęcia i przy rozsądnych kosztach, zmniejszając tym samym konieczność stosowania wyrównywania statystycznego.

W tym referacie zagadnienia braku odpowiedzi, niekompletności danych, włącznie z odpowiedziami błędnymi w badaniach gospodarstw domowych, będą przedstawione na tle źródeł ich powstawania, różnych metod redukowania, jak również wyrównania ich finalnych szacunków.

ZAGADNIENIE BRAKU ODPOWIEDZI

Pojęcie braku odpowiedzi może być zdefiniowane jako niepowodzenie w uzyskaniu możliwej do wykorzystania informacji od osoby ankietowanej, która wchodzi w skład próby badania. Wyróżnia się dwa rodzaje braku odpowiedzi:

a) podmiotowy brak odpowiedzi, gdy nie otrzymano od adresata kwestionariusza badania,
b) przedmiotowy brak odpowiedzi, kiedy otrzymano kwestionariusz badania od osoby ankietowanej, przy braku odpowiedzi na jedno lub więcej pytań.

Brak odpowiedzi jest spowodowany trudnościami operacyjnymi, ograniczeniami czasu i kosztów, brakiem współpracy z respondentem, niemożnością lub brakiem zaangażowania ankietera w odszukaniu osoby brakującej lub innymi przyczynami. Ostrość czynnika braku odpowiedzi mierzy się za pomocą wskaźnika obliczanego jako procent gospodarstw domowych, które nie udzieliły odpowiedzi wśród gospodarstw objętych próbą.

Wskaźniki braku odpowiedzi różnią się znacznie w poszczególnych badaniach. W niektórych badaniach sięgają one 40% lub więcej, a w innych mogą wynosić zaledwie około 4%. To czy wskaźniki braku odpowiedzi są zbyt wysokie lub za niskie zależy od celu danego badania. Jeżeli celem badania jest dokonanie szacunku informacji odpowiadającej 15% populacji, w zależności od dokładności z jaką należy dokonać szacunku tych 15%, to nawet 5% braku odpowiedzi może mieć istotny wpływ na wynik szacunku, a w szczególności, gdy charakterystyki nierespondentów są skorelowane z badanymi zmiennymi. Z drugiej strony wysoki wskaźnik braku odpowiedzi niekoniecznie oznacza, że badanie nie może dostarczyć użytecznej informacji. Przypuśćmy, że celem badania jest otrzymanie informacji na temat, czy 15% populacji kupiłoby dany produkt. Jeżeli w badaniu tym 20% udzieli odpowiedzi twierdzącej, 30% negatywnej, a 50% nie dałoby żadnej odpowiedzi, cel tego badania zostanie spełniony, nawet gdyby nie wszystkie osoby, które udzieliły odpowiedzi należały do grupy osób, która nie kupiłaby tego produktu. Ogólnie rzecz biorąc, im wyższy jest wskaźnik braku odpowiedzi tym wyższy jest możliwy błąd systematyczny szacunku i tym mniej prawdopodobne jest wykonanie celu badania.

Wielkość braku odpowiedzi nie może być rozwiązana przez zwiększenie próby, ponieważ w przypadku istnienia braków odpowiedzi próba nie będzie wówczas próbą losową. Trudność polega na tym, że nierespondenci różnią się w określonym i w różnym stopniu od respondentów. Można założyć, że każda wybrana osoba jest potencjalnym respondentem, to znaczy wybrana osoba udzieli odpowiedzi z prawdopodobieństwem δ_i i nie odpowie z prawdopodobieństwem $1-\delta_i$. Jeżeli prawdopodobieństwo odpowiedzi jest takie samo dla wszystkich i, sytuacja ta zostanie łatwo unormowana przez zastosowanie korekty (obniżania) wag respondentów. Ale prawdopodobieństwo odpowiedzi może zależeć od interesującej nas charakterystyki, a zastosowana korekta wag respondentów, w celu uwzględnienia braku odpowiedzi, spowoduje obciążenie. Wielkość systematycznego błędu z powodu braku odpowiedzi będzie zależała od związku pomiędzy interesującą nas cechą a prawdopodobieństwem odpowiedzi. Problem ten może być przedstawiony za pomocą następujących prostych przykładów, do różnych schematów losowania:

a. Jednostopniowe losowanie proste

Z populacji N jednostek wybrano n stosując losowanie proste bez zwracania, a przedmiotem badania jest ilościowa zmienna Y (np. dochód,

liczba wypadków itd.). Przypuśćmy, że prawdopodobieństwo uzyskania odpowiedzi zmniejszy się kiedy wartość y zwiększy się; gdyby nie było braków odpowiedzi wartość globalna $Y=\sum_{t=1}^{N} y_t$ oszacowana zostałaby właściwie jako $\hat{Y}=\frac{N}{n}\sum_{i=1}^{n} y_i$, ale gdy Y jest zaobserwowane dla $n'(<n)$ jednostek oraz $\bar{Y}=\frac{N}{n'}\sum_{i=1}^{n'} y_i$, użyte jest w celu oceny wartości globalnej Y sytuacja może być inna; $\bar{Y}$ nie byłoby obciążone, gdyby prawdopodobieństwo odpowiedzi nie zależało od wartości y, ponieważ jednak prawdopodobieństwa odpowiedzi są niskie dla wielkich wartości y, $\bar{Y}$ niedoszacowywuje wartości globalnej; inaczej mówiąc, mamy do czynienia z negatywną korelacją pomiędzy prawdopodobieństwem odpowiedzi i wielkością cechy, co powoduje ujemny błąd systematyczny z powodu występowania braku odpowiedzi.

b. Losowanie jednostopniowe z prawdopodobieństwami proporcjonalnymi do wielkości jednostki losowania

Z populacji N jednostek (firm) wybrano n w celu dokonania szacunku $Y=\sum_{t=1}^{N} y_t$, gdzie y jest cechą ilościową (np. produkcją itd.), a p_t jest miarą wielkości jednostki t, $t=1,2,...,N$; $\sum_{t=1}^{N} p_t=1$, gdzie $\Pi_t=np_t=$ prawdopodobieństwu, że jednostka i-ta jest w próbie, y_i jest zaobserwowaną wartością dla jednostki i, jeżeli została ona wybrana i odpowiedziała. Gdy nie wystąpiło zjawisko braku odpowiedzi $\hat{Y}=\sum_{i=1}^{n}\frac{y_i}{\Pi_i}$ jest nieobciążonym estymatorem Y. Przypuśćmy, że wartości y zaobserwowane są tylko dla $n'(<n)$ z wybranych n, wtedy $\bar{Y}=\left(\frac{n}{n'}\right)\sum_{i=1}^{n'}\frac{y_i}{\Pi_i}$ jest używane w celu dokonania szacunku Y i nie będzie obciążone, jeżeli prawdopodobieństwa odpowiedzi będą skorelowane ze wskaźnikiem y_i/p_i. Obciążenie będzie prawdopodobnie małe, w porównaniu z wartością globalną Y, jeżeli ilorazy y_i/p_i będą się charakteryzowały małą zmiennością. Zwykle rozważane losowanie jest stosowane, gdy istnieje pozytywna współzależność pomiędzy wartościami y_i a wielkościami p_i tych jednostek. Często ta wysoka współzależność powoduje małą zmienność w ilorazach y_i/p_i, kiedy regresja y_i na p_i przechodzi przez początek układu.

c. Wielostopniowe warstwowe losowanie z prawdopodobieństwami proporcjonalnymi do wielkości jednostki losowania

W badaniach gospodarstw domowych wybierane są gospodarstwa za pomocą schematu wielostopniowego, warstwowego (PPS) i schemat ten daje

zwykle próby automatycznie wyważone. W niektórych badaniach informacja o wszystkich wybranych członkach w tym gospodarstwie może być uzyskana od jednego z nich. Możliwe jest, że czynnik braku odpowiedzi będzie większy w przypadku gospodarstwa jednoosobowego niż wieloosobowego. W takich przypadkach, jeżeli w korekcie braku odpowiedzi, przy pomocy ważenia, nie weźmie się pod uwagę wielkości gospodarstwa domowego, powstały szacunek populacji będzie obciążony błędem o tendencji zwyżkowej. W tym kontekście rozważmy inny przykład. Przypuśćmy, że przedmiotem badania jest cecha jakościowa, tj. obecność lub brak wyróżnionego wariantu cechy. Przypuśćmy ponadto, że prawdopodobień stwo odpowiedzi jest wysokie, jeżeli wyróżniony wariant nie występuje i niskie, jeżeli występuje. Gdy poprzez korektę wag respondentów otrzymamy szacunek liczby osób wyróżnionych, wynik ten będzie niedoszacowaniem, gdyż będziemy mieli wówczas negatywną korelację pomiędzy prawdopodobieństwami odpowiedzi i posiadaniem wyróżnionego wariantu cechy.

Optymalną sytuacją, aby całkowicie usunąć czynnik braku odpowiedzi, byłoby podniesienie wagi każdej odpowiedzi z próby przez podzielenie jej przez iloczyn prawdopodobieństwa wyboru i prawdopodobieństwa otrzymania prawdziwej odpowiedzi. Jednakże, zadanie to nie jest wykonalne, ponieważ nie znamy prawdopodobieństwa otrzymania prawdziwej odpowiedzi od każdej osoby. W praktyce stosujemy przeciętne prawdopodobieństwa odpowiedzi oszacowane za pomocą wskaźników odpowiedzi, w celu skorygowania wag w ocenach. Rezultat takiej metody, stosowanej w celu zredukowania błędu spowodowanego brakiem odpowiedzi, będzie zależał od stopnia zależności pomiędzy prawdopodobieństwami odpowiedzi a ich wskaźnikami.

Najbardziej pożądanym trybem postępowania, w przypadku braku odpowiedzi, jest zminimalizowanie ich częstości. Jednakże każda systematyczna próba kontroli częstości braku odpowiedzi musi być oparta o dokładne rozpoznanie procesu jego powstawania.

PODEJŚCIE DO ZAGADNIENIA BRAKU ODPOWIEDZI

Plan badania składa się z trzech etapów: a) schematu losowania; b) zbierania danych; c) estymacji.

Żaden z tych etapów nie może być przeprowadzony li tylko na gruncie czysto technicznym lub czysto praktycznym. Plan badania powinien być opracowany w oparciu o założenie — co jest praktycznie wykonalne i pożądane z teoretycznego punktu widzenia, a jednocześnie musi sprostać wymaganiom użytkownika. Znaczenie czynników braku odpowiedzi, poprzez jego wpływ na wyniki badania, jest istotną sprawą planu badania i nie może być zlekceważone. Na każdym etapie badania muszą być ustalone zasady w celu kontrolowania częstotliwości braku odpowiedzi i zminimalizowania jego wpływu na wynik finalny.

Chociaż rzeczywisty brak odpowiedzi występuje na etapie zbierania danych, to na jego zmniejszenie wpłynąć można już na etapie planowania, przyjmując różne rozwiązania techniczne na brak odpowiedzi.

Ponadto, częstotliwość braku odpowiedzi można zmniejszyć poprzez staranne i właściwe przygotowanie zbierania danych, ze szczególnym zwróceniem uwagi na metody przeprowadzania wywiadu i motywację respondentów oraz dobór osób przeprowadzających wywiad.

Na etapie przetwarzania i estymacji próbuje się zminimalizować wpływ braku odpowiedzi na wyniki ostateczne w tym celu przeprowadza się imputację brakujących informacji.

Planowanie i rozwój

Jednym z najważniejszych czynników w planowaniu badania jest ustalenie poziomu tolerancji braku odpowiedzi. Doświadczony planista badania statystycznego może dość dokładnie określić poziom odpowiedzi jakiego należy się spodziewać w różnych warunkach danego badania. Na przykład, dla szacunków w skali krajowej z dużej próby, wpływ braku odpowiedzi na wiariancję losowania i wariancję odpowiedzi prawdopodobnie będzie znikomy i obciążenie będzie przypuszczalnie przeważającym składnikiem średniego błędu kwadratowego (Mean Square Error — MSE). Jednakże, do szacunków dla części kraju wariancje będą prawdopodobnie duże, więc obciążenie może mieć stosunkowo mniejszą wagę.

Koszt badania jest również ważnym elementem, który oddziałuje na wiele czynników w procesie rozwoju badania włącznie z czynnikiem braku odpowiedzi. Ważnym procesem jest porównanie tych czynników z kosztami, w celu osiągnięcia poziomu wskaźnika braku odpowiedzi wystarczająco niskiego, odpowiedniego do celu badania. Należy mieć świadomość, że lepiej jest czasami przyjąć nieco mniejszą próbę od planowanej, a środki pieniężne przeznaczyć na odpowiednią technikę zbierania danych, ich kontrolę i właściwe metody estymacyjne. To byłoby szczególnie korzystne, gdyby planujący badanie podejrzewał duże różnice charakterystyk respondentów i nierespondentów.

W planowaniu i rozwoju badania należy brać pod uwagę wiele czynników. Czynniki te mogą być sklasyfikowane w trzech grupach:

Grupa I: a) wielkość próby, b) warstwowanie, c) stopień grupowania w zespoły, d) podział próby (sample allocation), e) metoda wyboru.

Grupa II: a) operat losowania próby, b) metoda przeprowadzania wywiadu, c) selekcja, szkolenie i kontrola kadry, d) objętość kwestionariusza i liczba słów, e) drażliwość pytań, f) rodzaj obszaru, na którym przeprowadza się badanie, g) wykonalność i koszty powtórnych wizyt, h) reklama.

Grupa III: a) opracowanie i imputacja, b) estymacja, c) estymacja wariancji.

Wszystkie te działania wpływają na *MSE*, który dostarcza pewnej miary dokładności danych. Załóżmy, że *MSE* może być rozłożony na następujące składniki:

$$MSE = V_s + V_R + V_{CR} + (B_S + B_R)^2$$

gdzie: V_S — wariancja losowania,
V_R — wariancja odpowiedzi prostej,
V_{CR} — wariancja odpowiedzi skorelowanej,
B_S — obciążenie losowe,
B_R — obciążenie odpowiedzi (włącznie z obciążeniem z powodu braku odpowiedzi).

Na wariancję losowania (V_S) i obciążenie losowe (B_S) wpływają wszystkie czynniki z grupy I i III, jak również częstotliwość występowania braku odpowiedzi. Ważne jest stwierdzenie, że w każdym badaniu występują specyficzne wymagania co do schematu losowania próby, kwestionariusza oraz metody przeprowadzania wywiadu. Na przykład, schemat indywidualnego badania może powodować wyższy wskaźnik braku odpowiedzi niż schemat badania zespołowego. Może on wynikać z konieczności częstych podróży, w przypadku wywiadów osobistych, kiedy powtórne wizyty muszą być ograniczane z powodu wysokich kosztów. Im większa częstotliwość braku odpowiedzi tym większy ich wpływ na wariancję losowania i obciążenia. Na nielosowe składniki V_R, V_{CR}, B_R, średniego błędu kwadratowego (MSE), wpływają wszystkie czynniki z grupy II. Ponadto, oczywiste jest, że wszystkie czynniki z grupy II są potencjalnym źródłem i przyczyną braku odpowiedzi. Długość ankiety i drażliwość pytań niewątpliwie wpływają na częstotliwość występowania braku odpowiedzi. Dlatego też, jeżeli będziemy zwracać szczególną uwagę na czynniki z grupy II na etapie planowania, będzie można pomniejszyć błędy, wynikające z braku odpowiedzi.

Etap zbierania danych

Omawiając różne podejścia, dla zminimalizowania zjawiska braku odpowiedzi, możemy rozróżnić dwa ich rodzaje.

Rodzaj pierwszy określany jako „nikogo nie ma w domu" lub „czasowo nieobecny" jest w rzeczywistości sprawą „braku kontaktu". Drugim rodzajem, jest sprawa prawdziwego braku odpowiedzi, gdzie kontakt z respondentem zaistaniał, lecz nie uzyskano odpowiedzi możliwej do zaakceptowania.

Problem „braku kontaktu" zwykle jest poddawany rozwiązaniom operacyjnym. Podczas wywiadu telefonicznego lub osobistego ważny jest czas i sposób rozmowy z respondentem. W badaniu przeprowadzonym drogą pocztową istotne jest zapewnienie właściwych adresów na liście

adresowej, skuteczne działania sprawdzające oraz posiadanie odpowiednich materiałów. Przyczyny braku odpowiedzi spowodowane „nikogo nie ma w domu" lub „czasowo nieobecny" stanowią istotną wskazówkę w zakresie operatywnego działania.

Sprawa odmowy stanowi inny zespół problemów. Należy stwierdzić, że wskaźniki odmowy nie zawsze są tak oczywiste, jak należałoby przypuszczać. Osoba przeprowadzająca wywiad może zapisać odmowę jako „nikogo nie ma w domu", wówczas gdy respondent po prostu nie otwiera drzwi, jako sposób na nieudzielenie odpowiedzi. Dlatego ten przypadek może być zapisany jako „nikogo nie ma w domu". W badaniu przeprowadzonym drogą pocztową nie zawsze jest się pewnym, czy respondent otrzymał ankietę, a jeżeli ją otrzymał, to czy po prostu nie zapomniał jej odesłać. Dlatego też, nie łatwo jest rozróżnić prawdziwy brak odpowiedzi od innych powodów nieudzielenia odpowiedzi. Odpowiedź niepoprawna stanowi jeszcze inny zespół problemów, w przypadku gdy przeprowadzający wywiad jest osobą niedoświadczoną; może nie zdawać sobie sprawy z faktu, że informacje są niepoprawne lub nielogiczne do momentu, gdy zostanie to ujawnione w trakcie opracowywania materiału. Ponadto, osoba przeprowadzająca wywiad może przez pomyłkę zapisać odpowiedź w niewłaściwym miejscu na ankiecie, co powoduje, że informacje są niepoprawne i zostają odrzucone. Jednakże, bez względu na to, jak brak odpowiedzi zostanie zapisany, sprawa ta powinna mobilizować respondenta do udzielenia właściwej odpowiedzi.

Jaką motywacją kieruje się respondent, jako osoba o neutralnym podejściu do badania, że udzieli odpowiedzi albo nie? Takie czynniki, jak trudność w rozumieniu pytań, zabranie czasu respondentom, chęć ochrony prywatności, obojętność, trudności w przypomnieniu sobie informacji, skrępowanie lub pytania osobiste są przyczynami nieudzielania przez respondentów odpowiedzi. Z drugiej strony, motywacja do udzielenia odpowiedzi wynika z zainteresowania badaniem, chęci pomocy, obowiązku, rozumienia znaczenia wyników badania.

Problemem jest, jak zdobyć motywację pozytywną i zredukować motywację negatywną, aż do otrzymania pozytywnej postawy dla udzielania właściwej odpowiedzi. Istotnym elementem badania jest respondent i dlatego osoba organizująca badanie musi wziąć pod uwagę wszystkie okoliczności wpływające na pozytywną motywację do udzielenia odpowiedzi.

Wstępna korespondencja, podanie przykładów wykorzystywania informacji, broszury opisujące cel oraz znaczenie badania, często stanowią właściwy sposób na uniknięcie wrogości i braku zaufania.

Treść kwestionariusza może być związana z ingerencją w sprawy prywatne, toteż różnie reagować mogą poszczególni respondenci. Podejmuje się wiele działań dla zminimalizowania negatywnej postawy respondenta; działania te powinny być dostosowane do określonej sytuacji. W niektórych

przypadkach lepiej jest pozwolić respondentowi odpowiedzieć w całkowicie anonimowy sposób. Można to przeprowadzić przez ponumerowanie kwestionariuszy, bez personalnej identyfikacji. Ważne jest posiadanie oznaczenia kodowego danego obszaru lub oznaczenia próby w celu ważenia i oszacowania na etapie sprawdzania. Należy zatem zadbać o to, aby respondent nie odczuł tych działań jako sposobu na identyfikację udzielonych przez niego odpowiedzi z jego osobą.

Jedną z metod niwelowania wpływu czynnika braku odpowiedzi na etapie zbierania danych jest zastąpienie nierespondentów innymi osobami z danego terenu, wylosowanymi do próby (substytucja w terenie). Należy podkreślić jednak, że jest to zawsze brak odpowiedzi, który zastąpiony został formą imputacji. Istnieją dwa podstawowe rodzaje sybstytucji w terenie, które mogą być stosowane w badaniach: a) **dobór substytutu losowego,** b) **dobór specjalnego substytutu.**

W metodzie doboru substytutu losowego, aby zastąpić każdego nierespondenta, wybiera się losowo dodatkową osobę. W wielu działaniach doboru substytutu losowego zbieranie danych poprzedzone jest selekcją potencjalnych substytutów dla uniknięcia opóźnień i problemów, które mogłyby zaistnieć, gdyby substytuty były wybierane w czasie lub po zbieraniu danych.

Celem procedury, wykorzystującej specjalnie określone substytuty (np. sąsiad), jest znalezienie substytutu, mającego podobną do nierespondenta charakterystykę. Problemem, związanym z doborem specjalnego substytutu jest to, czy substytucja dostarczy lepszych informacji zastępczych dla nierespondentów niż informacje dostarczane przez alternatywne metody imputacji. Niewątpliwie istnieją dobre i złe strony stosowania procedury substytucji. Pierwsza korzyść wynika z faktu, że jest to wygodny sposób równoważenia próby w odniesieniu do wielkości próby. Inną korzyścią jest redukcja wariancji losowania w wyniku zwiększenia efektywnej wielkości próby.

Do negatywów stosowania substytucji zaliczyć należy tendencję do jej nadużywania zamiast starania się o uzyskanie odpowiedzi od uprzednio wylosowanych osób. Dlatego stosowanie metody substytucji wymaga dokonania odpowiedniej kontroli, w celu podjęcia maksymalnego wysiłku dla uzyskania informacji od wytypowanych jednostek próby. Kolejnym faktem negatywnym jest pomijanie wskaźnika poziomu i częstotliwości stosowania substytucji, w przypadku kiedy zostały one określone dla danego badania.

Do innych stosowanych metod należy zaliczyć powtórne wizyty, badania zastępcze, kontrolę badań itd. (szczegółowe informacje patrz: „Methodology and Treatment for Non-response" — „Metodologia i zajmowanie się zjawiskiem braku odpowiedzi", R. Platek 1986, Instituto Vasco de Estadistica).

Na etapie przetwarzania i dokonywania szacunków informacje klasyfikowane są według: całkowitego braku odpowiedzi, częściowego braku odpowiedzi i odpowiedzi niepoprawnej.

Ważnym zagadnieniem jest posiadanie skutecznego systemu kontroli, włączonego w schemat losowania, który zapewnia, że wywiad przeprowadzony został z jednostkami wylosowanymi do próby (a nie innymi), że jednostki, które nie udzielały informacji są właściwie sklasyfikowane, np. brak odpowiedzi, nie istnieje osoba wylosowana, luki w operacie losowania próby są zidentyfikowane oraz, że dane wyjściowe są skompletowane itd.

Istnieje wiele sposobów opracowywania odpowiedzi niekompletnych i niepoprawnych. Wynikiem każdego z nich jest przypisanie określonej wartości nieważnym lub brakującym danym chyba, że została podjęta decyzja o opublikowaniu surowych danych. Przypisanie dodatkowych wartości nazywa się **imputacją**, która zakłada, że dodana wartość odnosi się do danej cechy nierespondenta. Dlatego odtworzony zostaje „czysty" zbiór danych, to znaczy, że do każdej badanej jednostki przypisana jest jakaś wartość. Zanim przystąpimy do omówienia różnych metod imputacji zbadajmy kiedy dochodzi do imputacji.

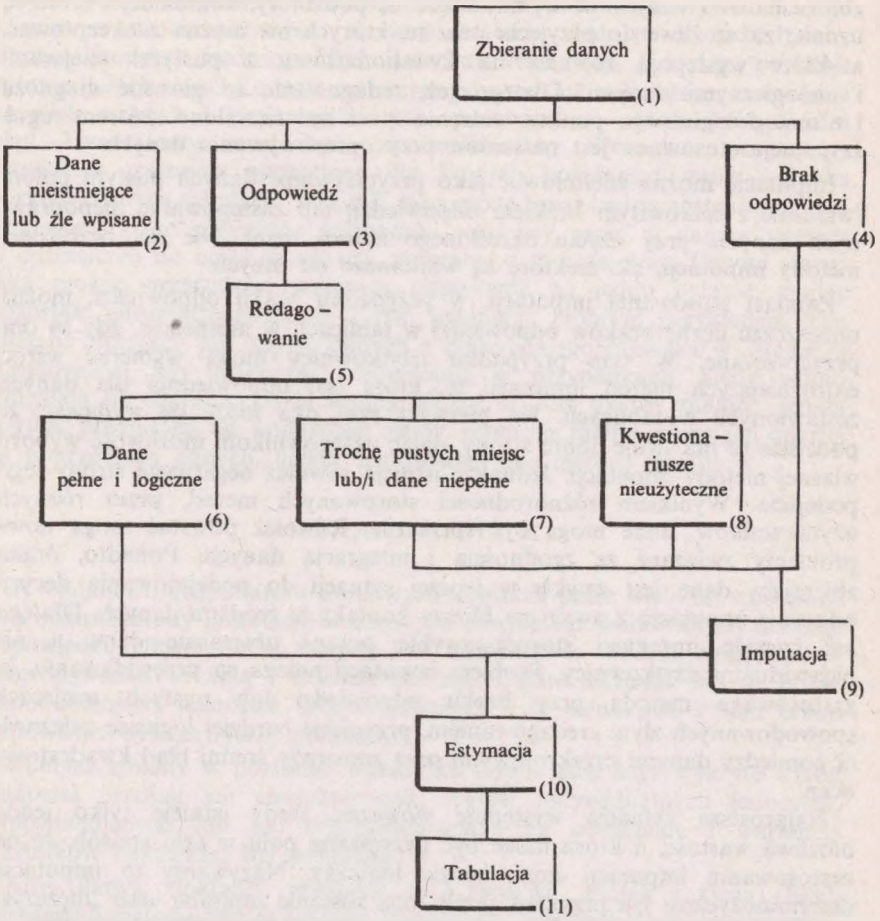
W trakcie przepływu informacji od zbioru danych do tabulacji otrzymujemy różne rodzaje odpowiedzi. Ich rodzaje ilustruje wykres.

Z wykresu wynika, że dwie spośród trzech grup kwestionariuszy, na etapie redagowania, wymagają dalszego działania, poprzedzającego estymację. Są to kwestionariusze nie nadające się do użycia (8), kwestionariusze zawierające puste miejsca (lub nielogiczne dane (7)) oraz kwestionariusze, zawierające brak odpowiedzi (4). Nie nadający się do użycia kwestionariusz może być sklasyfikowany, jako całkowity brak odpowiedzi lub może dotyczyć gospodarstwa domowego, którego odpowiedź zawiera miejsca puste lub dane nielogiczne. W obu przypadkach pożądane jest dalsze działanie w postaci imputacji (9). Nierespondenci, nie udzielający żadnych lub częściowych odpowiedzi, ważeni są zwykle w określony sposób we wszystkich badaniach, z wyjątkiem spisów powszechnych. Wyróżniamy dwie kategorie niekompletnych kwestionariuszy: a) dane nielogiczne, b) niezgodne zapisy.

Niezgodność zapisów może wynikać z niemożliwości logicznej lub wyjątkowej możliwości, ale mało prawdopodobnej. Naturalnym wydaje się, że jeżeli dane są logicznie niemożliwe, należy je skorygować nawet wówczas, gdy nie miały znaczącego wpływu na informacje. Skorygowanie to eliminuje kłopoty analityków w opublikowaniu wyników z danego badania.

W przypadku możliwych, lecz mało prawdopodobnych danych spotykamy się z trudnym wyborem pomiędzy pozostawieniem tego, co wydaje się być rozkładem nienaturalnym, a usunięciem ekstremalnych wartości tego rozkładu, które rzeczywiście mogą przedstawiać prawdziwą sytuację ży-

WYKRES 1. SCHEMAT PROCESU WŁAŚCIWEGO KAŻDEJ JEDNOSTCE W PRÓBIE



Wykres podany jest w formie uproszczonej i służy tylko dla naświetlenia tez referatu.

ciową. Ideałem byłoby dokonanie wyboru na podstawie doświadczenia, w zakresie mechanizmów błędu i istoty faktycznego rozkładu, opartego na wiedzy z danej dziedziny. W każdym razie, należy zidentyfikować ten fakt, a w przypadku napotkania niemożliwych lub bardzo mało prawdopodobnych danych, należy zastosować właściwą metodę ich opracowania (tzn. imputację) oraz właściwy sposób ich redagowania.

Istnieją zasadnicze rozróżnienia pomiędzy redagowaniem a imputacją. Przeanalizujmy zbiór wszystkich możliwych kombinacji kodowych na kwestionariuszu. Redagowanie (editing) można zdefiniować jako podział zbioru na dwa wzajemnie wykluczające się podzbiory: kombinacje, które są uznane za możliwe do przyjęcia oraz te, których nie można zaakceptować, a które występują również na kwestionariuszu z pustymi miejscami i nielogicznymi danymi. Dlatego też, redagowanie to głównie diagnoza i z metodologicznego punktu widzenia musi być określone zbiorem reguł. Imputacja stosowana jest natomiast przy opracowywaniu danych.

Imputację można zdefiniować jako przypisywanie danych pustym polom (włącznie z całkowitym brakiem odpowiedzi) lub zastępowanie niepoprawnych danych, przy użyciu określonego zbioru reguł. Nie ma bezbłędnej metody imputacji, ale niektóre są właściwsze od innych.

Zamiast stosowania imputacji, w przypadku braku odpowiedzi, można umieszczać liczbę braków odpowiedzi w tablicach w momencie, gdy są one przygotowane. W tym przypadku użytkownicy mogą wybierać wśród najróżniejszych metod imputacji tę, która jest odpowiednia dla danych zestawionych w tablicach. Na pierwszy rzut oka może się wydawać, że podejście to ma swoje dobre strony, dając użytkownikom możliwość wyboru własnej metody imputacji. Jednakże istnieją również negatywne strony tego podejścia. Wynikiem różnorodności stosowanych metod, przez różnych użytkowników, dane mogą być sprzeczne. Również powstać mogą nowe problemy związane ze zgodnością i integracją danych. Ponadto, organ zbierający dane jest zwykle w lepszej sytuacji do podejmowania decyzji odnośnie imputacji, z uwagi na bliższy kontakt ze źródłem danych. Dlatego też metodę imputacji stosują zwykle organa zbierające dane, a nie indywidualni użytkownicy. Problem imputacji polega na przewidywaniu, że zastosowana metoda przy braku odpowiedzi lub pustych miejscach spowodowanych złym zredagowaniem, przyniesie bardziej logiczne zależności pomiędzy danymi przekrojowymi oraz zmniejszy średni błąd kwadratowy ocen.

Najprostsza sytuacja występuje wówczas, kiedy istnieje tylko jedna możliwa wartość, a która może być przypisana polu w taki sposób, że po zastosowaniu imputacji zapis będzie logiczny. Nazywamy to imputacją deterministyczną. Na przykład, jeżeli żona zostanie zapisana jako „mężczyzna", będzie istniała tylko jedna wartość, umożliwiająca przypisanie płci, co sprawia, że będzie ona zgodna z informacją. Czasami może istnieć więcej wartości, które mogą powodować, że zapis będzie logiczny. Jeżeli tak się zdarzy, można wybrać konkretną wartość, która dominuje w stosunku do całkowitej częstotliwości lub jest bardziej możliwa do przyjęcia. Dobrym przykładem może być „Badanie Siły Roboczej" w Kanadzie w miesiącach od jesieni do wiosny. Dla osób w wieku 15 i 16 lat, gdyby nie zanotowano w tym badaniu żadnej informacji o pracy, przypisanoby im cechę „uczęszczają do szkoły", chociaż jest możliwe, że nie uczęszczają. Tak długo,

jak występowanie takich przypadków będzie wystarczająco małe, wynikiem takiej imputacji będzie niewielki wzrost obciążenia, jak i pewna redukcja w wariancji.

W innych sytuacjach, w których uzasadnione byłoby przypisywanie całego zakresu wartości, potrzebne są inne kryteria. Jednym z nich byłoby zminimalizowanie średniego błędu kwadratowego ocen. Jednakże, powstaje pytanie: jakich ocen? Ze stale rosnącym popytem na mikrodane, zestawione na wiele różnych i nieprzewidzianych sposobów, nie wiemy, który średni błąd kwadratowy musi być zminimalizowany. Ponadto, nie znamy wszystkich rodzajów agregatów, dla których pojedynczy zapis wchodzi w skład różnych tabulacji. Wskutek tego może lepiej będzie zastosować inne kryterium, które przyniesie najbardziej właściwy zapis w danym miejscu w odniesieniu do innej informacji, związanej z tym zapisem. Innymi słowy, jaką można przewidzieć najlepszą odpowiedź na jedno pytanie, znając pozostałe informacje związane z tym pytaniem. Dobrym przykładem tego rodzaju imputacji jest użycie danych z poprzedniego miesiąca w „Badaniu Kanadyjskiej Siły Roboczej", szczególnie w przypadkach, kiedy wolno zmieniają się charakterystyki demograficzne, trudno byłoby znaleźć lepszą wartość przypisaną dla konkretnej osoby. Jeżeli nie mamy informacji opartej na poprzednich danych, należy zastosować inne metody imputacji.

METODY IMPUTACJI

W badaniach gospodarstw domowych i spisach przez wiele lat stosowane były różne metody imputacji brakujących danych, spowodowanych brakiem odpowiedzi. Użycie danej metody jest uzasadnione wiedzą, umiejętnością przewidywania, intuicją i doświadczeniem. Często zakłada się, że prawdopodobieństwa jednostek odpowiadających były jednakowe i błąd braku odpowiedzi był zazwyczaj pomijany.

Chociaż zmiany w poziomie wskaźnika odpowiedzi były wykryte wśród jednostek według ich charakterystyk, wpływ indywidualnych jednostek, odpowiadających lub nie odpowiadających, na obciążenie i wariancję szacunków najczęściej był pomijany.

Aby ułatwić dokładne zbadanie tego wpływu, Platek i Gray (1980) opracowali metodologię w zakresie obciążenia i wariancji dla różnych metod imputacji. Wyprowadzenie wzoru dla obciążenia i wariancji ocen jest oparte na podstawowej zasadzie, że jednostka wybrana do próby odpowiada lub nie z pewnym prawdopodobieństwem odpowiedzi, przypisanym do tej jednostki. Jest to rozwinięcie podejścia, przyjętego przez Platka, Singh'a i Tremblay'a (1978), odnośnie spisów.

Definicja różnych metod imputacji zawiera:

1) wykorzystanie komórek dla imputacji; komórki te mogą być obszarami wyrównania (balancing areas) lub klasami ważenia (weighting classes) oraz

2) korekty wag, przy użyciu oszacowanych prawdopodobieństw odpowiedzi wewnątrz tych komórek.

Obszary wyrównania często określane są, dla zastosowania imputacji, jako równoważące się obszary uzależnione od schematu wyboru próby. Obszar wyrównania jest obszarem geograficznym, w którym wadliwa próba, wynikająca z brakujących danych jest powiększona do wyznaczonego poziomu za pomocą imputacji brakujących danych. Powszechnie, taki obszar jest warstwą, ale może stanowić inne, uzależnione od schematu losowania próby, obszary, takie jak: jednostka losowania pierwszego stopnia, zespół, grupy lub warstwy, lub nawet cała próba. Obszary wyrównania mogą być wyznaczone przed lub po przeprowadzeniu badania.

Klasy ważenia są to warstwy zdefiniowane po wylosowaniu próby, a utworzone na podstawie informacji o respondentach i nierespondentach danej próby. Informacja może być uzyskana od częściowych nierespondentów, których pewne cechy są znane, nawet gdyby cecha badana nie była znana tym jednostkom, jak na przykład, w przypadku przedmiotowego braku odpowiedzi. Wybór cech i wielkości obszarów wyrównania oraz klas ważenia jest bardzo ważny, ponieważ wariancja i obciążenie oceny, wynikające z próby, zależą od jednorodności cech respondentów i nierespondentów wewnątrz obszarów wyrównania i klas ważenia. Obszary wyrównania i klasy ważenia określa się również jako komórki lub jednostki ważenia.

„Ważenie" w komórkach korekcyjnych

Jedną z metod imputacji jest „metoda ważenia" stosowana w praktyce, dla uzupełnienia braku odpowiedzi. W metodzie tej wagi dla próby lub odwrotności prawdopodobieństwa wylosowania do próby są podwyższane przez odwrotność wskaźnika odpowiedzi w komórce. Wobec tego, przypisana brakującym danym wartość na kwestionariuszu każdego nierespondenta jest średnią ze wszystkich wartości udzielonych odpowiedzi w tej komórce, z pewnymi korektami prawdopodobieństw wyboru.

Jeżeli komórka „b" zawiera n_b jednostek w próbie i m_{by} z nich odpowiedziało na jakieś pytanie lub pytania, które określają wartość badanej cechy y, wówczas estymator Horvitza-Thomsona[2]) wartości globalnej cechy y w tej komórce byłby przedstawiony przez (opuszczając y w m_{by}):

$$\hat{Y}_{bl} = \frac{n_b}{m_b} \sum_{i=1}^{n_b} \delta_{iy} Y_i / \pi_i = \sum_{i=1}^{n_b} \delta_{iy} Y_i / (\pi_i \hat{\alpha}_{iy}) \qquad (1)$$

gdzie: δ_{iy} — 1 lub 0, odpowiednio, gdy jednostka i-ta udziela lub nie udziela informacji o wartości cechy y,

[2]) Ograniczając się tylko do tego estymatora.

$\hat{\alpha}_{iy}$ — m_b/n_b ocena prawdopodobieństwa odpowiedzi, przedstawiona jako $E\delta_{iy}$, określona przez α_{iy},

Y_i — odpowiedź jednostki i-tej, jak określono wcześniej.

Obszar, wewnątrz którego przeprowadza się korekty składa się ze wzajemnie wykluczających i wyczerpujących się komórek, aby całkowite szacunki wartości globalnej cechy y były sumą szacunków dla wszystkich komórek, tj.

$$\hat{Y}_1 = \sum_b \hat{Y}_{b1} \tag{2}$$

Jak wynika z (1), $\hat{Y}_{b1}$ może być uznany za estymator doważony, gdzie waga dla jednostki i-tej zawiera odwrotność prawdopodobieństwa wyboru do próby π_i^{-1} oraz ocena odwrotności prawdopodobieństwa odpowiedzi przyjęta jest jako $\hat{\alpha}_{iy}^{-1}$ lub $(m_b/n_b)^{-1}$. Ponieważ w momencie przetwarzania danych rzadko znamy prawdopodobieństwo indywidualnej odpowiedzi, musimy zastosować najlepszy szacunek wszystkich dostępnych prawdopodobieństw odpowiedzi i po dokonaniu wyboru właściwych obszarów wyrównania i klas ważenia, szacunek ten stanowi zwykle wskaźnik odpowiedzi w danej komórce lub oszacowaną przeciętną prawdopodobieństw odpowiedzi tych jednostek.

Metoda podwajania (Duplication Method)

Imputacyjna „metoda podwajania" jest metodą, w której niedobór próby w jakiejś komórce, spowodowany brakiem odpowiedzi, jest uzupełniony przez podwojenie (reprodukcję) wszystkich lub podpróby respondentów.

W metodzie podwojenia, jeżeli rozważymy jakąś klasę ważenia b z wybranymi n_b jednostkami do próby i m_b respondentami w odniesieniu do cechy y, ocena dla komórki „b" będzie przedstawiona następująco:

$$\hat{Y}_{b2} = \sum_{i=1}^{n_b} \delta_{iy} W_{iy} Y_i / \pi_i \tag{3}$$

gdzie: W_{iy} — liczba razy jaką jednostkę i-tą reprodukowano w celu dodania $(n_b - m_b)$ nierespondentów w komórce „b" tak, aby $\Sigma \delta_{iy} W_{iy} = n_b$.

Zauważmy, że W_{iy} jest określona tylko dla respondentów, to znaczy, kiedy $\delta_{iy} = 1$.

Istnieje kilka sposobów na podwajanie jednostek. Jednym z nich, a zarazem najprostszym z punktu widzenia rozwoju metodologicznego, jest losowy wybór bez zwracania jednostek do podwojenia.

Metoda dynamiczna (Hot Deck)

Metody dynamiczne są powszechnymi metodami, korygującymi zbiory danych, zawierające wartości brakujące. Metoda dynamiczna jest to proces podwajania. Odrzucony zapis jest zastąpiony przez zapis uprzednio

zaakceptowany w tej samej komórce. Typowy zbiór jest stale uaktualniany poprzez zastępowanie błędnych zapisów zapisami nowymi i możliwymi do zaakceptowania. Zabezpiecza to w pewien sposób charakterystykę rozkładową zapisów i jednocześnie wykorzystuje bieżące dane.

Jako metoda imputacji — metody dynamiczne mają dodatkowe atrakcyjne cechy, włącznie z następującymi: a) metody te umożliwiają stosunkowo prosty sposób tworzenia warstw po losowaniu (patrz I. P. Fellegi i Holt), b) dopasowywanie zapisów nie stanowi poważnego problemu, c) w celu dokonania szacunków indywidualnych wartości dla brakujących pozycji nie potrzeba opracowywać ostrych założeń.

Oceniając metody dynamiczne warto wiedzieć, jaki wpływ na obciążenie i dokładność głównych estymatorów mają: wielkość grup klasyfikacji (często określane klasami ważenia), częstotliwość brakujących danych, wybór dopasowywanych pozycji itd. Niektóre prace teoretyczne, omawiające metody dynamiczne zostały opracowane przez I. P. Fellegi'ego i Holta, Bailara i Bailara (1978) oraz B. Coxa i R. E. Folsoma (1978).

Przyszłe prace teoretyczne będą polegały na próbie uogólnienia metod dynamicznych, które zostały rozwinięte oraz na wyprowadzeniu wzorów dla obciążenia i wariancji ocen, opartych na tych metodach. Obciążenie wynika z odchylenia oszacowanego prawdopodobieństwa odpowiedzi dla danej jednostki od rzeczywistego prawdopodobieństwa odpowiedzi (którego oczywiście nie znamy) i współzależności pomiędzy wielkością cechy i prawdopodobieństwem odpowiedzi. Wariancja zawiera dodatkowe elementy poza tymi, które występują w przypadku stosowania ważenia, w klasie ważenia. W zależności od metod i ograniczeń nałożonych na metodę dynamiczną, określenie składników dodatkowej wariancji może okazać się bardzo złożone.

W każdym razie, rozszerzenie teorii właściwej dla wariancji, dotąd rozwiniętej przez Bailara i Bailara wydaje się najbardziej obiecującą drogą do osiągnięcia celu.

Metoda substytucji danych historycznych

Metoda substytucji danych historycznych jest metodą, w której historyczne lub zewnętrzne źródła danych, takie jak: spis, wcześniejsze badania lub dane administracyjne, są podstawiane w celu uzupełnienia brakujących danych spowodowanych nie udzieleniem odpowiedzi przez jednostkę. Należy rozważyć dwa następujące rodzaje metody substytucji danych historycznych:

1) pierwszy, gdzie historyczne lub pochodzące z zewnątrz źródła dostępne są wszystkim jednostkom, które nie odpowiedziały;
2) drugi, gdzie źródła pochodzące z zewnątrz dostępne są niektórym, ale nie wszystkim jednostkom i trzeba zastosować dla tych ostatnich inną metodę imputacji, np. „ważenie".

Kiedy dane, pochodzące ze źródeł zewnętrznych, dostępne są wszystkim jednostkom, wówczas wartość przypisana brakującym danym jest przedstawiona za pomocą Z_{iy}^1, która może być obserwowaną wartością badanej cechy y^1 w poprzednim badaniu, spisie lub z akt administracyjnych. Wówczas estymator jest przedstawiany, jako:

$$\hat{Y}_3 = \sum_{i=1}^{n} [\delta_{iy} y_i / \pi_i + (1 - \delta_i) Z_{iy}^1 / \pi_i] \tag{4}$$

Należy zauważyć, że zarówno y_i, jak i Z_{iy}^1 podlegają błędowi odpowiedzi (zawierającemu systematyczny błąd odpowiedzi i wariancję odpowiedzi), uzależnionemu od prawdziwej wartości cechy dla jednostki *i*. Jeżeli historyczne lub zewnętrzne źródła danych dostępne są każdej jednostce, która nie udziela odpowiedzi odnośnie cechy „*y*", wówczas komórki korekcyjne, używane w metodach ważenia i podwajania, są zbyteczne.

W przypadku, kiedy dane pochodzące z zewnętrznych źródeł są dostępne tylko niektórym nierespondentom, wówczas trzeba by zastosować inną metodę na przykład, jak ważenie wraz z substytucją danych historycznych.

Metoda substytucji zerowej

Imputacyjna metoda substytucji zerowej jest metodą, w której brakujące dane wynikające z nieudzielenia przez jednostkę odpowiedzi, są pomijane lub ich brakujące wartości są uzupełnione za pomocą metody substytucji zerowej.

W niektórych przypadkach, komórki brakujące są po prostu oznaczone jako „nie wiadomo" i wówczas analityk decyduje o tym, która metoda analizy danych najbardziej odpowiada jego celom.

OSZACOWANIE PRAWDOPODOBIEŃSTWA ODPOWIEDZI

W każdej metodzie imputacji obciążenie w ocenie (poza systematycznym błędem odpowiedzi) wynika z oszacowanych prawdopodobieństw odpowiedzi, różniących się od prawdziwych prawdopodobieństw odpowiedzi i korelacji pomiędzy wartościami cechy badanej i prawdziwymi prawdopodobieństwami odpowiedzi. Przez określenie komórek dla celów imputacji, na przykład warstw lub zespołów, próbuje się zastosować oszacowane prawdopodobieństwo odpowiedzi, jako najbliższe prawdziwym wartościom. Najbardziej powszechną procedurą jest zastosowanie wskaźników odpowiedzi, które są równoważne szacunkom przeciętnych prawdopodobieństw odpowiedzi w komórkach korekcyjnych. Oszacowane prawdopodobieństwa odpowiedzi przedstawione są za pomocą metody imputacji, w załączonej tablicy.

Oszacowane prawdopodobieństwa odpowiedzi $\hat{\alpha}_{iy}$ według metody imputacji

Metoda	Estymator numer wzoru (w nawiasie)		Oszacowane prawdopodobieństwo odpowiedzi
Ważenie	$\hat{Y}_{b1}$	(1)	m_b / n_b
Podwajanie	$\hat{Y}_{b2}$	(3)	$[n_b/m_b]^{-1}$ lub $[n_b/m_b+1]^{-1}$
Substytucja danych historycznych			
Przypadek (*i*)	$\hat{Y}_3$	(4)	(zawiera prawdopodobieństwo dostępnych danych historycznych)

We wszystkich przypadkach z wyjątkiem $\hat{Y}_3$ oszacowane prawdopodobieństwa odpowiedzi uzyskuje się przez wskaźnik odpowiedzi w komórce „b".

ZASTOSOWANIE METODY IMPUTACJI W BADANIU SIŁY ROBOCZEJ W KANADZIE

Metody imputacji, powszechnie stosowane w badaniach reprezentacyjnych i spisach w Kanadzie, zawierają korektę wag, podwajanie, substytucję danych historycznych lub pochodzących z zewnątrz oraz metody dynamiczne.

Jednym z głównych stałych badań, przeprowadzanych przez Kanadyjskie Biuro Statystyczne, jest Badanie Siły Roboczej, za pomocą którego otrzymuje się miesięczne szacunki, dotyczące bezrobocia, zatrudnienia i wielu innych. Dane te są publikowane w kilka tygodni po zbadaniu około 56 tysięcy gospodarstw domowych (około trzeciego tygodnia każdego miesiąca). Ścisłe zasady zbierania i przetwarzania danych uniemożliwiają skontaktowanie się ze wszystkimi gospodarstwami domowymi, z którymi powinniśmy się skontaktować. Niektóre gospodarstwa domowe są niedostępne przez cały tydzień, mieszkańcy są nieobecni za każdym razem kiedy przeprowadzająca wywiad osoba przyjdzie z wizytą lub też odmawiają udzielania odpowiedzi. Istnieją oczywiście i inne przyczyny braku odpowiedzi ale stanowią one niewielki procent wszystkich braków odpowiedzi (w większości wypadków utrzymują się one w granicach pięciu procent za wyjątkiem wzrostu do siedmiu lub ośmiu procent w miesiącach letnich, co spowodowane jest „czasową nieobecnością" wynikłą z wyjazdu na urlop).

Imputacja wszystkich gospodarstw domowych, które nie udzieliły odpowiedzi przeprowadzana jest według następującego kryterium: (*i*) dla około jednej trzeciej nierespondentów, gdzie jest to możliwe do zastosowania, przeprowadzana jest substytucja wartości z poprzedniego miesiąca (wraz z odpowiednimi transformacjami niektórych danych, dla zaktualizowania danych z ostatniego miesiąca), lub (*ii*) podwyższenie poprzez odwrotność wskaźnika odpowiedzi w jednostkach wyrównania („odpowiedź" w tym

przypadku zawiera podstawione wartości dla tych nierespondentów, którzy w poprzednim badaniu udzielili poprawnych odpowiedzi). W przypadku (*ii*) imputacja pozostałych dwóch trzecich nierespondentów jest średnią dla wszystkich respondentów w jednostce wyrównania.

Imputacja przedmiotowego braku odpowiedzi lub skreśleń w czasie redagowania przeprowadzona jest jednym z trzech sposobów w zależności od cechy lub cech z brakującymi lub niepoprawnymi danymi i od stanu odpowiedzi i cech jednostki w poprzednim badaniu.

(I) na podstawie pozostałych odpowiedzi na kwestionariuszu można wywnioskować właściwą odpowiedź w danej pozycji, jeżeli nie została ona udzielona (tablica decyzyjna zapewnia, że odpowiedź będzie jednoznaczna i logiczna);

(II) sybstytucja odpowiedzi danej pozycji poprzez wykorzystanie danych z poprzedniego badania, jeżeli są one dostępne i właściwe według tablicy decyzyjnej;

(III) zastosowanie metody dynamicznej, za pomocą której można otrzymać podobny zapis w tej samej jednostce losowania pierwszego stopnia, przy tej samej ścieżce (jednej z sześciu możliwych) w szeregu pytań i w tej samej grupie wiekowej o tej samej płci. W tym przypadku, aby znaleźć podobny rodzaj zapisu, koniecznym może być łączenie pewnych klas ważenia; dla celów imputacji, jeżeli jest to konieczne, grupuje się zwykle kategorie wieku i płci, a nie jednostki losowania pierwszego stopnia czy ścieżki.

Aby zilustrować procedurę imputacji dla częściowo wypełnionych kwestionariuszy, rozważmy następujące przypadki w „Badaniu Siły Roboczej" w Kanadzie.

Przykład 1

Każdej osobie w wieku 15 lat i powyżej, mieszkającej w gospodarstwie domowym wybranym dla celów tego badania, zadaje się pytanie 80 (Q80): „czy chodzi do szkoły czy nie?". Jeżeli chodzi odpowiedź oznacza się „1", jeżeli nie, odpowiedź oznacza się „2". Osobom nie uczęszczającym do szkoły nie zadaje się żadnych dodatkowych pytań. Jednakże tym, którzy uczęszczają zadaje się następujące pytania: pytanie 81 (Q81): „czy uczęszcza do szkoły o pełnym lub częściowym wymiarze godzin?". Jeżeli odpowiedź brzmi „tak", wówczas zadajemy pytanie 82 (Q82): „do jakiego typu szkoły uczęszcza"? Wówczas możemy mieć do czynienia z następującą sytuacją: $Q80 \neq 1$ lub 2 co jest błędne.

Zależność pomiędzy pytaniami musi być następującego rodzaju:

a) jeżeli $Q80 = 1$, to znaczy uczęszcza do szkoły, wówczas $Q81 = 1$ lub 2, to znaczy szkoła o pełnym lub częściowym wymiarze godzin i $Q82 = 1$ lub 2, lub 3, lub 4, to znaczy typ szkoły;

b) jeżeli $Q80 = 2$, to znaczy nie uczęszcza do szkoły, wówczas Q81 — puste miejsce (nie stosuje się) i Q82 — puste miejsce (nie stosuje się).

Istnieją również zapisy w odpowiedziach na inne pytania, które mogą odnosić się do Q80 i Q82. Są to Q14 i Q36: „uczęszczanie do szkoły", jako powód do pracowania mniej niż 30 godzin tygodniowo, Q58 „uczęszczanie do szkoły", jako odpowiedź na to co respondent robił tuż przed poszukiwaniem pracy i „uczęszczanie do szkoły" w pytaniu Q64, jako przyczyna niemożności podjęcia pracy w ostatnim tygodniu. W każdym przypadku odpowiedź oznaczona jest „3".

Jako przykład szczególnych warunków[3]) przypuśćmy, że nie uzyskaliśmy odpowiedzi na pytanie Q80 lub odpowiedź ta nie jest zgodna, z logicznego punktu widzenia, z odpowiedziami na dalsze pytania. Poniższa tablica odpowiedzi ilustruje sposób działania dla przypisania odpowiedniej wartości pytaniu Q80.

Wyszczególnienie	Reguła imputacji						
	1	2	3	4	5	6	7
Q81 = 1 lub 2	Y	N	N	N	N	N	N
Q82 = 1, 2, 3 lub 4	Y	N	N	N	N	N	N
Q14, 36, 58 lub 64 = 3		Y	N	N	N	N	N
Wiek = 15 lub 16 lat			Y	Y	N	N	N
Miesiąc badania — lipiec lub sierpień[a])			Y	N	N	N	N
Q80 w poprzednim mies. = 1 . .					Y	N	N
= 2 . .						Y	N
≠ 1 lub 2							Y
Q80 w bieżącym miesiącu . . .	1	1	2	1	1	2	Hot Deck

[a]) Dla osób nie będących w wieku 15 czy 16 lat nie ma to żadnego znaczenia czy miesiącem badania w celu zastosowania reguł imputacji nr 5, 6 i 7 będzie lipiec czy sierpień.

W tym przypadku Hot Deck (metoda dynamiczna) oznacza poszukiwanie wewnątrz tej samej jednostki losowania pierwszego stopnia, ścieżki, grupy wieku dla danej płci, ustalonych wcześniej.

Każda kolumna przedstawia oddzielny porządek działania, który musi być podjęty w celu przypisania właściwej wartości, określającej czy dana osoba uczęszcza do szkoły czy nie (1 lub 2). Y = Tak i N = Nie (na tablicy decyzyjnej korespondują z odpowiedziami).

Poniżej podany jest skrócony opis każdej metody imputacji:

1. Wartość przypisana oparta jest na wewnętrznej logice powiązanych ze sobą pytań. Dlatego, jeżeli na pytania Q81 i Q82 otrzymaliśmy ważne odpowiedzi, wówczas pytaniu Q80 przypiszemy wartość, która będzie

[3]) Eliminuje to inne uwarunkowania, na przykład, jeżeli Q81 nie równa się 1 lub 2, a Q82 równa się 1, 2, 3 lub 4 i odwrotnie.

zgodna z inną informacją, to znczy: jeżeli Q81 = 1 lub 2 i Q82 = 1, 2, 3 lub 4, wówczas pytaniu Q80 przypiszemy „1", tzn. uczęszcza do szkoły.

2. Jeżeli nie uzyskaliśmy odpowiedzi na pytania Q81 i Q82, musimy wówczas poszukać odpowiedniej informacji w innym miejscu na kwestionariuszu. Stawiamy pytania w zależności od „ścieżki", którą dana osoba mogła podążyć. Jeżeli osoba ta zaznaczyła wcześniej, że uczęszcza do szkoły, wówczas pytanie Q80 jest również zakodowane, tzn. jeżeli Q81 ≠ 1, 2 i Q82 ≠ 1, 2, 3, 4, ale Q14, 36, 58 lub 64 = 3, wówczas Q80 = 1, tzn. uczęszcza do szkoły.

3—4. Jeżeli żadna bezpośrednio związana informacja nie jest dostępna, wówczas wykorzystujemy informację, która nie wprost odnosi się do pytania o odpowiedzi uznanej, w czasie redagowania, za sprzeczną z innymi. W tym przypadku jest to wiek osoby i miesiąc, w którym badanie jest przeprowadzane. Jeżeli dana osoba jest w wieku lat 15 lub 16, wówczas najprawdopodobniej nie chodziła do szkoły w miesiącach lipcu i sierpniu i dlatego Q80 zakodowane jest jako 2, tzn. nie chodzi do szkoły. W przypadku wszystkich innych miesięcy osobę tę koduje się jako „chodzi do szkoły", tzn. Q81 = 1. Ponieważ przypadki imputacji podane na przykładach zdarzają się raczej rzadko, usprawiedliwionym jest wyznaczenie „najbardziej prawdopodobnej wartości". Nawet, gdyby metoda dynamiczna była metodą lepszą z teoretycznego punktu widzenia, metoda ta jest prawdopodobnie droższa.

5—6. Jeżeli nie mamy dostępu do żadnych danych z bieżącego miesiąca i dana osoba nie jest w wieku lat 15 lub 16, wówczas cofamy się do danych z poprzedniego miesiąca. Bez względu na odpowiedź jaka została udzielona w poprzednim miesiącu, jeżeli jest ona dostępna, zostaje przeniesiona na kwestionariusz z bieżącego miesiąca. Wykorzystanie danych z poprzedniego miesiąca jest usprawiedliwione głównie przez fakt, że współzależność pomiędzy miesięcznymi ocenami dla pewnych cech jest dość wysoka.

BIBLIOGRAFIA

1. Ashraf A. and Macredie I.: *Edit and Imputation in the Labour Force Survey*, Imputation and Editing of Faulty or Missing Survey Data, 114—119, selected papers presented at 1978 American Statistical Association meeting.
2. Bailar J.C. III and Bailar B.A. (1978): *Comparison of Two Procedures for Imputing Missing Survey Values*, Imputation and Editing of Faulty or Missing Survey Data, 65—75, selected papers presented at 1978 American Statistical Association meeting.
3. Ford Bary L. (1976): *Missing Data Procedures: A Comparative Study*, American Statistical Association, Proceedings of Social Statistics Section.
4. Hansen M.H., Hurvitz W.N. and Madow W.G. (1953): *Sample Survey Methods and Theory*, Volume II, Theory, 139—141, John Willy and Sons, Inc.
5. Joseph Naus: *Data Quality Control and Editing*, Marcel Dekker, New York.

6. Kalton G. and Kasprsy: *The Treatment of Survey Missing Data,* Survey Methodology Journal Volume 12, No. 1, 1986.
7. Madow William G., Unpublished: *Report on Nonresponse* (1980).
8. Morgan J.A. and Sonquist J.N. (1964): *Monograph 35 Survey Research Centre,* University of Michigan.
9. Platek R.: *Some Factors Affecting Nonresponse,* Survey Methodology Journal (Statistics Canada) Volume 3, Number 2 (December 1977), 191—214.
10. Platek R., Singh M.P. and Tremblay, V.: *Adjustment for Nonresponse in Surveys,* Survey Sampling and Measurement, Academic Press, New York, 1978, 157—175.
11. Platek R. and Gray G.B.: *Methodology of Adjustments for Nonresponse,* Invited Paper presented at the 42 Session of International Statistical Institute, Manila, Philippines, December 1979.
12. Platek R.: *Methodology and Treatment for Nonresponse,* INSTITUTO VASCO DE ESTATISTICA, 1986.
13. Platek R., Gray G.B.: *Definitions of Response Rates,* Survey Methodology Journal Volume 12, No. 1, 1986.
14. Statistics Canada: *Imputation for Household Survey in Statistics Canada.*

Konieczność uwzględniania błędów losowych i nielosowych w planowaniu badań reprezentacyjnych i wykorzystaniu ich wyników

doc. dr hab. Jan Kordos
Główny Urząd Statystyczny
Prezes Polskiego Towarzystwa Statystycznego

UWAGI WSTĘPNE

Wprawdzie metoda reprezentacyjna znalazła szerokie zastosowania w badaniach statystycznych, to jednak wykorzystanie jej w różnych typach badań jest dość zróżnicowane. Najczęściej stosowana jest w badaniach społecznych, a znacznie rzadziej w statystyce handlu, przemysłu, budownictwa i innych działach statystyki, które opierają się głównie na pełnej sprawozdawczości. Uznaje się ogromne potencjalne możliwości zastosowania metody reprezentacyjnej także w tych dziedzinach badań statystycznych, w których dotychczas, z różnych przyczyn, metoda ta nie znalazła szerszego zastosowania. Poznanie tych przyczyn i głębsze ich zbadanie może przyczynić się do efektywnego wykorzystania metody reprezentacyjnej w praktyce badań statystycznych.

Warto zaznaczyć, że podobne zasady teoretyczne, na których opierają się badania reprezentacyjne, wykorzystywane są w doświadczalnictwie, w statystycznej kontroli jakości, w symulacji statystycznej i różnego rodzaju badaniach naukowych. Jednakże podobieństwa te dotyczą tylko wspólnych podstaw teoretycznych, na których opierają się te dyscypliny. Metoda reprezentacyjna, stosowana w badaniach statystycznych, musi rozwiązać wiele innych zagadnień, aby mogła być efektywnie stosowana w praktycznych zastosowaniach. Chodzi tu głównie o to, że nie wystarczy uwzględnić, przy planowaniu badań statystycznych metodą reprezentacyjną, jedynie oczekiwanej precyzji ocen parametrów, ale także inne aspekty metodologiczne badań statystycznych. Dotyczy to w szczególności metod zbierania danych, narzędzi badawczych, systemu organizacyjnego, dostępności personelu w terenie oraz możliwości jego szkolenia i kontroli, dostępnych środków finansowych, a także innych aspektów, ograniczających możliwości przeprowadzenia badania. Problemy te związane są z jakością danych statystycznych, a więc ich przydatnością, terminowością i dokładnością. Uwzględnienie jedynie precyzji przy planowaniu badania nie może dać zadowalających wyników. Powstaje więc problem — w jaki sposób zagadnienia te powinny być uwzględniane w całokształcie planowania badania reprezentacyjnego? Odpowiedź jest dość złożona i poświęcimy jej nasze rozważania w dalszej części tego opracowania.

Rozwój badań reprezentacyjnych i ich potencjalne możliwości zwróciły uwagę wyższych uczelni i urzędów statystycznych na potrzebę udoskonalenia metod szkolenia statystyków, zajmujących się badaniami reprezentacyjnymi. Okazało się, że w wielu częściach świata, a także i w Polsce, nauczanie i szkolenie statystyków w tym zakresie nie jest zadowalające. Tego stanu rzeczy nie można wyjaśnić niedostatkiem literatury z tego zakresu, która może być wykorzystywana na wykładach i w czasie szkolenia zawodowego. W rzeczywistości braki takie nie występują. Istnieje przecież wiele doskonałych podręczników na temat metody reprezentacyjnej, głównie w języku angielskim [zob.: 2, 7, 8, 12]. Także periodyki statystyczne poświęcają dość dużo miejsca problematyce badań reprezentacyjnych. W języku polskim są dostępne trzy podręczniki z tego zakresu [zob.: R. Zasępa 18, 19] i Z. Pawłowski [13] oraz dwie monografie wydane przez Główny Urząd Statystyczny [1, 16]. Wydał także specjalne opracowanie, poświęcone badaniom reprezentacyjnym, Zakład Badań Statystyczno-Ekonomicznych GUS i PAN [11, 20]. Powstaje więc pytanie — dlaczego metoda reprezentacyjna napotyka na znaczne trudności szerszego jej stosowania w niektórych typach badań statystycznych. Wydaje się, że na ten stan rzeczy złożyło się wiele różnych przyczyn. Warto już na wstępie niektóre z nich zasygnalizować.

Pragnę zaznaczyć, że podane tu uwagi nie dotyczą osób, zajmujących się metodologią badań reprezentacyjnych i prowadzących w tym kierunku prace naukowo-badawcze. Odnoszą się one do problemów metodologicznych

badań statystycznych oraz instytucji, prowadzących takie badania. Pomimo, że poziom prac teoretycznych w tym zakresie może być wysoki, to jednak ich jednostronność nie może dać efektywnych wyników w praktycznych zastosowaniach.

Jakie więc przyczyny powodują niedostateczne wykorzystanie metody reprezentacyjnej, w niektórych typach badań statystycznych?

Wydaje się, że niewielkie grono statystyków i użytkowników danych statystycznych rozumie rzeczywiste zalety metody reprezentacyjnej i istotę jej zastosowania w praktyce. Wtedy, stosowane dotychczas metody wydają się bezpieczniejsze i mniej kłopotliwe, a szczególnie gdy nie badano ich efektywności z punktu widzenia kosztu badania i rzeczywistej dokładności wyników. Wydaje się wtedy, że metoda reprezentacyjna może dać wyniki niedokładne, których wartość może być niewielka. Przy niedostatecznej popularyzacji statystyki, a metody reprezentacyjnej w szczególności, takie obawy są w pełni uzasadnione. Przecież rzeczywiście, przy stosowaniu metody reprezentacyjnej w początkowej fazie badania, tj. w jego przygotowaniu, niezbędne są o wiele większe nakłady pracy niż przy metodach tradycyjnych. Wymagana jest dość wysoka wiedza i doświadczenie w tym zakresie, a przecież w wielu jednostkach organizacyjnych GUS brak jest specjalistów z tego zakresu. Zmiana tego stanu rzeczy wymaga albo zmiany organizacji jednostek branżowych statystyki, albo stworzenia silnego zaplecza naukowo-badawczego, które by te prace systematycznie prowadziło. Nie rozwiąże tych zagadnień zespół trzy lub czteroosobowy nawet o najwyższym przygotowaniu zawodowym. Potrzebne jest szersze zaplecze naukowo-badawcze, prowadzące w tym zakresie eksperymenty i analizy, które by doprowadziły do racjonalnych zastosowań metody reprezentacyjnej.

Pewne wyjaśnienie można również znaleźć w tym, że czytelnicy literatury o metodzie reprezentacyjnej mogą natrafić na trudności w zrozumieniu prezentowanych zagadnień. Tak więc poszczególni autorzy dążą do skupienia uwagi na różnych problemach teoretycznych i niektórych dziedzinach zastosowań, a ponadto używają różnej terminologii i oznaczeń, co w znacznym stopniu utrudnia czytanie i zrozumienie literatury przedmiotu.

Przyczyna tego stanu rzeczy może być również spowodowana nauczaniem tego przedmiotu na uczelniach i w urzędach statystycznych. Jak podaje T. Dalenius [4], w wielu krajach brakuje statystyków, posiadających głębokie przygotowanie teoretyczne i praktyczne z metodologii badań statystycznych, a w tym z metody reprezentacyjnej, którzy są w stanie poświęcić swój czas nauczaniu i szkoleniu. Niejednokrotnie nauczanie przedmiotu o metodzie reprezentacyjnej powierza się na wyższych uczelniach (dotyczy to innych krajów) pracownikom naukowym, którzy nie są do tego dostatecznie przygotowani, a stąd nie są w stanie właściwie nauczyć tego przedmiotu. Dlatego pracownicy różnych instytucji lub statystycy, którzy chcą dogłębnie

poznać metodę reprezentacyjną skazani są głównie na własną inicjatywę w tym zakresie. Wyraźnie ten fakt podkreśla wspomniany już T. Dalenius [4] w wydanym w 1986 r. przewodniku po problematyce badań reprezentacyjnych, który umożliwia studiowanie tych zagadnień przez szersze grono statystyków. Wydaje się celowe przetłumaczenie tego opracowania na język polski z odpowiednimi uzupełnieniami niezbędnymi dla naszej specyfiki.

Ogólnie uznaje się, że wyniki badań reprezentacyjnych powinny być inaczej traktowane niż wyniki badań pełnych. Zwykle podkreśla się występowanie błędów losowych, które mogą w istotny sposób zniekształcić uzyskane wyniki. Jednak jest to tylko jedna strona medalu. Wyniki badań reprezentacyjnych mogą być również obciążone, tak jak każde badanie statystyczne, błędami innego rodzaju, a mianowicie błędami nielosowymi. Błędy te powinny być także uwzględniane przy prezentacji i analizie wyników badań reprezentacyjnych, tak jak i badań pełnych i częściowych, opartych na innych zasadach doboru próbki niż wybór losowy. Problem ten będzie również przedmiotem naszych rozważań.

TEORIA I PRAKTYKA BADAŃ REPREZENTACYJNYCH

Powszechnie uważa się, że badania reprezentacyjne opierają się na mocnych podstawach teoretycznych. Trudno wyrazić odmienną opinię, jeśli uważa się, że dla badań reprezentacyjnych najważnieszy jest problem doboru próbki, która umożliwia wnioskowanie o całej zbiorowości generalnej, przy określonych założeniach.

Istnieje dość obszerna literatura na ten temat, która w zasadzie jest znana większości badaczom tego przedmiotu. Dotyczy ona abstrakcyjnych rozważań matematycznych, opartych głównie na rachunku prawdopodobieństwa i statystyce matematycznej, a także w pewnym stopniu na prakseologii i różnych metodach badawczych, przedstawianych w podręcznikach metody reprezentacyjnej.

Teoria badań reprezentacyjnych jest jednak w niewielkim stopniu wykorzystywana w praktycznych zastosowaniach. Wykorzystywane są jedynie podstawowe zasady losowania próbki i estymacji parametrów, bez uwzględniania głębszych rozważań teoretycznych. Optymalizacja badań reprezentacyjnych ogranicza się głównie do teorii. Podstawowy problem można by krótko sformułować następująco: znalezienie takiego schematu losowania, który umożliwi zminimalizowanie błędu przy danym koszcie badania (lub zminimalizowanie kosztu badania przy żądanej dokładności ocen parametrów), pod warunkiem różnych ograniczeń natury fizycznej, psychicznej i organizacyjnej w określonej sytuacji. Teoria badań reprezentacyjnych dostarcza przede wszystkim podstaw dla zdefiniowania problemu, a następnie umożliwia jego rozwiązanie zakładając, że mogą być określone odpowiednie parametry. Jest to ogólny model teoretyczny badania

reprezentacyjnego, który ma niewiele wspólnego z problemami, występującymi w większości badań reprezentacyjnych.

Teoretyczne sformułowanie problemu rozpoczyna się od próby sprecyzowania celu badania, określenia zbiorowości generalnej oraz dokładniejszego zdefiniowania „błędu", który powinien być minimalizowany. Zwykle wyniki badania składają się z wielu różnych zmiennych i musimy najpierw zdecydować, która z nich lub ich kombinacja ma ten błąd minimalizować. Musimy następnie odnieść ten błąd do schematu losowania wyrażonego funkcją matematyczną. Zwykle traktuje się tylko składniki błędów losowych, a składniki błędów nielosowych są rzadko uwzględniane (por.: [5 i 6]). Następnie wyraża się koszt badania jako funkcję schematu losowania. Wtedy dopiero można przystąpić do obliczeń optymalnych rozwiązań dla danego schematu, a te optymalne rozwiązania mogą być porównane z różnymi możliwymi schematami, co może teoretycznie prowadzić do absolutnego optimum. W końcu sprawdzamy je z dodatkowymi „ograniczeniami" i jeśli potrzeba, dokonujemy korekt do obliczonego optimum, w celu jego spełnienia. Jest to oczywiście uproszczona idea postępowania.

Procedura ta wymaga:

— jasnego sprecyzowania głównego celu badania,
— pewnych szczegółowych informacji o parametrach, oddziałujących na organizację badania: wariancję między warstwami oraz wewnątrz warstw i zespołów, błędy nielosowe, funkcję kosztów.

W praktyce dość rzadko spotykamy się z sytuacją, w której to wszystko jest spełnione. Określenie głównego celu badania w badaniach wieloprzedmiotowych i wielocelowych jest bardzo trudne. Chodzi głównie o to, aby do tego celu znaleźć pojedynczą zmienną lub kombinację zmiennych. Główny cel badania można jeszcze w pewien sposób określić w niektórych badaniach demograficznych (np. dla wskaźnika urodzeń lub wskaźnika przyrostu naturalnego na poziomie krajowym), ale nie jest łatwo w innych badaniach. Jeśli nawet można określić jedną najważniejszą zmienną, to mogą wystąpić trudności, przy określaniu względnej wagi, którą należy jej przypisać na poziomie krajowym, następnie dla regionów kraju lub innych podziałów. Znajomość parametrów błędów i kosztów może być dostępna tylko wtedy, gdy podobne badanie zostało przeprowadzone w tym samym lub podobnym kraju, a nawet gdy otrzymano parametry tylko dla pewnego schematu, który był poprzednio wykorzystany (np. otrzymano wariancje wewnątrz warstw dla poszczególnych warstw, które zostały wykorzystane, lecz może nie były one optymalne). Nie tylko trudno jest otrzymać dane dla optymalizacji schematu losowania, lecz również ograniczona jest praktyczna ważność takich optymalizacji. Po pierwsze, badania wykazały, że efektywność schematu losowania jest raczej niewrażliwa na dokładne wartości parametrów blisko optimum; można powiedzieć, że odchylenia od optymalnych wartości mogą być dość duże bez zbytniego oddziaływania na efektywność.

Zachęca to do kompromisowych rozwiązań przy planowaniu badania i uczynienia ich bardziej elastycznymi. Po drugie występują „ograniczenia", o których już wspomniano; różne opcje są ograniczone przez takie czynniki, jak: dostępność operatów losowania, dostępność personelu przeprowadzającego badania, możliwość kontroli badania w czasie jego realizacji itd. Niekiedy zdarzają się poważne ograniczenia planu badania, że niewiele pozostaje z pierwotnej optymalizacji. Tak więc w praktyce okazuje się, że większość badań reprezentacyjnych planuje się z niewielkim lub przy prawie żadnym uwzględnieniu metodologii optymalizacji, która została rozwinięta przez teoretyków badań reprezentacyjnych. Chociaż niekiedy wykonuje się obliczenia zgodnie z zasadami opisanymi powyżej, to jednak końcowe decyzje są zwykle oparte na „wyczucie" sytuacji, doświadczeniach z poprzednich badań oraz różnych ograniczeniach organizacyjnych, które limitują przyjęte rozwiązania. Tak więc w praktyce badania reprezentacyjne są tylko częściowo oparte na podstawach naukowych, a częściowo na innych zasadach, które można by zaliczyć raczej do sztuki niż nauki. W różnych opracowaniach o metodzie reprezentacyjnej można się spotkać z pytaniem, w jakim stopniu badania te zależą od nauki, a w jakim od sztuki? Nie ulega wątpliwości, że sztuka odgrywa tu ważną rolę.

Nie można nikogo winić za ten stan rzeczy, gdy chodzi o badania jednorazowe. Jednakże, gdy badania są powtarzalne, wtedy istnieją większe możliwości zastosowania teorii badań reprezentacyjnych, która powinna prowadzić do doskonalenia ogólnej metodologii badania. Niewiele podejmowano dotychczas tego rodzaju badań, a jeszcze mniej publikowano wyników, w których bada się wariancje, elementy kosztu badania oraz problemy organizacyjne, niezbędne do zwiększenia efektywności przyszłych badań.

Tak więc, chociaż gromadzi się doświadczenia, to jednak nie są one przekazywane, w wystarczającym stopniu, innym badaczom. Jest to sytuacja niepokojąca ze względu na efektywność przyszłych badań reprezentacyjnych, które w poważnym stopniu zależą od dotychczasowych doświadczeń zarówno w zakresie teorii, jak i praktyki.

Należałoby więc dążyć do tego, aby koncentrować uwagę nie tylko na bieżących problemach, ale także na zagadnieniach związanych z przyszłością. Chodzi tu także o bardziej efektywne powiązanie teorii z praktyką.

BŁĘDY LOSOWE A NIELOSOWE

Wiadomo, że błędy w badaniach reprezentacyjnych powstają z wielu różnych przyczyn, a błąd losowy jest tylko jednym ze składników błędu ogólnego. Jak wiadomo, ogólnym miernikiem błędów, tj. błędów losowych i nielosowych, jest **średni błąd kwadratowy**, który można przedstawić w następującej ogólnej postaci [14]:

$$MSE = WO + B^2$$

gdzie: *WO* oznacza warjancję ogólną, która zawiera wariancję losową (*WL*), prostą wariancję odpowiedzi (*PWO*), skorelowaną wariancję odpowiedzi (*SWO*) oraz kowariancję odpowiednich czynników (*KCZ*), tj.

$$WO = WL + PWO + SWO + KCZ,$$

przy czym *B* — obciążenie, które z kolei składa się z obciążenia losowania (*BL*), obciążenia odpowiedzi (*BO*), obciążenia ze względu na brak odpowiedzi (*BN*), tj. $B = BL + BO + BN$.

Już z tego prostego schematu ogólnego miernika błędu uwidacznia się, że różne czynniki oddziaływują na dokładność danych, a czynniki te powinny być uwzględniane przy planowaniu badania, gdy zakłada się racjonalne jego wykorzystanie. Uwzględnienie tylko błędu losowego może prowadzić do niewłaściwych rozwiązań. Niejednokrotnie w wielu badaniach rozmiary błędów nielosowych mają o wiele większe znaczenie niż błędy losowe. Jednakże faktu tego nie uwzględnia się przy planowaniu badań, a także przy wykorzystaniu ich wyników. Rozważmy problematykę błędów losowych i nielosowych nieco szerzej, w aspekcie planowania badań reprezentacyjnych i wykorzystywania ich wyników.

Jak wiadomo, celem badania reprezentacyjnego jest dokonanie określonych ocen i wnioskowanie, dotyczące zbiorowości generalnej na podstawie obserwacji uzyskanej z próbki wybranej, w odpowiedni sposób, z tej zbiorowości. W czasie tego badania możemy popełnić błędy różnego rodzaju. **„Błąd"** możemy określić jako **różnicę między faktyczną wartością z populacji** (zwykle nieznaną), a **wartością uzyskaną z obserwowanej próbki.** Rozróżniamy tu dwie szerokie kategorie błędu (zob. J. Kordos [9]). **Po pierwsze,** błędy powstają z faktu, że to co obserwowano odbiega od tego co zamierzano zmierzyć w badaniu. Takie **błędy pomiaru** koncentrują się na podstawowej treści badania; określenie celów badania, przetransportowanie tych celów w odpowiedni zestaw pytań, interpretacja tych pytań i przekazanie ich w odpowiedni sposób respondentom, zdolność respondentów i chęć do podania żądanych informacji, jakość zapisu na formularzu, jakość redagowania i kodowania odpowiedzi, itd. **Po drugie,** błędy powstają z **procesu uogólnienia wyników z próbki na całą badaną populację.** Błędy te zależą od sposobu doboru próbki, jej liczebności i metody estymacji. Wystąpią one nawet wtedy, jeśli nie popełniono błędów pomiaru badanych jednostek. Ważne jest, aby określić wyraźnie tę drugą grupę błędów, gdyż błąd losowy, który jest jej składnikiem, jest niekiedy mylony z tą grupą, jako całością. Aby wyciągnąć wnioski z próbki o całej populacji badanej w sposób obiektywny, wybór tej próbki powinien być oparty na pewnym schemacie probabilistycznym. Nawet, gdy wybór próbki jest oparty na takim schemacie, to w praktyce mogą być naruszone niektóre warunki tego wyboru, w czasie procesu realizacji próbki. Próbka taka, jak wiadomo, losowana jest z odpowiedniego operatu losowania w określony sposób. Jednakże może się zdarzyć, że nie wszystkie jednostki badanej populacji

znajdują się w wykorzystywanym operacie losowania, a niektóre z nich mogą wystąpić podwójnie. Mogą być również niedostateczne informacje do zidentyfikowania jednostek w terenie. Nie ze wszystkich jednostek, wybranych do próbki, można uzyskać żądane informacje. W wyniku oddziaływania takich czynników, struktura próbki może być poważnie naruszona, gdyż prawdopodobieństwa wyboru jednostek do próbki są różne od tych, jakich wymagał schemat losowania. Te błędy w realizacji próbki, jak również świadome odstępstwa od losowania próbki, mogą powodować obciążenia ocen otrzymanych z badania.

Błędy pomiaru (zwane również błędami odpowiedzi) i **błędy realizacji** próbki są wspólnie nazwane **błędami nielosowymi.** Z powyższego wynika, że błędy mogą powstać z różnych przyczyn i oddziaływać na wyniki badania w różny i złożony sposób. Problemy te przedstawione są obszerniej w pracach [5, 6, 14, 15 i 17], a w języku polskim w pracach [9 i 10]. W odróżnieniu od tych błędów, **błąd losowy** powstaje w procesie statystycznej estymacji parametrów populacji, na podstawie wyników uzyskanych z próbki.

Warto przypomnieć, że schemat losowania określa zasady, według których powinna być wylosowana próbka do badania oraz zasady estymacji parametrów populacji generalnej. Gdyby w badaniu nie wystąpiły nawet błędy pomiaru i błędy realizacji badania (tj. błędy nielosowe), to powtórne zastosowanie tego samego schematu losowania dałoby w wyniku różne oceny parametrów, w zależności od aktualnych jednostek badania, które wybrano do próbki. Błąd losowy estymatora jest miarą jego zmienności przy założeniu teoretycznie możliwych powtórzeń badania i braku błędów nielosowych.

Na ogół oddziaływanie jakiegoś źródła błędu (losowego lub nielosowego) na wyniki badania można rozłożyć na dwa składniki: **błąd zmienny** i **obciążenie.** W zasadzie rozróżnienie to jest oparte na założeniu możliwości powtórzenia badania, w tych samych warunkach. Niektóre z tych warunków, w których badanie jest przeprowadzane, można uważać jako **podstawowe.** Do nich można zaliczyć: ogólne warunki społeczne, cechy badanej populacji, jakość dostępnego operatu losowania, istotę i złożoność poszukiwanych informacji, metodę zbierania danych, typ personelu badania i inne podobne elementy. Oprócz warunków podstawowych, na wyniki badania oddziaływują także czynniki **przejściowe** lub **przypadkowe,** jak na przykład wybór poszczególnych jednostek do próbki, wykorzystanie w badaniu określonego personelu, warunki przeprowadzenia konkretnego wywiadu (lub pomiaru) itd. Gdy badanie będzie powtarzane w tych samych warunkach podstawowych, wtedy różne powtórzenia dadzą różne wyniki, na skutek wpływu czynników przypadkowych. **Zmienny składnik błędu** mierzy **zmienność między różnymi ocenami dokonywanymi w takich hipotetycznych powtórzeniach badania.** Jeśli weźmiemy średnią ze wszystkich możliwych powtórzeń to uzyskamy **wartość oczekiwaną** w danych warunkach pods-

tawowych. **Różnica między wartością oczekiwaną a wartością z populacji daje obciążenie.** Z pewnym uproszczeniem, można to wyrazić w ten sposób, że obciążenie jest stałym składnikiem błędu, który ma taki sam wpływ na każde powtórzenie badania.

Tak więc **błąd zmienny** mierzy zmienność między ocenami uzyskanymi **z różnych powtórzeń w tych samych podstawowych warunkach badania.** Można rozważać możliwe powtórzenia w dwóch „płaszczyznach": powtórne obserwacje dokonywane na jednostkach stałej próbki oraz powtórzenie badania na różnych próbkach. **Wariancja pomiaru** (lub wariancja odpowiedzi) jest miarą zmienności powtórnych pomiarów na jednostkach tej samej próbki. Zmienność przeciętnej lub wartości oczekiwanej tych pomiarów z różnych próbek daje **wariancję losową.**

Należy w końcu zauważyć, że **składnik losowy** „średniego błędu kwadratowego" składa się z **wariancji losowej** plus **kwadrat obciążenia estymatora.** To ostatnie obciążenie jest określone jako różnica między oczekiwaną wartością estymatora i estymowanego parametru populacji, przy założeniu braku błędów nielosowych. Obciążenie to jest czysto statystyczną właściwością zastosowanego estymatora, a przy racjonalnej liczebności próbki i odpowiednim planie losowania można go zwykle uniknąć, przez zastosowanie właściwych procedur estymacji. Powstaje pytanie — w jaki sposób uzyskać informacje o rozmiarach błędów nielosowych? Informacji i wskazówek o rozmiarach błędów nielosowych mogą dostarczyć różnego rodzaju badania próbne i kontrolne, które są niezbędne przy planowaniu badań reprezentacyjnych.

ZNACZENIE BADAŃ PRÓBNYCH PRZY PLANOWANIU BADAŃ REPREZENTACYJNYCH

Problematyka badań próbnych została szerzej przedstawiona w pracy [9, podrozdz. 1.7]. Rozważmy zatem tylko możliwość wykorzystania badań tego rodzaju w początkowej fazie badania. Badania takie przeprowadzamy zwykle dla sprawdzenia narzędzi badawczych, takich jak: kwestionariusze i instrukcje, ale i dla wielu innych celów, związanych z jakością przeprowadzonego badania. W zależności od metody zbierania danych, problematyki badawczej, personelu uczestniczącego w badaniu i jego organizacji można uzyskać dodatkowe informacje o różnego rodzaju błędach nielosowych, które mogą wystąpić w badaniu. Można oszacować oczekiwane rozmiary błędów kompletności badania w zależności od metody zbierania danych, badań próbnych oraz doświadczeń podobnych badań, przeprowadzonych w przeszłości. Tematyka badania wskazuje na możliwość popełnienia błędów odpowiedzi (pomiaru). Gdy są to problemy drażliwe lub godzące w pewien sposób w interesy osobiste respondentów, wtedy należy oczekiwać wystąpienia poważnych błędów odpowiedzi, a więc dla tych pozycji nie jest uzasadnione żądanie wysokiej precyzji ocen, przy planowaniu

badania, natomiast więcej uwagi i przeznaczonych zasobów finansowych należy poświęcić na minimalizację błędów nielosowych. Dotychczas, w niewielkim stopniu zagadnienia te były łącznie traktowane przy planowaniu badań reprezentacyjnych.

POTRZEBA INFORMOWANIA UŻYTKOWNIKA O JAKOŚCI DANYCH

Użytkownik danych statystycznych dotychczas był informowany w niewielkim stopniu o jakości prezentowanych wyników. Dotyczy to większości krajów i dlatego Konferencja Statystków Europejskich wydała zalecenia w sprawie informowania użytkownika o jakości danych statystycznych. Zalecenia te omówione są szerzej w publikacji [11, rozdz. 6]. Warto rozważyć jakiego rodzaju ostrzeżenia powinien uzyskać użytkownik danych statystycznych, aby mógł je prawidłowo wykorzystać. Użytkownik musi mieć informacje o głównych celach danych, które zostały opracowane i opublikowane, o pojęciach i definicjach oraz o ograniczeniach z nimi związanych, a także o liczebności próbki wraz z ogólnym schematem losowania. Należy poinformować użytkownika w jakich przekrojach podstawowych dane mogą być wykorzystywane. Gdy, na przykład, dane nie powinny być wykorzystane w przekroju wojewódzkim, to należy wyraźnie to zaznaczyć i podać przyczyny. Należy ostrzec przed ryzykiem jakie może wystąpić, gdy dane wykorzystane są dla innych celów niż pierwotnie zakładano.

Powstaje pytanie, czy publikowanie tylko błędów losowych, które są często jedynymi wskaźnikami jakości, nie wprowadza użytkownika w błąd, gdy losowanie wprowadziło w istocie nieznaczne błędy. Wspomnienie tylko o błędach losowych może być rzeczywiście mylące. Dlatego powinny być równolegle omówione błędy związane z nieuzyskaniem informacji, pominięciem jednostek, błędami odpowiedzi itd. Należy o tym zawsze pamiętać, gdyż dotychczas statystycy poświęcali za dużo uwagi błędom losowym jedynie z tego powodu, że mieli do dyspozycji narzędzie, które dobrze znali, tj. statystykę matematyczną.

Przy przygotowywaniu danych dla użytkowników należy wziąć pod uwagę dwie kategorie użytkowników:
— użytkowników ogólnych, którzy powinni uzyskać informacje o jakości danych w formie zwięzłej i jasnej,
— statystyków, którzy powinni znać całkowitą metodologię badania, wyniki sprawdzenia jakości danych wraz z wnioskami na przyszłość.

Wkraczamy tu w problematykę metadanych (tj. informacji o danych). Ogólnie można stwierdzić, że metadane powinny w szczególności zawierać:
— wykaz błędów i innych czynników tego rodzaju,
— definicje wskaźników (dobre lub złe wykonanie),
— elementy kosztu,

— istotę zależności,
— końcową analizę i względne priorytety różnych celów.

Dotychczasowa praktyka w tym zakresie nie jest zadowalająca i należy ją stopniowo zmienić, aby dane uzyskane z badań reprezentacyjnych były należycie wykorzystane.

UWAGI KOŃCOWE

W opracowaniu przedstawione zostały w ogólnym zarysie podstawowe problemy, występujące w planowaniu badań reprezentacyjnych, a także w analizie wyników uzyskanych z tych badań. Szczególną uwagę zwróciliśmy na potrzebę uwzględniania w planowaniu badań reprezentacyjnych zarówno błędów losowych, jak i nielosowych oraz na konieczność pewnych kompromisów, przy podejmowaniu decyzji odnośnie liczebności próbki i schematów losowania.

Dotychczas zajmowaliśmy się w Polsce w niewystarczającym stopniu metodologią badań statystycznych. Szczególnie zaniedbana jest początkowa faza badania, tj. przygotowanie i realizacja. Nie prowadzono szerszych prac badawczych nad poszukiwaniem efektywnych metod zbierania danych w różnych typach badań statystycznych. Ograniczono się głównie do sprawozdawczości statystycznej, budowy formularza i sposobów jego wypełniania. Sprawozdawczością statystyczną obejmuje się zwykle wszystkie jednostki badania, chociaż bardziej racjonalne mogłoby być zastosowanie badania częściowego, a szczególnie dla małych jednostek organizacyjnych, które są zwykle bardzo liczne w badanej zbiorowości. Nie prowadzono także szerszych badań nad dokładnością uzyskanych wyników i poszukiwaniem bardziej racjonalnych metod zbierania danych, przy danym koszcie badania. Niedostatek prac badawczych w tych zakresie spowodował stagnację w metodologii badań, a tym samym nie mógł oddziaływać na udoskonalenie jakości danych statystycznych.

W udoskonaleniu badań statystycznych mogą w istotny sposób pomóc prace z zakresu metodologii badań. Szczególną rolę może odegrać metoda reprezentacyjna, gdy będzie traktowana łącznie z całością przygotowywanego badania, a nie tylko niektórych jej elementów. Niezbędne są, jak już zaznaczono, integracyjne wysiłki zarówno sfery naukowo-badawczej, jak i urzędów statystycznych, realizujących te badania. Pomimo wysokiego poziomu prac teoretycznych z zakresu metody reprezentacyjnej nie uzyska się efektywnego jej zastosowania w praktyce, gdy pomijane są inne aspekty metodologiczne badań statystycznych. Dlatego wydaje się celowe, aby w praktyce badania reprezentacyjne były przygotowywane nie tylko przez teoretyków metody reprezentacyjnej, ale wspólnie z metodologami badań statystycznych, którzy znają problematykę zbierania danych, efektywność narzędzi badawczych stosowanych w badaniach statystycznych oraz przedmiot badań. Dopiero takie połączenie specjalności może dać efektywne

ustawienie badania i właściwe jego przeprowadzenie, opracowanie i wykorzystanie.

Należy w zakończeniu także podkreślić **potrzebę popularyzacji metody reprezentacyjnej,** gdyż dotychczasowa praktyka wskazuje na znaczne zaniedbania w tej dziedzinie. Powinna to być akcja systematyczna, którą powinny prowadzić **urzędy statystyczne, wyższe uczelnie,** a także **Polskie Towarzystwo Statystyczne.** Należy mieć nadzieję, że seminarium poświęcone metodzie reprezentacyjnej przyczyni się do jej popularyzacji oraz podjęcia ciągłych prac badawczych, które umożliwią doskonalenie jakości danych statystycznych.

BIBLIOGRAFIA

[1] *Badania statystyczne metodą reprezentacyjną w krajach socjalistycznych,* „Biblioteka Wiadomości Statystycznych", t. 14, GUS, Warszawa 1971.
[2] Cochran W.G.: *Sampling Techniques,* New York 1963.
[3] Dalenius T.: *Bibliography on non-sampling errors,* „International Statistical Review", vol. 45, 1977.
[4] Dalenius T.: *Elements of Survey Sampling,* Stockholm 1986.
[5] Fellegi I.P.: *The evaluation of the accuracy of survey results: some Canadian experience,* „International Statistical Review", vol. 41, 1973.
[6] Fellegi I.P., Sunter A.B.: *Balance between different sources of survey errors – some Canadian experiences.* „Sankhya", series C, vol. 36, 1974.
[7] Hansen M.H., Hurwitz W.N., Madow W.G.: *Sample Survey Methods and Theory,* vol. I i II, New York 1953.
[8] Kish L.: *Survey Sampling,* New York 1965.
[9] Kordos J.: *Dokładność danych w badaniach społecznych,* „Biblioteka Wiadomości Statystycznych", GUS, t. 35, Warszawa 1987.
[10] Kordos J.: *Jakość danych statystycznych,* PWE (w druku).
[11] *Metodologia badań reprezentacyjnych w GUS* — prace Komisji Matematycznej, „Biblioteka Wiadomości Statystycznych", t. 29, GUS, Warszawa 1979.
[12] Moser C.A., Kalton G.: *Survey Methods in Social Investigation (revised edition),* New York: Basic Books, 1978.
[13] Pawłowski Z.: *Wstęp do statystycznej metody reprezentacyjnej,* PWN, Warszawa 1972.
[14] Platek R.: *Methodology and treatment for non-response,* EUSTAT, Instituto Vasco de Estadistica, 10, 1988.
[15] United Nations. *Non-sampling Errors in Household Surveys: Sources, Assessment and Control (Preliminary version),* New York 1982.
[16] *Wybrane problemy metodologiczne badań reprezentacyjnych,* „Biblioteka Wiadomości Statystycznych", GUS, t. 15, Warszawa 1971.
[17] Zarkovich S.S.: *Quality of Statistical Data,* FAO, Rome 1966.
[18] Zasępa R.: *Badania statystyczne metodą reprezentacyjną.* PWN, Warszawa 1962.
[19] Zasępa R.: *Metoda reprezentacyjna,* PWE, Warszawa 1970.
[20] *Zastosowanie metody reprezentacyjnej w badaniach statystycznych GUS (1981–1986),* „Z prac Zakładu Badań Statystyczno-Ekonomicznych", GUS, z. 166, Warszawa 1987.

Wybrane zagadnienia jakości informacji statystycznych

doc. dr hab. Bogdan Stefanowicz
Szkoła Główna Planowania i Statystyki

Jakość informacji w ogóle, w tym także informacji statystycznych, jest jednym z tych tematów, którym poświęcono już wiele rozpraw, lecz które dotąd nie zostały do końca rozwiązane. Jest to temat wieloaspektowy i wielowątkowy. W niniejszym artykule ograniczymy się zaledwie do kilku kwestii, mianowicie: do przedstawienia propozycji pewnego sposobu pomiaru tej jakości oraz do krótkiego nakreślenia kilku wybranych sposobów jej podniesienia. Rozważania te pragniemy oprzeć na tzw. infologicznym podejściu do pojęcia informacji.

INFOLOGICZNA INTERPRETACJA INFORMACJI

Pojęcie informacji nie zostało dotąd zdefiniowane jednoznacznie. Toteż w różnych opracowaniach można spotkać różną jego interpretację. Jedną z nich jest tzw. podejście infologiczne, zapoczątkowane przez B. Sundgrena [5] i B. Langeforse'a [4]. Kierując się zasadą przedstawioną przez tych autorów, podejście to można w pewnym rozwinięciu ująć następująco.

Niech będzie dany układ K:

$$(P, x \in X, t, \mathbf{q})$$

gdzie: K — jest oznaczeniem układu elementów zawartych w nawiasie,
P — jest identyfikatorem (symbolem, nazwą lub innym jednoznacznym określeniem) pewnego obiektu, którym może być dowolny wycinek rzeczywistości, jak obiekt fizyczny, zdarzenie, proces, zjawisko itp., tj. wszystko, co może być przedmiotem obserwacji w badaniach statystycznych,
X — jest oznaczeniem (identyfikatorem, nazwą) cechy obiektu P,
t — oznacza czas, w którym obiekt P jest rozpatrywany ze względu na cechę X,
x — jest wartością cechy X w momencie t,
$\mathbf{q}$ — oznacza wektor dodatkowych charakterystyk, związanych z obiektem P, cechą X i czasem t, jak np. jednostka miary, metoda pomiaru i rejestracji x itd.

Układ K można rozpatrywać dwojako:

a) z punktu widzenia strukturalnego i w tym ujęciu K będziemy nazywać komunikatem;

b) z punktu widzenia znaczeniowego (semantycznego), jako opis obiektu P ze względu na cechę X w czasie t; relację wiążącą elementy komunikatu K nazwiemy informacją; K jest nośnikiem tej informacji.

W celu odróżnienia komunikatu K od zawartej w nim informacji tę ostatnią zapiszemy $I(K)$.

Jeżeli obiekt P, występujący w komunikacie K, jest elementem pewnej zbiorowości statystycznej lub sam jest taką zbiorowością, to I(K) będziemy traktować jako informację statystyczną.

Zbiór $\boldsymbol{P}$ wszystkich rozpatrywanych obiektów $\boldsymbol{P}$ wraz ze zbiorem $\boldsymbol{X}$ branych pod uwagę cech $\boldsymbol{X}$ tych obiektów w przedziale czasu $\boldsymbol{T}$ wyznaczają przestrzeń informacyjną $\boldsymbol{I}$, którą można określić jako iloczyn kartezjański:

$$\boldsymbol{I} = \boldsymbol{P} \times \boldsymbol{X} \times \boldsymbol{T}$$

gdzie: T jest traktowane jako zbiór wyróżnionych punktów czasu w rozpatrywanym okresie, zaś „$\times$" jest znakiem iloczynu kartezjańskiego.

Każdy element przestrzeni $\boldsymbol{I}$ jest określoną informacją $I(K)$, którą będziemy nazywać informacją jednostkową. Innymi słowy, przestrzeń $\boldsymbol{I}$ traktujemy jako zbiór informacji jednostkowych $I(K)$.

Przedstawiona interpretacja informacji i przestrzeni informacyjnej oparta jest na obiektywnym istnieniu elementów składowych odpowiednich komunikatów K. W tym sensie istnienie informacji $I(K)$ należy uznać za fakt obiektywny, niezależny od obserwatora obiektu P. Podobnie można uznać obiektywne istnienie przestrzeni informacyjnej $\boldsymbol{I}$, przy wyróżnionym zbiorze $\boldsymbol{P}$, zbiorze $\boldsymbol{X}$ oraz okresie T.

W praktyce użytkownika U, rozwiązującego określony problem R, interesuje nie cała przestrzeń $\boldsymbol{I}$, lecz tylko pewien jej podzbiór $\boldsymbol{I'} \subset \boldsymbol{I}$, do którego należą takie informacje $I(K,U,R)$, które są niezbędne lub przydatne dla U przy rozwiązywaniu R.

W tym ujęciu informacje $I(K,U,R)$ nabierają charakteru subiektywnego, co w szczególności wyraża się w sposobie określania podzbioru $\boldsymbol{I'}$ oraz w formułowaniu ocen przydatności i jakości elementów tego podzbioru, tzn. jednostkowych informacji $I(K,U,R)$, które doń należą. Z tego powodu obiektywnie istniejąca ta sama informacja $I(K)$ może być różnie oceniana przez różnych użytkowników U_1 i U_2, jako różne w sensie subiektywnym elementy $I(K,U_1,R)$ i $I(K,U_2,R)$ różnych podzbiorów I_1' i I_2'.

JAKOŚĆ INFORMACJI

Prezentując infologiczne podejście do pojęcia informacji, ukazaliśmy jej subiektywny i obiektywny charakter. Z tego też względu o jakości informacji można mówić w ujęciu obiektywnym i subiektywnym. W dalszej części o jakości informacji będziemy mówić wyłącznie przy założeniu jej

rozpatrywania z uwzględnieniem użytkownika, któremu informacja ma służyć do rozwiązania konkretnego zadania.

Do jakości informacji można odnieść kryteria, jakie brane są pod uwagę przy analizie jakości w ujęciu ogólnym:

1) przydatność, tzn. zbiorczy, wynikowy stopień spełnienia wymagań formułowanych przez użytkownika w zakresie przeznaczenia rozpatrywanego produktu (w naszym przypadku — informacji $I(K,U,R)$),
2) poprawność rozumianą jako stopień spełnienia wymagań, dotyczących wytworzenia produktu (informacji),
3) użyteczność (operatywność, skuteczność), oznaczająca stopień spełnienia zdefiniowanych przez użytkownika wymagań użytkowych,
4) doznaniowość (zadowolenie, satysfakcja) traktowana jako stopień spełnienia wymagań doznaniowych,
5) opłacalność (oszczędność, efektywność) interpretowana jako stopień spełnienia wymagań ekonomicznych.

W odniesieniu do informacji (w tym także statystycznych) te pięć grup ogólnych rozwija się i uściśla stosownie do potrzeb i wymagań formułowanych przez użytkownika. W literaturze specjalistycznej można znaleźć różne wykazy szczegółowych cech jakościowych informacji (będziemy je dalej nazywać cechami pożądanymi). Na ogół jednak w ich określaniu można dostrzec pewne nieścisłości. Wynika to prawdopodobnie z dwóch powodów. Po pierwsze, samo pojęcie informacji nie zostało do końca wyjaśnione. Po drugie, nieprecyzyjne i niejednolite są zasady, według których są definiowane cechy pożądane. Prowadzi to czasami do mylenia cechy pożądanej informacji z właściwościami źródeł tej informacji, z charakterystyką metod zbierania informacji itp.

Trudności w tym zakresie, w pewnym stopniu, dają się pokonać przy infologicznym podejściu. Mianowicie, pozwala ono przybliżyć pojęcie informacji bez konieczności odwoływania się do innych pojęć, jak: entropia, odbicie, moc zbioru, prawdopodobieństwo itd. Pozwala też określić informację jednostkową, opisującą jeden wybrany obiekt, jak i informację odnoszącą się do całego zbioru takich obiektów, przy wskazaniu konkretnej badanej cechy i określonego momentu obserwacji. Wyróżnienie zaś poszczególnych elementów składowych w formule $I(K,U,R)$, gdzie komunikat K składa się z kolei z identyfikatora obiektu P, identyfikatora cechy X, jej wartości x, oznaczenia wyodrębnionego momentu czasu t i innych charakterystyk ujętych w wektorze $\mathbf{q}$, pozwala odnieść każdą cechę pożądaną informacji do odpowiedniego „adresata", co ułatwia dokładniejsze jej określenie.

POMIAR JAKOŚCI INFORMACJI STATYSTYCZNEJ

Jednym z istotnych zagadnień jakości informacji statystycznej jest kwantyfikacja osądów w tym zakresie. Do interesujących propozycji pod tym względem trzeba zaliczyć pracę [3]. Jej autor prezentuje pewne sposoby

oceny dokładności danych i precyzji stosowanych estymatorów z uwzględnieniem występowania w zgromadzonym materiale statystycznym takich błędów, jak tzw. błędy pokrycia, błędy wynikające z braku danych (odpowiedzi) w niektórych pozycjach dokumentów źródłowych oraz błędy obciążające poszczególne zebrane dane jednostkowe. Tutaj pragniemy zaprezentować inną próbę kwantyfikacji ocen w sprawie jakości informacji statystycznej. Oprzemy ją na infologicznej interpretacji pojęcia informacji i związanym z nią pojęciem komunikatu. W tym celu na początku zbiór $\boldsymbol{C}$ wszystkich cech pożądanych podzielimy na dwa podzbiory:

a) cechy odnoszące się zarówno do informacji jednostkowych $I(K,U,R)$, zawartych w pojedynczych komunikatach K, jak i do informacji zawartych w zbiorach $\boldsymbol{K}$ komunikatów K; do cech takich można odnieść np. aktualność informacji, terminowość, dokładność, zrozumiałość itp.; zbiór ten oznaczymy $\boldsymbol{C}_1$;
b) cechy odnoszące się wyłącznie do informacji zawartych w zbiorach $\boldsymbol{K}$ komunikatów, jak np. redundancja (nadmierność informacyjna), kompletność, niesprzeczność; zbiór tych cech oznaczymy $\boldsymbol{C}_2$.

Sprawdzenie, czy pewna jednostkowa informacja $I(K,U,R)$ ma cechę $C \in \boldsymbol{C}_1$, można sprowadzić do badania, czy cechę tę ma odpowiedni element komunikatu K w stopniu satysfakcjonującym użytkownika U rozwiązującego problem R. Jeżeli każde takie wymaganie U określi zawczasu w postaci warunków W (np. tak jak to jest czynione przy opracowywaniu reguł kontrolnych dla celów automatycznej kontroli danych wejściowych), to każda jednostkowa informacja $I(K,U,R)$ może być jednoznacznie oceniona ze względu na wyróżnioną cechę C; jeżeli warunek W jest spełniony, to dana informacja ma cechę C, w przeciwnym przypadku informacja $I(K,U,R)$ tej cechy nie ma.

Ocena informacji dostarczonych przez zbiór $\boldsymbol{K}$ komunikatów pod względem cechy $C \in \boldsymbol{C}_1$ lub cechy $C \in \boldsymbol{C}_2$ — przy jednakowym znaczeniu (ważności) dla użytkownika U wszystkich jednostkowych informacji $I(K,U,R)$ — oznacza badanie, ile komunikatów $K \in \boldsymbol{K}$ cechę C ma w stopniu satysfakcjonującym U oraz ile komunikatów należących do $\boldsymbol{K}$ tej cechy nie ma. Jeżeli liczbę wszystkich komunikatów $K \in \boldsymbol{K}$ oznaczymy przez n, zaś liczbę komunikatów zbioru $\boldsymbol{K}$, nie mających cechy C, oznaczymy przez m, wtedy miarą spełnienia przez podzbiór informacji $\boldsymbol{I}(\boldsymbol{K},\boldsymbol{U},\boldsymbol{R})$ wymagania sformułowanego przez użytkownika U w postaci cechy C może być funkcja $f(n,m)$:

$$f(n,m)=\frac{n-m}{n}$$

Funkcja ta przyjmuje wartości z przedziału od 0 do 1 i pozwala na ocenę informacji $\boldsymbol{I}(\boldsymbol{K},\boldsymbol{U},\boldsymbol{R})$ bez względu na semantyczne znaczenie cechy C. Przy tym liczbę m, występującą w funkcji $f(n,m)$, można oszacować na podstawie

analizy błędów podczas ręcznej kontroli dokumentów źródłowych lub po automatycznej kontroli danych wejściowych.

Funkcja $f(n,m)$ pozwala na przeprowadzenie oceny jakości informacji tylko ze względu na pojedyncze cechy C. Jeżeli użytkownik U bierze pod uwagę pewien zestaw cech pożądanych i pragnie ocenić jakość informacji łącznie ze względu na te cechy, to procedura postępowania może przedstawiać się następująco:

1. Abstrahując od liczby i rodzaju branych pod uwagę cech C, użytkownik U wyznacza pewną skalę S o wartościach od 1 do N, według której będzie prowadzona dalsza analiza jakości informacji, ze względu na wyróżnione cechy C. Rozpatrywany zbiór tych cech oznaczmy przez $\boldsymbol{C}$.

2. Cechy $C \in \boldsymbol{C}$ są porządkowane przez U stosownie do ich „ważności" z uwzględnieniem skali S. Każda cecha C otrzymuje swój numer k_i równy numerowi i-tej pozycji na skali S, gdzie ją użytkownik umieścił. Przyjmijmy, że cecha uznana przez U za najważniejszą otrzyma numer bliższy początkowi skali S (np. nr 1), zaś cecha mniej ważna zostanie oznaczona numerem zbliżającym się do końca skali (np. ostatni numer skali równy N). Cechy uznane przez U za równorzędne otrzymają ten sam wspólny numer, tzn. zostaną odniesione do wspólnej i-tej grupy. Niektóre wartości na skali S mogą pozostać nie wykorzystane, jeżeli U uzna, iż żadna z cech $C \in \boldsymbol{C}$ nie zasługuje na jej umieszczenie na danym miejscu.

3. Dla każdej cechy $C_j \in \boldsymbol{C}$ oblicza się dwa parametry:

a) $v_j = \dfrac{n - m_j}{n}$

gdzie: m_j — oznacza liczbę komunikatów nie mających cechy C_j, zaś n jest liczbą wszystkich rozpatrywanych komunikatów;

b) $w_j = \dfrac{N - (k_i - 1)}{N}$

gdzie: w_j — oznacza wagę cechy C_j,
k_i — jest numerem grupy, do której cecha C_j została zaliczona na skali S w kroku 2.

4. Oblicza się łączny wskaźnik v jakości informacji $I(K, U, R)$:

$$v = \frac{w_1 v_1 + \ldots + w_N v_N}{w_1 + \ldots + w_N}$$

Dane do wyliczenia v_j można uzyskać z różnych źródeł. Na przykład w odniesieniu do informacji wejściowych można wyróżnić:

— przeprowadzenie dodatkowych badań wyrywkowych, których celem jest zweryfikowanie, czy zebrane dane źródłowe posiadają określone cechy pożądane (np. czy są rzetelne);

— ręczną analizę (kontrolę) dokumentów źródłowych, w toku której można dokonać sprawdzenia wszystkich lub wyrywkowo wybranych dokumentów, w celu oceny zawartych w nich danych pod względem wyróżnionych cech C_j;
— analizę wykazu błędów, drukowanego w procesie kontroli automatycznej (programowej) danych, po ich zapisaniu na nośniku maszynowym; trzeba tylko w tym przypadku zadbać o to, aby wykaz wykrytych błędów zawierał niezbędne informacje pozwalające ustalić, której z cech C_j kontrolowane dane nie posiadają; uczynić to można poprzez odpowiednie przygotowanie algorytmu tej kontroli, a więc podczas projektowania procesów przetwarzania danych.

W odniesieniu do informacji wynikowych można wymienić:
— analizę wyników w oparciu o odpowiedni aparat statystyki matematycznej (np. do oceny dokadności wyników uzyskanych drogą badań reprezentacyjnych — patrz np. [3]);
— porównanie otrzymanych informacji z innymi analogicznymi informacjami, pochodzącymi z innych źródeł;
— pozyskanie opinii zainteresowanego użytkownika (np. odnośnie dokładności, kompletności i przydatności informacji).

CECHY NIEPOŻĄDANE

Rozpatrując jakość informacji statystycznych braliśmy pod uwagę tylko ich cechy pożądane, tzn. takie, które informacja powinna mieć w stopniu określonym przez użytkownika. Tymczasem jakość tę determinują także cechy, których informacja mieć nie powinna, tj. cechy niepożądane. Interesujący ich przegląd podał S. Garczyński [2]. Autor ten wymienia w szczególności takie cechy, jak jednostronność informacji, jej nieprzystosowanie do rzeczywistych potrzeb użytkownika, eksponowanie i wyolbrzymianie jednych informacji kosztem innych, „rozmydlanie" informacji, przez podawanie nadmiernej liczby różnych szczegółów, co odwraca uwagę odbiorcy od istoty opisu danego wycinka rzeczywistości i szereg innych.

Analiza jakości informacji wymaga tedy uwzględnienia dwóch czynników: wpływu na nią cech pożądanych oraz wpływu cech niepożądanych. Nie będziemy jednak bliżej zajmować się tutaj tym zagadnieniem.

SPOSOBY PODNIESIENIA JAKOŚCI INFORMACJI STATYSTYCZNYCH

Jakość informacji jest uzależniona od szeregu czynników, takich jak: czas, założenia organizacyjne, założenia metodologiczne badania statystycznego, wykorzystywany sprzęt itp. Szerszy przegląd tych czynników można znaleźć

w rozprawach poświęconych tym zagadnieniom [patrz np. 1 lub 3]. Tutaj pragniemy przedstawić wybrane rozwiązania nie zawsze dostrzegane w innych pracach i zastosowaniach praktycznych.

▲ **Automatyczna korekta błędów.** Jednym z ważnych sposobów podniesienia jakości informacji statystycznych jest usunięcie błędów, jakie mogą wystąpić w zebranych danych źródłowych. W procesie tym można wyróżnić dwa główne etapy: wykrywanie błędów w tych danych oraz ich usuwanie, przez odpowiednie skorygowanie. W praktyce od wielu już lat jest stosowana tzw. automatyczna (programowa) kontrola tych danych, tzn. kontrola, którą wykonuje komputer na podstawie odpowiedniego programu. Początkowo statystycy nieufnie odnosili się do niej, nie wierząc, że maszyna może być zdolna do wykonywania pracy związanej z przeprowadzaniem analizy każdej danej i dokonywaniem jej oceny pod względem poprawności. Wkrótce jednak okazało się, że komputer pod tym względem jest niezastąpiony przez człowieka, a nawet liczne zespoły specjalistów nie mogą z nim konkurować w tym zakresie.

Pozostaje wszakże kwestia korygowania wykrytych przez maszynę błędów. Nadal ten etap przetwarzania danych pozostaje tradycyjnie nie zautomatyzowany, tzn. odpowiednie prace muszą być wykonywane przez człowieka. Korekta ręczna ma swoje zalety; pozwala indywidualnie rozpatrywać każdy jednostkowy błąd oraz podejmować indywidualne działania, w celu ustalenia prawdziwej wartości, jaka powinna być wstawiona na jego miejsce.

Ale też sposób ten ma wiele wad. Do ważniejszych trzeba zaliczyć jego czasochłonność oraz częste niekonsekwencje w dobieraniu wartości korygujących. W rezultacie korekta ręczna, zamiast podnosić jakość informacji, obniża ją poprzez opóźnianie wyników (obniżenie jakości informacji pod względem terminowości i aktualności) oraz przez wprowadzanie wartości, korygujących według subiektywnej oceny osoby wykonującej tę pracę. Trzeba też pamiętać o pracochłonności i kosztach przetwarzania danych, co także jest jedną z ocen jakości informacji.

Z tego względu celowe jest odwołanie się do korekty maszynowej (automatycznej), wykonywanej przez komputer na podstawie odpowiedniego programu. Okazuje się, że w wielu przypadkach programy takie można opracować w sposób zapewniający automatyczne skorygowanie wykrytych błędów bez obciążenia (w sensie statystycznym) wyników przetwarzania. Dowodzą tego różnorodne badania oraz doświadczenia praktyczne GUS, jak i innych krajów. Przy tym cała praca jest wykonywana nieporównywalnie szybciej niż potrafi uczynić to człowiek. Zachodzi też możliwość zapewnienia pełnej kontroli tej opreacji oraz zebrania niezbędnych informacji do oceny jej skutków.

Nadal jednak pozostaje nieuzasadniona nieufność do jej stosowania. Być może dlatego, że w istocie statystycy nie zajmują się bezpośrednio korektą

błędów i z tego względu nie odczuwają uciążliwości pracy ręcznej w tym zakresie, co najwyżej narzekając na opóźnienie wyników.

Nie bez winy pozostają też informatycy; zazwyczaj zbyt biernie oczekują na inicjatywę ze strony statystyków. W rezultacie długie wykazy błędów nadal koryguje się ręcznie ze wszystkimi ujemnymi skutkami.

▲ **Projektowanie formularzy statystycznych.** Dane statystyczne, w obecnych warunkach, są zbierane z wykorzystaniem odpowiednich formularzy statystycznych. Przy ich projektowaniu zazwyczaj dąży się do ich minimalizacji pod względem ilości rejestrowanych danych (z zachowaniem niezbędnej treści informacyjnej). Zapewnia to skrócenie czasu sporządzania takich formularzy, zmniejszenie liczby popełnianych błędów itp. Warunek ten pozostanie zapewne w mocy nawet wtedy, kiedy zostanie rozpowszechniony zwyczaj przekazywania źródłowych informacji sprawozdawczych do systemu statystycznego za pomocą maszynowych nośników danych, jak taśmy magnetyczne lub dyskietki.

Ta minimalizacja musi jednak uwzględnić nie tylko wspomniane warunki, tzn. zachowanie niezbędnej informacji, zmniejszenie kosztów opracowania, zmniejszenie liczby popełnionych błędów itp., lecz także dalsze etapy procesu przetwarzania danych. Do tych etapów należą w szczególności wspomniane wyżej — automatyczna kontrola oraz automatyczna korekta wykrytych błędów. Jak zauważono, automatyzacja tych prac jest wskazana z powodów praktycznych i może przyczynić się do podniesienia jakości informacji, ze względu na różne cechy pożądane. Aby jednak było możliwe opracowanie odpowiednich programów komputerowych, trzeba zapewnić niezbędny nadmiar informacyjny, potrzebny do zbudowania algorytmów kontroli automatycznej oraz algorytmów korekty automatycznej. Nadmiar taki musi być zaprojektowany właśnie podczas prac przygotowawczych nad formularzem statystycznym. Zaniedbanie tego prowadzi do przypadkowych rozwiązań, które nie w pełni wykorzystują możliwości podniesienia jakości informacji wynikowych przez prostą i skuteczną eliminację błędów.

▲ **Prezentacja wyników.** Jedną z cech jakości informacji jest jej czytelność i przejrzystość. Na to zaś składa się m.in. sposób prezentacji wyników przetwarzania. Dostępny obecnie sprzęt, zwłaszcza mikrokomputery, stwarza duże możliwości pod tym względem. Trzeba tylko pewnej dozy wyobraźni, pomysłowości i cierpliwości (głównie ze strony informatyków), aby praktycznie wydobyć całą pełnię form przedstawienia wyników w sposób najbardziej dostosowany do potrzeb użytkowników.

Warto nadmienić, że na ogół celowe jest ostateczne redagowanie wyników bezpośrednio przez zainteresowanych statystyków. Oni to bowiem znają przeznaczenie wytwarzanych informacji oraz dokładniej znają ich finalnego odbiorcę. W takim razie jest im łatwiej dobrać odpowiednią postać i układ graficzny wyniku. Wymaga to jednak bliższego zainteresowania się przez

statystyków (a ogólniej — przez użytkowników informacji statystycznych) współczesnym sprzętem i możliwościami jego wykorzystania.

▲ **Bazy danych.** Technologia przetwarzania danych w oparciu o bazy danych rozpowszechniła się w świecie informatycznym w wyniku wielu swoich zalet. W szczególności sprzyja ona: przyspieszeniu dostarczenia informacji selektywnie dostosowanych do specyficznych potrzeb konkretnych użytkowników, przyczynia się do zmniejszenia kosztów i pracochłonności przygotowania tych informacji, ułatwia rozwiązanie problemów spójności i porównywalności danych, odnoszących się do różnych obiektów lub pochodzących z różnych okresów, źródeł itp. Zwłaszcza wobec możliwości wykonania wielu obliczeń, z wykorzystaniem mikrokomputerów, celowe staje się budowanie takich baz z ogólnym przeznaczeniem i zapewnieniem możliwości wybierania dowolnych podzbiorów stosownie do potrzeb zainteresowanych użytkowników-statystyków. Trzeba tylko zauważyć, że musi nastąpić zrewidowanie dotychczasowych sposobów budowania i wdrażania tych baz w praktyce krajowej. Z jednej strony, informatycy muszą dostosować się do realnych uwarunkowań pracy statystyków i oddawać im do eksploatacji bazy uwzględniające rzeczywiste, a nie teoretycznie wyobrażane potrzeby. Z drugiej zaś, użytkownicy, w tym także statystycy, muszą uświadomić sobie korzyści, wynikające z zastępowania pracy człowieka pracą maszyny (komputera), w tym także w zakresie wyszukiwania i opracowywania danych statystycznych.

▲ **Wykorzystanie systemów ekspertowych.** Organizacja procesu ustalania potrzeb informacyjnych użytkownika, przygotowywanie odpowiedniego badania statystycznego, zbieranie i opracowywanie danych, a także interpretacja uzyskanych wyników wymagają posiadania rozległej wiedzy teoretycznej oraz doświadczeń praktycznych. Niestety, nie zawsze warunek ten może być spełniony. Toteż dużego znaczenia nabiera zapewnienie niezbędnych konsultacji i pomocy ze strony odpowiednich specjalistów.

W ostatnim dziesięcioleciu w informatyce poczęły się rozwijać koncepcje budowania „sztucznych specjalistów", tzn. systemów ekspertowych. Zadaniem takich systemów jest rozwiązywanie nietrywialnych zadań z wykorzystaniem metody sztucznej inteligencji. Wprawdzie w statystyce tego rodzaju rozwiązania zaczynają sobie dopiero nieśmiało torować drogę, to jednak wykazały już pewną praktyczną użyteczność; pozwalają szybciej opracować niektóre analizy statystyczne, dopomagają we właściwym opracowaniu planów obliczeń itp.

Wyniki te zachęcają do sformułowania tezy o celowości upowszechnienia tego rodzaju rozwiązań informatycznych w różnych zastosowaniach w statystyce; przy projektowaniu badań statystycznych z zachowaniem uwarunkowań, wynikających z przesłanek teoretycznych, przy planowaniu procesu przetwarzania danych źródłowych, do ustalania zakresu i metod analiz statystycznych danych itp.

Pewne prace w tym zakresie zostały już podjęte w GUS. Szersze upowszechnienie systemów eksportowych wymaga jednak zainteresowania i udziału ze strony statystyków, a także pozyskania odpowiednich środków finansowych. Jest to zadanie czasochłonne i pracochłonne, lecz w razie wdrożenia tego rodzaju systemu w praktyce pozwoli zapewnić większą fachowość w organizowaniu szeregu procesów, związanych ze zbieraniem i opracowywaniem informacji statystycznych, a więc podnieść ich jakość.

▲ **Wykorzystanie wykazów błędów do eliminowania przyczyn ich powstawania.** Pomocą w uniknięciu błędów jest poznanie przyczyn i źródeł ich powstawania. Jednym ze sposobów takiego poznania są studia i rozważania teoretyczne. Przykładem takiego podejścia są schematy analizy błędów losowych, pojawiających się wraz ze stosowaniem metody reprezentacyjnej w badaniach statystycznych [patrz np. 3]. Jakkolwiek ma ono szereg ważnych zalet, to jednak nie zawsze może być zrealizowane. W szczególności trudno jest tą drogą poznać rzeczywiste przyczyny pojawienia się błędów, w zebranym materiale źródłowym. Wtedy celowa staje się analiza błędów wykrytych w poprzednio przeprowadzonych badaniach. Może to stać się podstawą do określenia mechanizmów, wywołujących określony typ zniekształceń zbieranych danych. Sformułowane wnioski mogą dopomóc w określeniu tych środków zaradczych, których zastosowanie pozwoli na uniknięcie lub przynajmniej zminimalizowanie analogicznych błędów w kolejnych badaniach podobnego rodzaju. Praktyka dowodzi, że takie postępowanie jest nader skuteczne nie tylko w odniesieniu do badań opartych na cyklicznej sprawozdawczości statystycznej, lecz także wszędzie tam, gdzie są podejmowane pokrewne badania, ze względnie dużą częstotliwością ich realizacji.

UWAGI KOŃCOWE

Przedstawione w artykule rozważania na temat jakości informacji statystycznej stanowią nieśmiałą próbę włączenia się autora w nurt dyskusji na ten temat. Nie absolutyzując przedstawionych propozycji, warto zwrócić uwagę na fakt, iż o jakości informacji można mówić z różnych punktów widzenia, wśród których ważnymi wydają się być następujące:

a) podkreślające złożoność i wieloaspektowość pojęcia jakości informacji, w tym informacji statystycznej;
b) uzyskujące subiektywny charakter pojęcia jakości informacji;
c) proponujące pewną obiektywizację ocen;
d) wymuszające zarówno użytkowników informacji, jak i specjalistów od ich zbierania, gromadzenia, przetwarzania i udostępniania do określania i uzgadniania ilościowo wyrażanego poziomu jakości, co może przyczynić się do zracjonalizowania szeregu prac związanych z procesami wytwarzania informacji. Cechy te, jak się wydaje, mają podejście infologiczne.

BIBLIOGRAFIA

[1] Buśko B., Filipek H., Śliwieński J.: *Wiarygodność informacji ekonomicznej w systemach informatycznych;* PWE, Warszawa 1980.
[2] Garczyński S.: *Z informacją na bakier;* Inst. Wyd. Zw. Zaw., Warszawa 1984.
[3] Kordos J.: *Jakość danych statystycznych;* PWE, Warszawa 1988.
[4] Langeforse B.: *Infological models and information users view;* Information Systems 1980, vol. 5.
[5] Sundgren B.: *An infological approach to data bases;* Lund, Skriftseries Statistiska Centralbyran, Sztokholm 1973.

Współczynniki rzetelności danych wielowymiarowych

dr Janusz Wywiał
Akademia Ekonomiczna w Katowicach

Wynik wnioskowania statystycznego w dużej mierze zależy od jakości danych o populacji, które są uzyskiwane podczas badania reprezentacyjnego. Poznanie źródeł błędów, popełnianych przy gromadzeniu danych oraz mechanizmu ich wpływu na wyniki np. estymacji, jest etapem wstępnym przy poszukiwaniu metod, prowadzących do polepszenia precyzji wnioskowania statystycznego. Rzeczą ważną staje się problem pomiaru stopnia rzetelności danych statystycznych, co z kolei pozwala ocenić wpływ jakości danych na dokładność np. estymacji.

Wśród błędów, pojawiających się na etapie gromadzenia danych, wyróżnia się te, które są spowodowane niedokładnością obserwacji wartości zmiennych w populacji. Ponadto mogą wystąpić już na etapie wnioskowania statystycznego, tzw. błąd pokrycia lub błąd związany z brakiem odpowiedzi. Pierwszy z nich jest konsekwencją losowania próby z części populacji zamiast z całej, a z drugim mamy do czynienia wtedy, gdy obserwacje zmiennych w próbie są niekompletne.

W niniejszym referacie nasze rozważania ograniczamy do określenia współczynników rzetelności, mierzących dokładność obserwacji wielowymiarowej zmiennej.

MODEL BŁĘDÓW OBSERWACJI

Błąd obserwacji wartości zmiennej polega na tym, że uzyskana informacja o poziomie wartości badanej zmiennej nie jest zgodna z jej stanem rzeczywistym. Źródeł nierzetelności [10] i [13], można dopatrywać się np.

w niedokładności przyrządów pomiarowych użytych do obserwacji wielkości fizycznych, a w przypadku badań opartych na wywiadzie — w niedoskonałości sposobu jego prowadzenia, co m.in. może być spowodowane niewłaściwym formułowaniem pytań ankietowych.

Generowanie się błędów obserwacji jest zwykle opisywane modelem addytywnym z jednoczesnym wprowadzeniem pojęcia nadpopulacji [2], [3], [7], [8]. Oznaczmy przez $u=\{u_j,\ j=1,2,...,N\}$ zbiór nazywany populacją skończoną. Przez $\mathbf{y}=[y_{ij}]$ $(i=1,...,m;\ j=1,...,N)$ oznaczamy macierz, której j-ta kolumna jest jednoznacznie przyporządkowana j-temu elementowi populacji. Kolumna ta stanowi obserwację m-wymiarowej zmiennej. Elementy i-tego wiersza macierzy $\mathbf{y}$ to wartości i-tej zmiennej, przypisane kolejnym elementom populacji.

Oznaczmy przez $\mathbf{x}$ macierz o wymiarach $m\times N$, której elementy są błędnymi obserwacjami elementów macierzy $\mathbf{y}$. Zbiór wszystkich możliwych wyników obserwacji elementów macierzy $\mathbf{y}$ oznaczamy symbolem U, który nazywany jest nadpopulacją populacji u. Zbiór U może składać się ze skończonej bądź nieskończonej liczby elementów, będących macierzami typu $\mathbf{x}$. Macierz $\mathbf{x}$, należąca do nadpopulacji U jest traktowana jako wartość pewnej macierzy losowej $\mathbf{X}=[X_{ij}]$. Rozkład prawdopodobieństwa zmiennej losowej $\mathbf{X}$ określa szanse pojawienia się błędnej obserwacji $\mathbf{x}$ macierzy $\mathbf{y}$. Przyjmujemy, że dla $i=1,...,m$ oraz $j=1,...,N$

$$\begin{cases} X_{ij}=y_{ij}+B_{ij} \\ B_{ij}=b_{ij}+Z_{ij} \end{cases} \tag{1}$$

Symbolem B_{ij} oznaczono błąd obserwacji wartości y_{ij}, który rozpada się na składnik nielosowy b_{ij} i składnik losowy Z_{ij}. Nielosową część błędu nazywamy obciążeniem obserwacji wartości zmiennej. Zwykle zakłada się, że rozkład losowy błędów obserwacji charakteryzują parametry:

$$\begin{cases} E(Z_{ij})=0 \\ C_{ov}(Z_{ij}, Z_{tv})=c(Z_{ij}, Z_{tv}) \\ i,t=1,2,...,m;\ j,v=1,...,N \end{cases} \tag{2}$$

W dalszej analizie wygodnie będzie wartość i-tej zmiennej oznaczać symbolem $y_i=[y_{i1}, y_{i2}, ..., y_{iN}]$, jej obciążenia przez $b_i=[b_{i1}, b_{i2}, ..., b_{iN}]$, natomiast losowy błąd obserwacji i-tej zmiennej oznaczamy przez $Z_i=[Z_{i1}, Z_{i2}, ..., Z_{iN}]$, a błędną obserwację przez $X_i=[X_{i1}, X_{i2}, ..., X_{iN}]$.

Wektory wartości średnich w populacji określonych zmiennych oznaczamy odpowiednio symbolami $\bar{\mathbf{y}}=\bar{y}_1, \bar{y}_2, ..., \bar{y}_m]$, $\bar{\mathbf{b}}=[\bar{b}_1, \bar{b}_2, ..., \bar{b}_m]$, $\bar{\mathbf{Z}}=[\bar{Z}_1, \bar{Z}_2, ..., \bar{Z}_m]$ i $\bar{\mathbf{X}}=[\bar{X}_1, \bar{X}_2, ..., \bar{X}_m]$. Macierze kowariancji w populacji są następujące:

$$\mathbf{C}_{yy}=[c(y_i,y_t)], \quad \mathbf{C}_{bb}=[c(b_i,b_t)], \quad \mathbf{C}_{ZZ}=[c(Z_i,Z_t)], \quad \mathbf{C}_{XX}=[c(X_i,X_t)],$$
$$\mathbf{C}_{yb}^T=\mathbf{C}_{by}=[c(b_i,y_t)] \quad i,t=1,2,...,m.$$

przy czym np.:

$$c(y_i,y_t)=\frac{1}{N}\sum_{j=1}^{N}(y_{ij}-\bar{y}_i)(y_{tj}-\bar{y}_t); \qquad \bar{y}_i=\frac{1}{N}\sum_{j=1}^{N}y_{ij}$$

Wariancję danej zmiennej oznaczamy przez s^2, czyli np. $s^2(y_i)=c(y_i,y_i)$.

Wartości oczekiwane wektora $\bar{\mathbf{X}}$ oraz macierzy $\mathbf{C}_{XX}$, wyznaczane na podstawie łącznego rozkładu błędów losowych obserwacji, są następujące:

$$E(\bar{\mathbf{X}})=\bar{y}+\bar{b} \tag{3}$$

$$E(\mathbf{C}_{XX})=\mathbf{C}_{yy}+C_{by}+\mathbf{C}_{yb}+\mathbf{C}_{bb}+E(\mathbf{C}_{ZZ}) \tag{4}$$

gdzie: $E(C_{zz})=[E(\mathrm{C}(Z_i,Z_t))]$ (5)

$$E(C(Z_i,Z_t))=\frac{N-1}{N}(k(Z_i,Z_t)-k_w(Z_i,Z_t)) \tag{6}$$

$$k(Z_i,Z_t)=\frac{1}{N}\sum_{j=1}^{N}c(Z_{ij},Z_{tj}) \tag{7}$$

$$k_w(Z_i,Z_t)=\frac{1}{N(N-1)}\sum_{j=1}^{N}\sum_{h\neq j}^{N}c(Z_{ij},Z_{th}) \tag{8}$$

Na podstawie wyrażenia (3) wnioskujemy, że nawet wyczerpujące badanie populacji nie eliminuje obciążenia oszacowania wektora średnich $\bar{y}$ spowodowanego nielosowymi błędami obserwacji. Wzór (4) określa związki występujące między wartością oczekiwaną macierzy wariancji i kowariancji w populacji $\mathbf{C}_{XX}$ a macierzą wariancji i kowariancji w populacji $\mathbf{C}_{yy}$ mierzonych zmiennych.

WSPÓŁCZYNNIKI RZETELNOŚCI

W przypadku jednowymiarowym klasyczny współczynnik rzetelności [4], [6], [14], obliczany jest jako współczynnik korelacji liniowej między wartościami rzeczywistymi badanej zmiennej, a ich obserwacjami, które mogą być błędne. Uważa się, iż wzrostowi dokładności pomiarów wartości zmiennej towarzyszy wzrost wartości tak określonego współczynnika rzetelności.

Formułowane dalej sposoby oceny dokładności pomiaru danych wielowymiarowych bazują również na idei mierzenia ścisłości zależności liniowej

między wektorem mierzonym zmiennych a wektorem pomiarów tych zmiennych.

Oznaczmy przez C wartość oczekiwaną macierzy wariancji i kowariancji w populacji mierzonego wektora zmiennych i wektora jego pomiarów. Macierz tę w postaci bloków można zapisać następująco:

$$\mathbf{C}=\begin{bmatrix} E(\mathbf{C}_{XX}) & E(\mathbf{C}_{Xy}) \\ E(\mathbf{C}_{yX}) & \mathbf{C}_{yy} \end{bmatrix} \tag{9}$$

Macierz $E(\mathbf{C}_{XX})$ określają wzory (4)—(8), natomiast

$$E(\mathbf{C}_{Xy})=\mathbf{C}_{yy}+\mathbf{C}_{by}=E(\mathbf{C}_{yX}^{T}) \tag{10}$$

Pierwszy współczynnik rzetelności określamy na podstawie zamieszczonej w [5] definicji współczynnika alienacji wektorowej, w następujący sposób:

$$l_1=1-\frac{det(\mathbf{C})}{det(E(\mathbf{C}_{XX}))\,det(\mathbf{C}_{yy})}, \tag{11}$$

przy czym zakłada się, iż $det\ (E(\mathbf{C}_{XX}))>0$ i $det(\mathbf{C}_{yy})>0$. Korzystając z określonego np. w [11], str. 50 rozkładu macierzy blokowej wzór (11) przekształcamy do następującej postaci

$$l_1=1-\frac{det(E(\mathbf{C}_{XX})-E(\mathbf{C}_{xy})\,\mathbf{C}_{yy}^{-1}\,E(\mathbf{C}_{xy}))}{det(E(\mathbf{C}_{XX}))}. \tag{12}$$

Stąd i na podstawie wzorów (4) i (10) otrzymamy:

$$l_1=1-\frac{det(\mathbf{C}_{bb}-\mathbf{C}_{by}\mathbf{C}_{yy}^{-1}\mathbf{C}_{yb}+E(\mathbf{C}_{zz}))}{det(E(\mathbf{C}_{XX}))} \tag{13}$$

Oznaczmy przez $q_1^2\geqslant q_2^2\geqslant \ldots q_m^2$ pierwiastki następującego równania:

$$det(E(\mathbf{C}_{xy})\mathbf{C}_{yy}^{-1}E(\mathbf{C}_{yx})-q^2E(\mathbf{C}_{XX}))=0 \tag{14}$$

Wielkości $|q_i|\leqslant 1$ są nazywane współczynnikami korelacji kanonicznej [1], [11].

Nie wdając się w dość złożone szczegóły interpretacji współczynników korelacji kanonicznej można w kontekście naszych rozważań przyjąć, że współczynniki te, co do modułu, są tym większe im silniejsze są liniowe związki, występujące między wektorem mierzonych zmiennych a wektorem ich pomiarów. Korzystając z ogólnych wyników zamieszczonych w [1], [11]

i [12] prawą stronę wyrażenia (12) można sprowadzić do następującej postaci:

$$l_1 = 1 - \prod_{i=1}^{m} (1 - q_i^2) \tag{15}$$

Stąd i z nierówności $|q_i| \leqslant 1 \quad i = 1, ..., m$ wynika, że $0 \leqslant l_1 \leqslant 1$. Wzrostowi przynajmniej jednej z wartości $|q_i| \quad i = 1, ..., m$ odpowiada wzrost wartości parametru l_1, który osiąga swą maksymalną wartość, gdy przynajmniej jeden ze współczynników korelacji kanonicznej jest, co do modułu, równy jedności. Aby ten warunek był spełniony, wystarczy by tylko jedna spośród badanych zmiennych była mierzona bezbłędnie. Wtedy w konsekwencji $l_1 = 1$, mimo że nie wszystkie zmienne są obserwowane bezbłędnie. Po drugie współczynnik $l_1 = 1$ również wtedy, gdy występują tylko nielosowe błędy obserwacji, będące liniową funkcją wartości mierzonych zmiennych, czyli gdy np. $x_{ij} = a_1 y_{ij} + a_0 \quad i = 1, ..., m; \; j = 1, ..., N$. Zatem, w obu wymienionych teraz sytuacjach równa jedności wartość wskaźnika l_1 może mylnie sugerować idealną dokładność obserwacji zmiennych. Należy więc zaznaczyć, że równa jedności lub bliska jedności wartość współczynnika l_1 bynajmniej nie świadczy o dużym stopniu rzetelności obserwacji wektora mierzonych zmiennych. W tej sytuacji można jedynie wnioskować o dużym stopniu nierzetelności obserwacji wektora zmiennych, gdy wartość l_1 jest bliska zeru.

W pracy [12] H. Theil proponuje, by stopień zależności liniowej, pomiędzy parą wektorów zmiennych, mierzyć przy pomocy wartości średniej współczynników korelacji kanonicznej. Korzystając z tej propozycji **drugi współczynnik rzetelności** obserwacji zmiennych określamy następująco:

$$l_2 = \frac{1}{m} \sum_{i=1}^{m} q_i^2 \tag{16}$$

Parametr l_2 przyjmuje oczywiście wartość z przedziału $\langle 0;1 \rangle$. Z rachunkowego punktu widzenia najwygodniej współczynnik l_2 wyliczyć można ze wzoru [1], [11]:

$$l_2 = \frac{1}{m} tr(E(\mathbf{C}_{xy}) \, \mathbf{C}_{yy}^{-1} \, E(\mathbf{C}_{yx}) \, (E(\mathbf{C}_{xx}))^{-1}) \tag{17}$$

przy czym przez *tr* oznaczono ślad macierzy. Podobnie, jak w przypadku parametru l_1 możemy jedynie zasadnie wnioskować o występowaniu małego stopnia rzetelności obserwacji zmiennych, gdy współczynnik l_2 przyjmuje wartości bliskie zeru.

Oznaczmy przez $0 \leqslant d_i \leqslant 1 \quad i=1, ..., m$ wskaźnik determinacji równy kwadratowi współczynnika korelacji wielorakiej, mierzącego stopień ścisłości zależności liniowej, między zmienną będącą obserwacją i-tej zmiennej a wszystkimi m — mierzonymi bezbłędnie zmiennymi. Parametr d_i wskazuje na stopień dopasowania obserwacji X_i zmiennej y_i do regresji liniowej obserwacji X_i względem mierzonych zmiennych $y_1, y_2, ..., y_m$. Teraz kolejny **(trzeci) współczynnik rzetelności** określamy następująco:

$$l_3 = \sum_{i=1}^{m} d_i w_i \tag{18}$$

gdzie: $$w_i = \frac{E(s^2(X_i))}{\sum_{i=1}^{m} E(s^2(X_i))} \tag{19}$$

$$d_i = \frac{E(\mathbf{C}_{Xy}^{(i)})\, \mathbf{C}_{yy}^{-1}\, E(\mathbf{C}_{yX}^{(i)})}{E(s^2(X_i))} \tag{20}$$

przy czym przez $\mathbf{C}_{xy}^{(i)}$ oznaczono i-tą kolumnę macierzy $\mathbf{C}_{xy}$. Stąd, że $0 \leqslant d_i \leqslant 1$, $w_i > 0$ dla każdego $i=1, ..., m$ oraz $\sum_{i=1}^{m} w_i = 1$ wynika, że $0 \leqslant l_3 \leqslant 1$. Małe wartości współczynnika l_3 będą świadczyły o nierzetelności obserwacji wektora zmiennych w populacji.

Interpretacja wartości współczynników l_1, l_2 i l_3 nie jest jednoznaczna, co w szczególności wyjaśniono na przykładzie własności współczynnika l_1. Staje się ona bardziej jasna, gdy uprościmy przyjęty w uprzedniej części pracy model generowania się błędów obserwacji. Załóżmy, że nielosowe błędy obserwacji zmiennych nie są skorelowane z tymi zmiennymi, czyli że $\mathbf{C}_{by} = 0$. Ponadto przyjmijmy, iż błędy losowe obserwacji danej zmiennej nie są skorelowane z błędami losowymi obserwacji innej zmiennej, co analitycznie zapisujemy następująco:
$c(Z_{ij}, Z_{tv}) = 0$, gdy $i \neq t = 1,2, ..., m$. Wówczas na podstawie wzorów (5)—(8) otrzymujemy, że:

$$E(\mathbf{C}_{zz}) = [k(Z_i, Z_t)] = \frac{1}{N} \sum_{j=1}^{N} [c(Z_{ij}, Z_{tj})] \tag{21}$$

Przy określonych teraz dodatkowych dwóch założeniach, podane we wzorach (13) i (17)—(20) współczynniki redukują się do następujących postaci:

$$l_1 = 1 - \frac{det(\mathbf{C}_{bb} + E(\mathbf{C}_{zz}))}{det(\mathbf{C}_{yy} + \mathbf{C}_{bb} + E(\mathbf{C}_{zz}))} \tag{22}$$

$$l_2 = \frac{1}{m} tr(\mathbf{C}_{yy}(\mathbf{C}_{yy} + \mathbf{C}_{bb} + E(\mathbf{C}_{ZZ}))^{-1}) \tag{23}$$

$$l_3 = \frac{\sum_{i=1}^{m} s^2(y_i)}{\sum_{i=1}^{m} (s^2(y_i) + s^2(b_i) + N^{-1} \sum_{j=1}^{N} s^2(Z_{ij}))} \tag{24}$$

Stopień rzetelności obserwacji wartości zmiennych, określonej przy pomocy współczynnika l_3, rośnie, gdy podnosi się dokładność obserwacji zmiennych, tzn. gdy maleje przynajmniej jedna z wariancji błędów obserwacji. Korzystając z własności wyznacznika macierzy zamieszczonych w pracy [11], str. 89, można wykazać, że gdy macierz $\mathbf{C}_{yy}$ jest dodatnio określona i rośnie wartość przynajmniej jednej z wariancji błędu pomiaru oraz niediagonalne elementy macierzy $\mathbf{C}_{yy}$, $\mathbf{C}_{bb}$ i $E(\mathbf{C}_{ZZ})$ nie zmieniają się, to wartości współczynników l_1 i l_2 maleją bądź nie zmieniają się. Zatem w tym przypadku, wraz ze spadkiem dokładności pomiaru poszczególnych zmiennych, określany przy pomocy współczynników l_1 i l_2 syntetyczny poziom rzetelności obserwacji wszystkich zmiennych nie polepsza się, czyli maleje bądź nie zmienia się.

Analizowane współczynniki rzetelności można szacować na podstawie próby bezzwrotnie losowanej z populacji. Estymatory tych parametrów otrzymujemy, zastępując definiujące ją momenty centralne z populacji odpowiednimi momentami z próby.

Na podstawie znanych w statystyce matematycznej twierdzeń granicznych można wykazać, że tak określone estymatory są asymptotycznie nieobciążone i zgodne, gdy liczebność populacji i próby zmierza w nieskończoność.

BIBLIOGRAFIA

[1] Anderson T.: *Wwiedienije w mnogamiernyj statisticzeskij analiz.* GIFML Moskwa 1963 r.

[2] Biggeri L.: *On the distinction between sampling and non sampling errors and use of mathematical models for their evaluation.* Proceedings of the 40th session of the International Statistical Institute, Warszawa 1975.

[3] Cassel C.M., Särndal C.E., Wretman J.H.: *Foundation of inference in survey sampling.* John Wiley and Sons, Inc., New York — London — Sydney — Toronto 1977

[4] Guilford I.P.: *Podstawowe metody statystyczne w psychologii i pedagogice.* PWN Warszawa 1960

[5] Kendall M.G., Buckland W.R.: *Słownik terminów statystycznych,* PWE Warszawa 1975

[6] Kolonko J., Wywiał J.: *Metody analizy rzetelności danych, w materiałach na konferencję: Organizacja i badania rynku artykułów wyposażenia mieszkań.* PTE Katowice 1980
[7] Konijn H.S.: *Statistical theory of sample survey design and analysis.* North — Holland Publishing Campany — Amsterdam — London, American Elsevier Publishing Company, Inc. — New York 1973
[8] Little R.J.A.: *Models for nonresponce in sample survey.* Journal of the American Statistical Association 1982, s. 237—250
[9] Lutyńska K.: *Wywiad kwestionariuszowy.* PAN—Ossolineum Wrocław — Warszawa — Kraków — Gdańsk — Łódź 1984
[10] Morgenstern O.: *On the accurancy of economic observations.* Princentor University Press 1963
[11] Rao C.R.: *Modele liniowe statystyki matematycznej.* PWN Warszawa 1982
[12] Theil H.: *Zasady ekonometrii.* PWN Warszawa 1979
[13] Zarkovich S.S.: *Quality of statistical data.* FAO Roma 1966
[14] Zasępa R.: *Szacowanie niedokładności wyników badań statystycznych.* Biblioteka Wiadomości Statystycznych, tom 7. GUS Warszawa 1972

Możliwość zastosowania badań reprezentacyjnych do oceny stopnia rozprzestrzenienia się alkoholizmu w Polsce

Zofia Mielecka-Kubień
Akademia Ekonomiczna w Katowicach

Z doświadczeń wielu krajów wynika, iż koszty ponoszone przez społeczeństwo, w związku z nadużywaniem alkoholu przez niektórych jego członków są bardzo wysokie. Przykładowo koszty te oszacowane w 1983 r. dla Anglii i Walii wyrażają się kwotą 1615,5 mln funtów[1]), a podobnie wyliczone dla St. Zjedn. Ameryki w 1971 r. wynosiły około 31424 mln dolarów[2]). Mimo że w Polsce nie było jak dotąd tego rodzaju badań, można przypuścić, że kwota ta jest znaczna.

Jednym z podstawowych warunków umożliwiających ocenę, a także przewidywanie ekonomicznych skutków alkoholizmu, jak również kontrolę nad ilością i jakością spożywanych napojów alkoholowych są studia nad wzorami konsumpcji napojów alkoholowych oraz stopniem rozprzestrzenienia się alkoholizmu. Stopień ten można w początkowym etapie badań wyrazić poprzez liczbę osób pijących nadmiernie.

[1]) Por. [3] s. 33.
[2]) [2].

Szczególne znaczenie posiada znajomość rozkładu spożycia alkoholu według ilości konsumowanej w ciągu roku (z uwzględnieniem, w miarę możności, demograficznych i społeczno-zawodowych cech konsumentów napojów alkoholowych), ponieważ na tej podstawie można m.in. oszacować liczbę osób pijących szczególnie dużo.

Najpopularniejszą metodą badania wzorów spożywania napojów alkoholowych są badania ankietowe, które wykazują jednak zazwyczaj duży stopień niedoszacowania konsumpcji alkoholu, w porównaniu z danymi o sprzedaży napojów alkoholowych. Oszacowana, na podstawie badań ankietowych, ilość spożywanego alkoholu waha się najczęściej w granicach 40—60% ilości sprzedanej[3]); tak duże niedoszacowanie konsumpcji stawia pod znakiem zapytania możliwość prawidłowego wnioskowania o parametrach poszukiwanych rozkładów spożycia napojów alkoholowych w populacji, na podstawie wyników standardowych badań ankietowych.

W Polsce badania ankietowe nad spożyciem alkoholu są prowadzone w nieregularnych odstępach czasu, począwszy od 1961 r.[4]). Badanie z 1980 r. pozwoliło określić przykładowo[5]) 42,9% wypitej wódki, 37,1% wina oraz 76,2% piwa, co stanowi łącznie 46,8% napojów alkoholowych sprzedanych w Polsce w roku badania; rząd dokładności jest tu podobny do powszechnie uzyskiwanego w standardowych badaniach ankietowych, dotyczących spożycia napojów alkoholowych. Można zauważyć, że oceny wielkości spożycia napojów alkoholowych, uzyskane w badaniach budżetów gospodarstw domowych, cechuje jeszcze większy stopień niedoszacowania; na przykład spożycie napojów alkoholowych z dochodów osobistych ludności na 1 mieszkańca w 1980 r. wynosiło 5484 zł, oszacowane spożycie na podstawie budżetów gospodarstw domowych jest około 5 razy niższe; w 1986 r. różnica ta była jeszcze większa — spożycie z dochodów osobistych wynosiło 23906 zł na 1 mieszkańca, natomiast przeciętne roczne wydatki na napoje alkoholowe na 1 osobę w badanych gospodarstwach domowych wynosiły około 4207,8 zł[6]), czyli 5,68 razy mniej.

Do najważniejszych powodów istnienia różnic, pomiędzy uzyskiwaną w badaniach ankietowych wielkością spożycia napojów alkoholowych a danymi o sprzedaży tych napojów, zaliczyć można odmowy odpowiedzi oraz świadome zatajanie przez respondentów części rzeczywistego spożycia. Dużą rolę odgrywa również fakt, iż respondenci zapominają o różnych okazjach spożywania napojów alkoholowych, a w badaniach dodatkowo pomijane są najmłodsze grupy konsumentów. Na wyniki badań prowadzo-

---

[3]) [12].

[4]) Badania te prowadzone były przez Ośrodek Badania Opinii Publicznej później Ośrodek Badania Opinii Publicznej i Studiów Programowych przy Komitecie do Spraw Radia i Telewizji w Warszawie w latach 1961, 1962, 1968, 1980.

[5]) Por. [8] s. 9. W rzeczywistości rozbieżność ta jest jeszcze większa ponieważ dane o sprzedaży nie uwzględniają napojów alkoholowych produkowanych domowym sposobem oraz salda indywidualnego handlu zagranicznego w tej dziedzinie.

[6]) Obliczenia własne na podstawie Roczników Statystycznych GUS 1981 i 1987.

nych w naszym kraju niekorzystny wpływ miał sposób doboru respondentów do próby. Był to dobór celowy, polegający na tym[7]), że ankieterom określano cechy, jakimi powinni charakteryzować się respondenci, a zadaniem ankietera było znalezienie takich osób i uzyskanie od nich odpowiedzi na pytania ankiety. Jak wskazuje J. Jasiński [g] — „istnieje uzasadniona obawa..., że osoby intensywnie pijące miały zapewne mniejsze szanse dostania się do grupy badanej, niżby to wynikało z ich liczebności wśród ludności dorosłej, gdyż nie są to osoby, z którymi łatwo się skontaktować oraz uzyskać od nich odpowiedzi na zadane pytanie"[8]). Można dodać, iż ten sposób doboru próby mógł spowodować również inne, trudne do przewidzenia zniekształcenia wyników, czego można by uniknąć, dobierając odpowiednio liczną próbę w sposób losowy.

Problemy odmów odpowiedzi oraz uzyskiwania odpowiedzi rozmyślnie fałszywych, które występują w pewnym stopniu w każdym badaniu ankietowym, nabierają szczególnej wagi, jeśli w ankiecie występują pytania drażliwe, dotyczące spraw intymnych lub postaw społecznie nieakceptowanych. Udzielenie odpowiedzi staje się dla respondenta kłopotliwe, szczególnie jeśli odpowiedź stawia go w niekorzystnym świetle. Problem ten, oprócz aspektu praktycznego, ma również pewien wydźwięk moralny — prawo respondenta do nieujawniania swoich nie aprobowanych społecznie cech powinno być respektowane również w badaniu ankietowym. Pytania dotyczące spożywania napojów alkoholowych można zaliczyć do kategorii pytań drażliwych.

Zaproponowana w 1965 r. przez Stanley'a L. Warner'a[9]) metoda randomizacji (ulosowienia) odpowiedzi (Randomized Response Model) stwarza możliwość uniknięcia lub znacznego zmniejszenia występujących w badaniach ankietowych błędów, wynikających z dwu wyżej wymienionych powodów, zapewniając respondentowi całkowitą anonimowość. Najogólniej mówiąc, metoda ta polega na wprowadzeniu mechanizmu losowego, który powoduje, że osoba ankietowana odpowiada na zadane drażliwe pytanie z pewnym z góry określonym prawdopodobieństwem. Postępowanie takie nie ujawnia, w oczach badacza, sytuacji osoby ankietowanej w odniesieniu do rozważanego drażliwego problemu, pozwala jednak na oszacowanie parametrów rozkładu badanej zmiennej. W metodzie Warner'a anonimowość respondenta zagwarantowana jest przez samą procedurę badawczą, a nie osobę ankietera, jak to ma miejsce w zwykłym badaniu ankietowym.

Najprostsza i chronologicznie pierwsza wersja tej metody dotyczy szacowania wskaźnika struktury w populacji generalnej na podstawie jednej prostej próby losowej (losowanie indywidualne, nieograniczone, zwrotne).

---

[7]) [8] s. 10.
[8]) Ibid.
[9]) [15].

Załóżmy, że wylosowano próbę n — elementową z populacji, której członkowie należą wyłącznie do grupy A lub $B=\bar{A}$. Celem badania jest oszacowanie udziału osób należących do grupy A, przy czym przyznanie, że należy się do tej grupy jest dla respondenta kłopotliwe. W każdym wywiadzie osoba ankietowana wprawia w ruch mechanizm losowy, którego wskazania są niewidoczne dla osoby przeprowadzającej wywiad. W cytowanej pracy [15] zaproponowano użycie nakręconego bączka, którego wskazówka pokazuje literę A z prawdpopodobieństwem p, a literę B z prawdopodobieństwem $1-p$. Osoba ankietowana udziela odpowiedzi „tak" lub „nie" w zależności od tego, czy wskazówka pokazała grupę, do której dana osoba należy, ale nie ujawnia, czy była to grupa A czy B. Przy założeniu, że odpowiedzi respondentów były zgodne z prawdą, można metodą największej wiarygodności uzyskać estymator wskaźnika struktury. Oznaczenia:

Π_A — prawdopodobieństwo występowania w populacji osób należących do grupy A,

p — prawdopodobieństwo, że wskazówka bączka wskaże literę A,

$$X_i = \begin{cases} 1 \text{ jeśli } i\text{-ta wylosowana osoba odpowie „tak",} \\ 0 \text{ jeśli } i\text{-ta wylosowana osoba odpowie „nie",} \end{cases}$$

wtedy:

$$P\{X_i=1\}=\Pi_A \cdot p+(1-\Pi_A)(1-p)$$
$$P\{X_i=0\}=(1-\Pi_A)p+\Pi_A(1-p)$$

funkcję wiarygodności można więc zapisać:

$$L=[\Pi_A \cdot p+(1-\Pi_A)(1-p)]^{n_1} \cdot [(1-\Pi_A)p+\Pi_A(1-p)]^{n-n_1}$$

gdzie: n_1 oznacza liczbę osób w próbie, które odpowiedziały „tak", a $n-n_1$ liczbę osób, które odpowiedziały „nie".

Najlepszy, w sensie metody największej wiarygodności, estymator parametru Π_A, równy wartości maksimum funkcji L względem Π_A jest dany:

$$\hat{\Pi}_A=\frac{p-1}{2p-1}+\frac{n_1}{(2p-1)n}=\frac{1}{2p-1}\left\{p-1+\frac{n_1}{n}\right\};\ p\neq\frac{1}{2}$$

Jest to estymator nieobciążony, a jego wariancja wyraża się wzorem:

$$D^2(\hat{\Pi}_A)=\frac{1}{n}\left[\frac{1}{16(p-0{,}5)^2}-(\Pi_A-0{,}5)^2\right]$$

Wariancję powyższą można przedstawić w alternatywnej postaci jako sumę dwóch wariancji, z których pierwsza wyraża błąd losowy, jaki powstaje

w zwykłym badaniu ankietowym pod warunkiem, że respondenci udzielają wyłącznie prawdziwych odpowiedzi, a wystąpienie drugiej spowodowane jest zastosowaniem mechanizmu losowego:

$$D^2(\hat{\Pi}_A)=\frac{\Pi_A(1-\Pi_A)}{n}+\frac{p(1-p)}{n(2p-1)^2} \tag{1}$$

Można zauważyć, że dla $p=1$ powyższa procedura badawcza sprowadza się do standardowego badania, w którym respondent bezpośrednio odpowiada na pytanie, czy należy do grupy A.

Rozważając zagadnienie wyznaczenia wielkości próby, zapewniającej wymaganą precyzję szacunku, można spostrzec, iż wielkość próby zależy między innymi od parametru p. Jeżeli wartość p bliska jest 0 lub 1 próba może być mniejsza niż w przypadku, gdy p bliskie jest 0,5. Powodem tego jest, iż druga część składowa wyrażenia (1) jest szczególnie wysoka, gdy p bliskie jest 0,5 (a n niezbyt duże). Dlatego, w każdym przypadku należy rozważyć, czy zastosowanie metody Warner'a jest opłacalne z punktu widzenia dokładności szacunku, tj. czy korzyść, wynikająca z wyższego udziału odpowiedzi prawdziwych i mniejszej ilości braków odpowiedzi równoważy zmniejszoną dokładność szacunku, spowodowaną wystąpieniem dodatkowego błędu związanego z zastosowaniem w wywiadach mechanizmu losowego. Jeżeli przyjmiemy przykładowo, iż $\Pi_A=0{,}5$, a $p=0{,}75$ to wariancja estymatora $\hat{\Pi}_A D^2(\hat{\Pi}_A)=\frac{1}{n}$. Przy założeniu pożądanej dokładności szacunku parametru Π_A, na poziomie 0,05, można obliczyć, że próba, w której zastosujemy metodę randomizacji odpowiedzi powinna liczyć około 400 osób.

W zwykłym badaniu ankietowym (to znaczy dla $p=1$) wystarczyłaby natomiast próba 100-elementowa, pod warunkiem, że wszyscy respondenci mówiliby prawdę. Jeżeli jednak członkowie grupy A udzieliliby odpowiedzi odnośnie swojej przynależności do grupy A z prawdopodobieństwem T_A, a członkowie grupy B z prawdopodobieństwem T_B to, jak wykazał S. L. Warner, estymator:

$$\hat{\Pi}=\frac{\sum^{n} Y_i}{n}$$

gdzie: Y_i $\begin{cases}1, & \text{gdy } i\text{-ta osoba w próbie zgłasza przynależność do grupy } A,\\ 0, & \text{gdy } i\text{-ta osoba nie zgłasza przynależności do grupy } A,\end{cases}$

jest estymatorem obciążonym, a jego obciążenie wynosi:

$$B=\Pi[T_A+T_B-2]+[1-T_B]$$

Wariancja estymatora $\hat{\Pi}$ równa jest w tym przypadku:

$$D^2(\hat{\Pi}) = \frac{[\Pi \cdot T_A + (1-\Pi)(1-T_B)]\,[1-\Pi \cdot T_A - (1-\Pi)(1-T_B)]}{n}$$

Można dodać, że metoda Warner'a jest tym bardziej opłacalna, z punktu widzenia dokładności szacunku, im bardziej drażliwe jest pytanie A, tj. im większy byłby prawdopodobnie udział nieprawdziwych odpowiedzi w standardowym sposobie ankietyzacji. Metoda Warner'a została zmodyfikowana w 1967 r. przez Walt'a R. Simmons'a w ten sposób, iż w trakcie wywiadu respondent odpowiada na jedno z dwu pytań, z których jedno jest pytaniem drażliwym, a drugie pytaniem emocjonalnie obojętnym i nie związanym z pierwszym. U podstaw takiego postępowania leży przekonanie, że w takim przypadku respondent miał większy stopień pewności, że prowadzący badanie nie będzie mógł odgadnąć, drażliwej dla respondenta cechy A. Metoda ta znana jest w literaturze pod nazwą „Model randomizacji odpowiedzi z neutralnym pytaniem" (Unrelated Question Randomized Response Model). Załóżmy, iż respondent badany metodą randomizacji odpowiedzi udziela wyłącznie prawdziwych odpowiedzi. W metodzie Simmons'a, zamiast potwierdzenia lub zaprzeczenia jednemu z dwu stwierdzeń: „należę do grupy A", „nie należę do grupy A", co było charakterystyczne w metodzie Warner'a, respondent potwierdza lub zaprzecza jednemu z dwu wybranych, przy pomocy mechanizmu losowego, stwierdzeń następującej treści:

„należę do grupy A"
„należę do grupy Y"

z tym, że przyznanie przynależności do grupy Y nie jest dla respondenta kłopotliwe. Metoda ta okazuje się zazwyczaj bardziej efektywna od metody Warner'a. Przykładowo drażliwe pytanie A brzmi: „czy pije Pan codziennie alkohol?", a pytanie neutralne: „czy urodził się Pan w maju?". Zakładamy, że próba liczy 200 osób, a element losowości do odpowiedzi respondentów jest wprowadzony za pomocą rzutu monetą — jeśli respondent wyrzuci orła odpowiada na drażliwe pytanie A, gdy wyrzuci reszkę — odpowiada na neutralne pytanie Y. Przeprowadzający badanie nie wie, czy dana osoba wyrzuciła orła czy reszkę, nie wie więc, na które pytanie otrzymała odpowiedź.

Załóżmy dalej, że w ten sposób uzyskano 35 odpowiedzi „tak", a 165 „nie". Ponieważ prawdopodobieństwo wyrzucenia orła czy reszki jest 0,5, można przypuścić, że około 100 osób odpowiadało na pytanie A i 100 na pytanie Y. Jeśli przyjmiemy, że mniej więcej $200 \cdot \frac{1}{12} \approx 17$ osób urodziło się w maju to można wyliczyć, że 17 z 35 twierdzących odpowiedzi dotyczyło pytania Y, a 18 pytania A. Jako, że na pytanie A odpowiadało około 100 osób, oszacowany udział pijących codziennie alkohol $\hat{\Pi}_A \approx 18\%$. Wprowadzenie drugiego pytania Y powoduje konieczność oszacowania, obok Π_A,

również parametru Π_y, chociaż istnieje również możliwość takiego zdefiniowania grupy Y, że Π_y jest z góry znane (może tak być w przypadku osób urodzonych w maju). Jeśli Π_y jest nie znane muszą być wylosowane niezależnie od siebie dwie próby, liczące odpowiednio n_1 i n_2 elementów. W każdej z nich przeprowadza się wywiad metodą randomizacji odpowiedzi Simmons'a[10]) z tym, że mechanizm losowy w każdej z prób zapewnia inne prawdopodobieństwo wyboru pytania A.

Niech dla $i=1,2$ — p_i oznacza prawdopodobieństwo wyboru pytania A w i-tej próbie, przy czym $p_1 \neq p_2$,
— $1-p_i$ oznacza prawdopodobieństwo wyboru pytania Y w i-tej próbie,
— λ_i prawdopodobieństwo otrzymania odpowiedzi „tak",
— $\hat{\lambda}_i=\frac{n_i}{n_i}$ udział odpowiedzi „tak" w próbie i-tej,

można teraz zapisać:

$$\lambda_i=p_i \cdot \Pi_A+(1-p_i)\Pi_y=p_i \cdot \Pi_A+\Pi_y-p_i\Pi_y=p_i(\Pi_A-\Pi_y)+\Pi_y \qquad (2)$$

Podobnie budujemy funkcję największej wiarygodności, a rozwiązując ją względem Π_A oraz Π_y otrzymujemy:

$$\Pi_{Au}=\frac{\lambda_1(1-p_2)-\lambda_2(1-p_1)}{p_1-p_2}$$

$$\Pi_{yu}=\frac{p_2\lambda_1-p_1\lambda_2}{p_2-p_1}$$

Podstawiając $\hat{\lambda}_i$ w miejsce λ_i uzyskujemy estymator parametru Π_A:

$$\hat{\Pi}_{Au}=\frac{\hat{\lambda}_1(1-p_2)-\hat{\lambda}_2(1-p_1)}{p_1-p_2} \qquad (3)$$

Estymator ten jest estymatorem nieobciążonym, a jego wariancja równa jest:

$$D^2(\hat{\Pi}_{Au})=\frac{1}{(p_1-p_2)^2}\left\{\frac{\lambda_1(1-\lambda_1)(1-p_2)^2}{n_1}+\frac{\lambda_2(1-\lambda_2)(1-p_1)^2}{n_2}\right\}$$

Należy zaznaczyć, że w obu opisanych metodach zarówno punktowa, jak i przedziałowa[11]) ocena parametru Π_A może znaleźć się poza przedziałem (0,1). W modelu Simmons'a niebezpieczeństwo takie powstaje, jeśli przyjmie

[10]) [7].

[11]) Ponieważ $\hat{\Pi}_A$ jest estymatorem otrzymanym metodą największej wiarygodności, a n zazwyczaj duże, można przyjąć, że $\hat{\Pi}_A$ ma rozkład normalny i dokonać estymacji przedziałowej parametru Π_A.

się bliskie sobie wartości p_1 i p_2 (mianownik wyrażenia (3) będzie wtedy małą liczbą, a całe wyrażenie większe od 1).

Aby uniknąć tego rodzaju sytuacji Greenberg, et al.,[12]) postulują przyjęcie, jako generalnej zasady wyboru, aby prawdopodobieństwa p_1 i p_2 były liczbami możliwie od siebie oddalonymi w ramach przedziału (0,1), aż do granicy, która może wzbudzić nieufność respondenta. Prostym sposobem spełnienia tego postulatu jest taki wybór p_1 i p_2, aby $p_1+p_2=1$, aczkolwiek tak duży wpływ subiektywnego przecież wyboru prawdopodobieństw p_1 i p_2 na wartość estymatora $\hat{\Pi}_{Au}$ jest sprawą trochę niepokojącą. Greenberg, et al., zalecają więc następujący, korzystny również z punktu widzenia minimalizacji wariancji estymatora $\hat{\Pi}_{Au}$, sposób postępowania[13]), zgodnie z propozycją Warner'a wartość p_1 należy ustalić tak daleko od 0,5 w przedziale (0,1), jak to tylko jest w danym badaniu możliwe, ze względu na możliwość wzbudzenia podejrzliwości respondenta; najczęściej więc p_1 będzie równe 0,80 lub $0{,}20\pm0{,}10$. Następnie wyznacza się wartość p_2 w taki sposób, aby zminimalizować wariancję estymatora $\hat{\Pi}_{Au}$, przy zadanym Π_y, n_1, n_2 oraz warunku, że p_2 nie może leżeć bliżej końców przedziału (0,1) niż p_1. W tej sytuacji $D^2(\hat{\Pi}_{Au})$ jest funkcją p_2, a pierwsza pochodna tego wyrażenia, ze względu na p_2, wynosi:

$$\frac{d\,D^2(\hat{\Pi}_{Au})}{d\,p_2}=\frac{1-p_1}{(p_1-p_2)^3}\left\{\frac{2\lambda_1(1-\lambda_1)(1-p_2)}{n_1}+\frac{(\lambda_1+\lambda_2-2\lambda_1\lambda_2)(1-p_1)}{n_2}\right\} \quad (4)$$

Przy założeniu, że p_1 różne jest od 0 i 1 analiza wyrażenia (4) pozwala na stwierdzenie, że aby wartość jego była możliwie mała p_1 powinno jak najbardziej różnić się od p_2. Tak więc z punktu widzenia minimalizacji wariancji estymatora praktyczna reguła wyboru p_1 i p_2 może opierać się na podanym wyżej postulacie, aby $p_1+p_2=1$.

Reguła wyboru prawdopodobieństwa p_2 została zmodyfikowana przez J. J. A. Moors'a[14]). Proponuje on wybór $p_2=0$, co oznacza, że respondenci drugiej próby odpowiadają wyłącznie na neutralne pytanie Y. Metodę randomizacji odpowiedzi stosuje się więc tylko w jednej z prób, a druga służy do oszacowania wartości pomocniczego parametru Π_y. Jeżeli pytanie Y jest tak dobrane, że Π_y jest z góry znane, druga próba staje się w ogóle niepotrzebna. Zdaniem J. J. A. Moors'a takie postępowanie zapewnia szereg korzyści:

1) zmniejsza się wariancja estymatora,
2) zmniejsza się koszt badania, ponieważ tylko jedna z prób badana jest droższą metodą randomizacji odpowiedzi,
3) ponieważ wybór pytania Y zależy od badacza, można je tak dobrać, aby Π_y mogło być szacowane tańszymi metodami niż wywiad bezpośredni.

---

[12]) [5].
[13]) Ibid.
[14]) [10].

J. J. A. Moors udowodnił ponadto[15]), że przy omówionym poniżej, optymalnym wyborze n_1 i n_2, przy ustalonym n, optymalny, ze względu na postulat minimalizacji wariancji estymatora Π_{Au}, jest taki wybór p_1 i p_2, aby p_1 było jak najwyższe, a $p_2=0$. Odnośnie optymalnej alokacji n_1 i n_2, przy ustalonej całkowitej liczebności próby n, można stwierdzić, że wybór ten dokonany powinien być w taki sposób, aby suma obu umieszczonych w nawiasie składników wyrażenia (4) była minimalna (przy założeniu, że liczniki tych składników nie są sobie równe, co w praktyce na ogół bywa spełnione); prowadzi to do następującej proporcji[16]):

$$\frac{n_1}{n_2}=\sqrt{\frac{\lambda_1(1-\lambda_1)(1-p_2)^2}{\lambda_2(1-\lambda_2)(1-p_1)^2}} \tag{5}$$

Wartości parametrów λ_1 i λ_2 nie są oczywiście znane przed rozpoczęciem badania, powstaje więc problem, występujący na ogół w zagadnieniach optymalnej alokacji próby, jak w praktyce ustalić proporcję $n_1:n_2$. Można zauważyć, że ponieważ λ_1 i λ_2 są średnimi ważonymi parametrów Π_A i Π_y wystarczy, w celu dokonania optymalnej alokacji, znać choćby przybliżone oceny tych parametrów (na przykład z wcześniejszych badań). Warto również zauważyć, że wzór (5) w taki sposób rozdziela liczebność n między obie próby, że większa liczba elementów przypada tej próbie, w której wyższe jest prawdopodobieństwo wyboru pytania A. Jak wykazano w [5] optymalna alokacja próby znacznie poprawia dobroć oszacowania parametru Π_A metodą Simmons'a w porównaniu z metodą Warner'a.

Wielkie znaczenie dla skuteczności metody randomizacji odpowiedzi Simmons'a posiada wybór neutralnego pytania Y oraz wartość parametru Π_y. Optymalny, z punktu widzenia minimalizacji wariancji estymatora $\hat{\Pi}_{AU}$, wybór wartości Π_y został zaproponowany przez J. J. A.Moors'a. Minimalizując wariancję $D^2(\hat{\Pi}_{AU})$ względem optymalnie wybranych dla ustalonego n wartości n_1 i n_2 otrzymał on następujące wyrażenie[17]):

$$D^2(\hat{\Pi}_{AU})^*=\left\{\frac{(1-p_1)\sqrt{\lambda_2(1-\lambda_2)}+(1-p_2)\sqrt{\lambda_1(1-\lambda_1)}}{(p_1-p_2)\sqrt{n}}\right\}^2 \tag{6}$$

W wyrażeniu (6) parametr Π_y zawarty jest w wartościach λ_1 i λ_2, co wynika ze wzoru (2). Ponieważ funkcja $\sqrt{\lambda_i(1-\lambda_i)}$ posiada maksimum dla $\lambda_i=0{,}5$ i jest symetryczna oraz malejąca po obu stronach tego punktu, λ_i powinno mieć wartość jak najbardziej odległą od 0,5. Biorąc pod uwagę wyrażenie (2) można stwierdzić, że z punktu widzenia uzyskania jak najmniejszej wariancji

[15]) Ibid.
[16]) [5].
[17]) [10].

estymatora $\hat{\Pi}_{AU}$, zmienna Y powinna być tak dobrana, aby Π_y było po tej samej stronie wartości 0,5 co przypuszczalna wartość Π_A, a wyrażenie $|\Pi_y - 0{,}5|$ jak największe. Jeżeli przypuszcza się, że $\Pi_A = 0{,}5$, nie ma znaczenia po której stronie punktu 0,5 będzie leżeć wyrażenie $|\Pi_A - 0{,}5|$, lecz jak poprzednio powinno być ono możliwie duże.

Przy wyborze zmiennej Y i parametru Π_y nie można jednak kierować się wyłącznie postulatem minimalizacji wariancji estymatora $\hat{\Pi}_{AU}$, wielokrotnie wskazywano bowiem, że zarówno bardzo wysokie, jak również bardzo niskie wartości Π_y mogą wzbudzić podejrzenie respondenta, że jego przynależność do grupy A może zostać odkryta, co stawia pod znakiem zapytania celowość stosowania metody randomizacji odpowiedzi. Jest to szczególnie niebezpieczne, gdy przypuszczalna wartość Π_A jest mała, a wybrana wartość Π_y bliska 0. W tym przypadku bowiem, respondent może obawiać się, że ponieważ odpowiedzi „tak" będzie w ogóle bardzo mało, niezależnie od tego czy będą to odpowiedzi na pytanie A czy Y, osoba która odpowiedziała „tak" będzie mocno podejrzana o przynależność do grupy A; można więc oczekiwać, że respondenci będą udzielali nieprawdziwych odpowiedzi. Jeśli natomiast, przy małym udziale w próbie osób należących do grupy A, Π_y będzie duże, członek grupy A chętniej przyzna się, że do niej należy, ponieważ odpowiedź „tak" nie postawi go w grupie podejrzanych o posiadanie cechy A.

Tak więc z punktu widzenia uzyskania zaufania respondenta korzystne są stosunkowo wysokie wartości parametru Π_y, co w niektórych przypadkach może być sprzeczne z wymogami osiągnięcia jak najmniejszej wariancji estymatora $\hat{\Pi}_{AU}$. Trzeba jednak pamiętać, że odpowiedź „tak" jest zawsze bardziej kłopotliwa dla respondenta i to tym bardziej, im bardziej sugeruje, że może on należeć do grupy A. Można nawet przypuszczać, że wartość warunkowego prawdopodobieństwa, zdefiniowanego w następujący sposób: P (respondent należy do grupy A) respondent odpowie „tak", co można oznaczyć $P(A/\text{tak})$ zostaje podświadomie użyta przez respondenta, jako podstawa do podjęcia decyzji o udzieleniu prawdziwej odpowiedzi[18]). Istotne zatem staje się spełnienie warunku:

$$P(A/\text{tak}) \leqslant \theta,$$

gdzie: θ jest z góry ustaloną niewielką liczbą z przedziału [0,1].

Równie ważną sprawą jest wybór i sposób sformułowania naturalnego pytania Y. Wydaje się, że niektóre z zastosowanych w konkretnych badaniach pytania neutralne, na przykład: „czy Pan jest leworęczny?", „czy Pan urodził się w maju?", „czy urodziła się Pani w Północnej Karolinie?", mogą wzbudzić podejrzenie respondenta, iż ankieter może wiedzieć lub odgadnąć przynależność respondenta do grupy Y, i w związku z tym domyślić się, czy respondent udzielił odpowiedzi na pytanie A, czy Y (np.

[18]) [9].

w trzecim pytaniu przynależność do grupy Y może zdradzać sposób wymowy osoby ankietowanej). W takim przypadku respondenci nie będą skłonni przyznać, iż należą do grupy A.

Problem ten można ominąć przy zastosowaniu tej wersji metody randomizacji odpowiedzi z neutralnym pytaniem, gdy wartość parametru Π_y jest z góry znana. Dla oszacowania parametru Π_A wystarcza wtedy jedna próba. Prawdopodobieństwo otrzymania odpowiedzi „tak" będzie w tym przypadku równe:

$$\lambda_1 = p_1 \cdot \Pi_A + \Pi_y(1-p_1)$$

Jeśli zastosujemy[19]), jak poprzednio, metodę największej wiarygodności estymatorem parametru Π_A będzie następujące wyrażenie:

$$\hat{\Pi}_{AU/\Pi y} = \frac{\hat{\lambda}_1 - \Pi_y(1-p_1)}{p_1}$$

We wzorze wariancji tego estymatora pominięta zostaje część, która była spowodowana koniecznością szacowania Π_y:

$$D^2(\hat{\Pi}_{AU/\Pi y}) = \frac{\lambda_1(1-\lambda_1)}{n \cdot p_1^2}$$

Jak wskazuje Greenberg, et al.,[20]) ta wersja metody Simmons'a zwiększa precyzję oszacowania parametru Π_A w porównaniu z wersją, gdy Π_y nie jest znane. Powodem tego jest, że obserwacje z próby nie są zużywane dla oszacowania pomocniczego parametru Π_y, lecz jedynie dla oszacowania Π_A, co ma szczególne znaczenie, gdy prawdopodobieństwo pojawienia się osoby należącej do grupy A jest małe. Korzyść ta występuje jednak tylko przy założeniu, że wartość Π_y znana jest bez błędu.

Niestety w wielu przypadkach założenie to nie jest spełnione — zdarza się, że badana populacja osób różni się nieco co do wartości parametru Π_y od szerszej populacji, dla której parametr ten był oszacowany (czy jest znany). Miało to miejsce na przykład w badaniu opisanym w [4], gdzie losowanie pytania A i Y odbywało się na podstawie ostatniej cyfry numeru telefonu respondenta. Okazało się, że w badanej populacji studentów, końcowe cyfry numerów telefonów nie były rozłożone w sposób losowy. W takiej sytuacji — jeśli Π_y^* oznaczać będzie zastosowaną ocenę Π_y, estymator $\Pi_{AU/\Pi y}$ przybierze formę:

$$\hat{\Pi}_{AU/\Pi_y^*} = \frac{\hat{\lambda}_1 - \Pi_y^*(1-p_1)}{p_1}, \quad \text{gdzie: } \Pi_y = \Pi_y^* - c$$

[19]) [5].
[20]) Ibid.

i będzie estymatorem obciążonym, a jego obciążenie będzie równe:

$$B=\frac{c(1-p_1)}{p_1}$$

Opisane wyżej metody pozwalają jedynie na oszacowanie udziału respondentów, posiadających cechę A. Częstym przypadkiem jest chęć poznania rozkładu zmiennej A w populacji, co wiąże się z koniecznością ustalenia — w jakim stopniu wylosowane do próby osoby posiadają drażliwą charakterystykę A.

W [1] A—L, A, Abul—Ela, et al., podają rozszerzenie modelu Warner'a na przypadek, gdy respondenci mogą należeć do jednej z 3 wzajemnie wykluczających się grup, z których jedna lub dwie posiadają w różnym stopniu cechę A, a następnie uogólnienie na przypadek $j>3$ grup, z których co najwyżej $j-1$ posiada tę cechę w różnym natężeniu. W modelu tym zakłada się, że każda osoba w badanej populacji należy do jednej z j wzajemnie wykluczających się grup, a celem badania jest oszacowanie udziału każdej z tych grup w populacji, co wymaga wylosowania $s=t-1$ niezależnych prób.

W modelu tym komplikuje się sprawa wyboru wartości prawdopodobieństw p_{ij}, a oprócz tego ze względu na zwiększoną liczbę parametrów rozmiary każdej z i prób muszą być znaczne. Z tego względu korzystniejsze wydaje się stosowanie innego modelu randomizacji odpowiedzi, pozwalającego również na otrzymanie danych ilościowych zaproponowanego przez B. G. Greenberga, et al., w [6].

Dodatkowym argumentem przemawiającym na korzyść tego ostatniego modelu jest, że jest on uogólnieniem efektywniejszego zazwyczaj niż model Warner'a modelu Simmons'a, a także to, że w modelu B. G. Greenberga, et al., respondent sam podaje liczbowe wartości posiadanej cechy A, a nie potwierdza lub zaprzecza swoją przynależność do z góry ustalonego przedziału wartości zmiennej A, co może się wiązać z pewnymi stratami informacji. Autorzy [21]) zastosowali swoją metodę do badania dwóch kwestii — jednej bardziej drażliwej, dotyczącej ilości przerywanych ciąż, a drugiej mniej drażliwej, dotyczącej dochodów gospodarstwa domowego. W pierwszym badaniu pytania A i Y brzmiały:

A. Ile razy w życiu usuwała Pani ciążę?

Y. Jeśli kobieta musi pracować w pełnym wymiarze godzin, ile dzieci powinna ona zdaniem Pani posiadać?

W drugim natomiast pytania były następujące:

A. Ile mniej więcej pieniędzy (w dolarach) głowa tego gospodarstwa domowego zarobiła w zeszłym roku?

---

[21]) [6].

Y. Ile mniej więcej pieniędzy (w dolarach) zarabia w ciągu roku Pana zdaniem głowa domu gospodarstwa domowego wielkości pańskiego gospodarstwa domowego?

Mechanizm losowy wprowadzony był w postaci niewielkiego pudełeczka zawierającego 35 czerwonych i 15 niebieskich kulek. Respondent potrząsał pudełkiem, aż jedna z kulek pojawiła się w widocznym dla niego (ale nie dla ankietera) okienku pudełka. Jeżeli pojawiła się czerwona kulka respondent odpowiadał na pytanie *A*, gdy niebieska na pytanie *Y*. Następnie respondent ponownie potrząsał pudełkiem, w celu przemieszania kulek, a dopiero potem oddawał je ankieterowi. W każdym badaniu użyto po dwie niezależne od siebie próby. Teoretyczny model, odpowiadający takiemu postępowaniu jest następujący (oznaczenia):

p_i — prawdopodobieństwo, że w i-tej próbie respondent będzie odpowiadał na drażliwe pytanie $A\,(i=1,2\ p_i \neq p_2)$,

$1-p_i$ — prawdopodobieństwo, że w i-tej próbie respondent będzie odpowiadał na pytanie $Y(i=1,2)$,

z_{ij} — odpowiedź j-tej osoby w i-tej próbie $(1=1,2\ j=1,2,\ldots,n_i)$,

$f(z)$ — funkcja rozkładu prawdopodobieństwa (lub funkcja gęstości) zmiennej A odpowiadającej pytaniu A, $E_f(Z)=\mu_A$,

$g(z)$ — funkcja rozkładu prawdopodobieństwa (lub funkcja gęstości) zmiennej Y odpowiadającej pytaniu Y, $Eg(Z)=\mu_y$,

$\hat{\mu}_A$ — estymator wartości przeciętnej zmiennej A,

$\hat{\mu}_y$ — estymator wartości przeciętnej zmiennej Y,

$\bar{z}_1$ — średnia wartość z odpowiedzi w próbie 1,

$\bar{z}_2$ — średnia wartość z odpowiedzi w próbie 2.

Funkcję rozkładu prawdopodobieństwa (funkcję gęstości) dla każdego elementu w próbie można zapisać:

Próba 1: $\psi_1(z_1)=p_1 \cdot f(z_1)+(1-p_1)\,g(z_1)$

Próba 2: $\psi_2(z_2)=p_2 \cdot f(z_2)+(1-p_2)\,g(z_2)$

stąd:

$$\mu_{z_1}=E(z_1)=p_1 \cdot \mu_A+(1-p_1)\,\mu_y$$

$$\mu_{z_2}=E(z_2)=p_2 \cdot \mu_A+(1-p_2)\,\mu_y$$

z czego wynika:

$$\mu_A=\frac{(1-p_2)\,\mu_{z_1}-(1-p_1)\mu_{z_2}}{p_1-p_2},$$

$$\mu_y=\frac{p_2\mu_{z_1}-p_1\mu_{z_2}}{p_2-p_1}$$

Podstawiając w miejsce μ_{z_1} i μ_{z_2} średnie z próby, otrzymujemy nieobciążone estymatory μ_A i μ_y:

$$\hat{\mu}_A = \frac{(1-p_2)\,\bar{z}_1 - (1-p_1)\bar{z}_2}{p_1 - p_2}$$

$$\bar{\mu}_y = \frac{p_2\bar{z}_1 - p_1\bar{z}_2}{p_2 - p_1}$$

Wariancje tych estymatorów są dane:

$$\left.\begin{aligned} D^2(\hat{\mu}_A) &= \frac{1}{(p_1-p_2)^2}\left[(1-p_2)^2 \cdot D^2(\bar{Z}_1) + (1-p_1)^2 \cdot D^2(\bar{Z}_2)\right] \\ D^2(\hat{\mu}_y) &= \frac{1}{(p_2-p_1)^2}\left[p_2^2 D^2(\bar{z}_1) + p_1^2 \cdot D^2(\bar{z}_2)\right] \end{aligned}\right\} \qquad (7)$$

gdzie:

$$D^2(\bar{Z}_i) = \frac{1}{n_i}\left[\sigma_y^2 + p_i(\sigma_A^2 - \sigma_y^2) + p_i(1-p_i)(\mu_A - \mu_y)^2\right] \qquad i = 1,2 \qquad (8)$$

Wariancje $D^2(z_i)$ można oszacować na podstawie próby: $\alpha^2(z_i) = \frac{S_i^2}{n_i}$ $(i = 1,2)$.

W opisanej metodzie zasady wyboru prawdopodobieństw p_1 i p_2 pozostają takie, jak w modelu Simmons'a. Podobnie, jak poprzednio, szczególne znaczenie posiada warunek, aby p_1 było możliwie odległe od p_2, gdyż w przeciwnym wypadku oceny parametrów μ_A i μ_y mogą być zawyżone, zresztą, jak wykonano w [6] wariancja estymatora $\hat{\mu}_A$ maleje w miarę wzrostu wyrażenia: $|p_2 - p_1|$.

Podstawową zasadą, pozwalającą na uzyskanie wiarygodnych danych ilościowych omawianą metodą, jest aby neutralne pytanie Y było tak sformułowane, że odpowiedź na nie wyrażać się będzie w tych samych jednostkach miary co odpowiedź na pytanie A. Podobny powinien być również rząd wielkości poszczególnych wariantów odpowiedzi, a także wartości przeciętne zmiennych A i Y. Ze wzorów (7) i (8) wynika zresztą, że przy ustalonych wartościach σ_A, σ_y, p_1, p_2, n_1, n_2 wariancje estymatorów $\hat{\mu}_A$ i $\hat{\mu}_y$ rosną w miarę wzrostu wyrażenia: $|\mu_A - \mu_y|$.

Z punktu widzenia minimalizacji wariancji estymatorów $\hat{\mu}_A$ i $\hat{\mu}_y$ korzystne jest również, aby wariancja zmiennej y (σ_y^2) była możliwie mała, to jest aby możliwe na pytanie Y odpowiedzi niewiele się od siebie różniły. Postulat ten jest jednak w pewnym stopniu sprzeczny z warunkami uzyskania zaufania respondentów co do ich całkowitej anonimowości, jeśli bowiem wariancja zmiennej A byłaby stosunkowo duża, a respondent podawałby wartości zmiennej A daleko odległe od μ_A, mógłby się na tej podstawie uważać za

podejrzanego o posiadane cechy A, co spowodować mogłoby nieprawdziwą odpowiedź. Z tego powodu wybór zmiennej Y, ze względu na jej wariancję, powinien być kompromisem między wymogiem statystycznym a wymogiem uzyskania zaufania respondenta; zaleca się aby σ_y^2 było co najmniej równe σ_A^2.

Reguła optymalnego rozdziału n pomiędzy dwie próby liczące n_1 i n_2 elementów jest w przypadku, gdy $\sigma_A^2 = \sigma_y^2$ lub gdy obie wariancje nie różnią się zbyt wiele od siebie, bardzo prosta:

$$\frac{n_1}{n_2} = \frac{p_1}{p_2}, \quad \text{gdy} \quad p_1 + p_2 = 1 \tag{9}$$

Jeśli równość $\sigma_y^2 = \sigma_A^2$ nie jest spełniona nawet w przybliżeniu reguła ta ma bardziej skomplikowaną postać[22]), ponieważ jednak bliskie sobie wariancje zmiennych A i Y są warunkiem dobrego wyboru zmiennej (i pytania) Y, przytoczoną regułę (9) można uznać za wystarczającą dla celów praktycznych. Można wyobrazić sobie przypadki, gdy μ_y oraz σ_y^2 są z góry znane; nie ma wtedy potrzeby stosowania dwóch prób. Estymator parametru μ_A i jego wariancja są wtedy równe:

$$\hat{\mu}_{A/\mu y} = [\bar{z} - (1-p)\,\mu_y$$

$$D^2(\hat{\mu}_{A/\mu y}) = \frac{D^2(\bar{z})}{p^2} = \frac{D^2(z)}{np^2}$$

Znajomość parametrów rozkładu zmiennej Y pozwala jak widać na znaczne zmniejszenie wariancji estymatora $\hat{\mu}_A$.

Metoda randomizacji odpowiedzi, w różnych swoich odmianach, była wielokrotnie stosowana w badaniach, dotyczących kwestii drażliwych[23]). We wspomnianych badaniach uzyskiwano z jednej strony bardzo mały procent braków odpowiedzi, co świadczy o zrozumieniu i akceptacji zasad randomizacji odpowiedzi przez respondentów, a z drugiej istotnie wyższe na ogół niż w zwykłych badaniach ankietowych oceny parametrów drażliwej zmiennej A.

W [14] opisano badanie ankietowe, dotyczące spożycia napojów alkoholowych z zastosowaniem metody randomizacji odpowiedzi, w którym kolejno zastosowano wersję Simmons'a, dla oszacowania udziału osób pijących napoje alkoholowe oraz wersję Greenberg'a pozwalającą uzyskać ilościowe dane o tym spożyciu.

W pierwszym przypadku pytania A i Y brzmiały:

A. Czy w ostatnim tygodniu pił (a) Pan (i) piwo, wino lub inny napój alkoholowy, przynajmniej jeden raz w ciągu dnia?

---

[22]) Ibid.

[23]) Por. na przykład [4], [6], [11], [14].

Y. Czy w ostatnim tygodniu pił (a) Pan (i) mleko przynajmniej jeden raz w ciągu dnia?

W drugim natomiast:

A_1 W ciągu ilu dni w tygodniu wypił (a) Pan (i) przynajmniej jeden kieliszek alkoholu w ciągu dnia?

A_2 W tych dniach, w których pił (a) Pan (i) alkohol, ile mniej więcej kieliszków wypił (a) Pan (i) w sumie?

Y_1 W ciągu ilu dni w tygodniu, Pana (i) zdaniem, przeciętna osoba w Pana (i) wieku wypija przynajmniej 1 kieliszek alkoholu (piwa, wina lub innego?)

Y_2 W tych dniach, w których przeciętna osoba w Pana (i) wieku pije choć 1 kieliszek alkoholu, ile łącznie kieliszków wypija ona Pana (i) zdaniem?

Randomizacja odpowiedzi respondentów zapewniona była w następujący sposób: każdy respondent powinien pomyśleć liczbę naturalną z przedziału [1,6] nie informując o jej wartości ankietera, a następnie rzucić kostką do gry; jeżeli wybrana przez niego liczba równa była wyrzuconej na kostce odpowiadał na pytania, dotyczące swojego spożycia alkoholu (A). W badaniu tym lepsze wyniki uzyskano stosując metodę drugą (dla danych ilościowych), przy okazji ujawniły się ciekawe różnice w niedoszacowaniu spożycia alkoholu (w porównaniu do zwykłych badań ankietowych prowadzonych poprzednio w tej samej populacji) w różnych grupach respondentów[24]).

Porównanie ocen średniej liczby wypijanych w tygodniu kieliszków alkoholu przy zastosowaniu metody randomizacji odpowiedzi oraz zwykłego badania ankietowego według płci, wieku i wykształcenia

Grupa respondentów	Liczba kieliszków w tygodniu, metoda randomizacji odpowiedzi	Liczba kieliszków w tygodniu, zwykłe badanie ankietowe
Cała próba	8,7 (±1,3)	3,9 (±1,4)
Płeć		
mężczyźni	10,6 (±1,9)	7,2 (±3,2)
kobiety	5,9 (±1,7)	1,9 (±0,8)
Wiek		
⩽ 52 lata	11,7 (±2,5)	6,6 (±2,9)
> 52 lata	6,6 (±1,3)	2,1 (±1,0)
Wykształcenie		
⩽ 12 lat nauki	6,5 (±2,2)	4,5 (±2,8)
> 12 lat nauki	10,3 (±1,6)	4,2 (±1,6)

U w a g a. W nawiasach podano wartości odchylenia standardowego.

[24]) [14] s. 747.

Statystycznie istotne były różnice ocen uzyskanych obiema metodami w całej populacji, a także dla kobiet, osób starszych niż 52 lata życia oraz tych, których nauka trwała więcej niż 12 lat, co wskazuje, że w tych grupach konsumentów alkoholu występuje znacznie silniejsza tendencja do zaniżania swojego spożycia niż w pozostałych.

Rozważając możliwość zastosowania opisanych wyżej metod do badania spożycia napojów alkoholowych w Polsce należy zwrócić uwagę , że pytania dotyczące spożycia napojów alkoholowych muszą być sformułowane w nieco inny sposób, inny jest bowiem w naszym kraju model spożywania alkoholu. Spożycie alkoholu w Polsce ma głównie charakter okazjonalny — zazwyczaj wypija się duże ilości alkoholu przy małej częstotliwości spożycia. Z tego powodu pytanie o ilość wypijanego alkoholu nie może odnosić się do jednego tygodnia, czy ostatnich 24 godzin jak w krajach, gdzie spożywa się jednorazowo mniejsze ilości alkoholu, ale za to częściej. Można więc pytać o częstość i ilość spożycia alkoholu w dłuższym okresie — w ciągu miesiąca, kwartału, roku co powoduje jednak negatywne dla badania konsekwencje; z jednej strony większe są szanse, że respondent nie przypomni sobie wszystkich okazji spożywania alkoholu w tak długim czasie, a z drugiej wymaga się od respondenta podania wiarygodnego wyniku skomplikowanych dosyć obliczeń.

W polskiej praktyce badań ankietowych nad spożyciem napojów alkoholowych pytanie, dotyczące spożycia brzmiało[25]): „Kiedy ostatnim razem przed naszą rozmową pił(a) Pan(i) następujące napoje ...?" Sposób sformułowania tego pytania oparto na wzorach fińskich[26]), (model spożywania napojów alkoholowych w Finlandii jest podobny do polskiego), przyjmując założenie, że okres, który dzieli kolejne okazje spożywania alkoholu, jest średnio dwa razy dłuższy, niż czas dzielący ostatnią okazję od chwili przeprowadzenia ankiety oraz, że ilość wypitego wówczas alkoholu odpowiada ilości „zwykle" wypijanej. Taki sposób formułowania pytania ma z kolei tę wadę, że „ostatnia okazja" może dość znacznie odbiegać od przeciętnego sposobu picia respondenta, może więc zostać on zaliczony do niewłaściwej klasy konsumentów. Szczególnym problemem w ankietowych badaniach nad spożyciem napojów alkoholowych jest grupa osób najwięcej pijących. Jest to grupa, w której trudno na ogół stosować metody wywiadu bezpośredniego włącznie z metodą randomizacji odpowiedzi, wymagającą przecież dość wysokiego poziomu intelektualnego respondenta. Z drugiej strony grupa ta przysparza najwięcej kłopotów i kosztów pozostałym członkom społeczeństwa; trzeba też zdawać sobie sprawę, że niezbyt liczna grupa osób najwięcej pijących spożywa w Polsce ponad połowę wypijanego alkoholu[27]), dlatego też warto podjąć próbę, by grupę tę w jakiś sposób uwzględnić w badaniu reprezentacyjnym.

---

[25]) [8].
[26]) [13].
[27]) [8].

Można zaproponować, stosowaną już do badania spożycia napojów alkoholowych, w grupie najwięcej pijących metodę informatorów[28]) (Informant Method). W metodzie tej wykorzystuje się inny, niż w metodzie randomizacji odpowiedzi sposób na uzyskanie informacji od respondenta, bez ujawniania jego osobistej sytuacji. Metoda informatorów polega na wyborze osób lub grup osób, zwanych informatorami, które udzielają informacji nie o sobie, lecz o dobrze sobie znanej grupie społecznej. Na przykład w Finlandii[29]) zastosowano powyższą metodę do badania spożycia wyrobów tytoniowych. Informatorem była pani domu, która podawała informacje o zakupach wyrobów tytoniowych w całym gospodarstwie domowym. Oszacowane w ten sposób spożycie wyrobów tytoniowych wynosiło 94,3% ilości sprzedanej, podczas gdy w zwykłych badaniach ankietowych jest to na ogół 60—70% sprzedaży.

W [12] zastosowano metodę informatorów do oszacowania spożycia napojów alkoholowych. Grupy informatorów liczyły 5 do 7 osób i w każdej grupie była wyznaczona osoba nią kierująca. Każda grupa dostawała kwestionariusz ankiety, który był wspólnie wypełniany po uzgodnieniu jednomyślnej odpowiedzi wszystkich członków grupy — dyskusje te trwały średnio 2,5 godziny.

Wybór informatorów powinien być reprezentatywny dla badanej populacji, w omawianym badaniu byli oni wylosowani z uwzględnieniem podziału geograficznego i grupy zawodowej. W porównaniu do zwykłego badania ankietowego wyniki uzyskane w [12] metodą informatorów wskazują na następujące różnice:

1) uzyskano wyższy udział osób często spożywających alkohol i wyższy poziom spożycia napojów alkoholowych;
2) mniejsze zróżnicowanie spożycia w zależności od płci (istotnie wyższe były udziały kobiet).

Dla porównania podano tamże, że przeciętne roczne spożycie alkoholu na 1 dorosłego mieszkańca w badanym okręgu Toronto, uzyskane z różnych źródeł, wynosiło:

1) na podstawie danych o sprzedaży napojów alkoholowych (15 lat i więcej) — 10,23 l;
2) metodą informatorów (18 lat i więcej) — 8,26 l;
3) standardowe badanie ankietowe (18 lat i więcej) — 3,94 l.

Szczególnie korzystną stroną tej metody, w badaniach nad spożyciem napojów alkoholowych, jest możliwość uniknięcia bezpośredniego ankietowania osób najwięcej pijących, co stwarza ogromne trudności w zwykłych badaniach ankietowych, a jednocześnie uwzględnić takie osoby w próbie, co jest konieczne ze względu na reprezentatywność próby.

---

[28]) [12].
[29]) Ibid.

Reasumując, można więc zaproponować, mimo pewnych mankamentów tych metod, zastosowanie w warunkach polskich metody Simmons'a do oszacowania udziału osób wypijających więcej niż pewną ustaloną ilość alkoholu lub metod opisanych w [1,6] dla poznania rozkładu spożycia napojów alkoholowych i oszacowania parametrów tego rozkładu. Badana próba powinna być próbą losową. Pożądanym, ze względu na zachowanie reprezentatywności tej próby, uzupełnieniem mogłaby stać się metoda informatorów, którą można by zastosować w badaniu spożycia osób pijących najwięcej.

Każdą z opisanych wyżej metod można by również zastosować do oszacowania wielkości produkcji i spożycia wytwarzanych w warunkach domowych napojów alkoholowych, przede wszystkim bimbru, który stanowi znaczącą część spożycia. W [8] podano na przykład[30]), że na 12,2 l wypitej w ciągu roku wódki 4,6 l przypadało na bimber.

Można dodać, że nawet jeśli proponowane wyżej metody okażą się tylko w pewnym stopniu efektywniejsze niż metody ankietyzacji spożycia napojów alkoholowych dotychczas stosowane w naszym kraju, można by drogą porównań, uzyskać na ich podstawie pewne dane o rozkładzie niedoszacowania deklarowanej konsumpcji alkoholu, co byłoby w badaniach problemów alkoholizmu cenną informacją.

---

[30]) s. 42.

LITERATURA

[1] A. L. A. Abul-Ela, B. G. Greenberg, D. G. Horvitz: *A. Multi-proportions Randomized Response Model,* Journal of the American Statistical Association, 1967, s. 990—1008.

[2] R. E. Berry; *Estimating the Economic Costs of Alcohol Abuse,* The New England Journal of Medicine, 9/1976 s. 620—621.

[3] R. Mc Donnell, A. Maynard: *The Costs of Alcohol Misuse,* British Journal of Addiction, 1985, s. 27—35.

[4] M. S. Goodstadt, V. Gruson: *The Randomized Response Technigue, a Test on Drug Use,* Journal of the American Statistical Association, 32/1975.

[5] B. G. Greenberg, Abdel-Latif A. Abul-Ela, W. R. Simmons, D. G. Horvitz: *The Unrelated Question Randomized Response Models. Theoretical Framework,* Journal of the American Statistical Association, 1969, s. 520—539.

[6] B. G. Greenberg, R. R. Kuebler, J. R. Abernathy, D. G. Horvitz: *Application of the Randomized Response Technigue on Obtaining Quantitive Data,* Journal of the American Statistical Association, 1971, s. 243—250.

[7] D. G. Horvitz, B. V. Shah, W. R. Simmons: *The Unrelated Question Randomized Response Model, Social Statisties Proceedings of the American Statistical Association,* 1967, s. 65—72.

[8] J. Jasiński: *Badania ankietowe nad spożyciem alkoholu w Polsce w 1980 roku,* Społeczny Komitet Przeciwalkoholowy, W-wa 1984.

[9] J. Lanke: *On the Choice of the Unrelated Question in Simmons' Version of Randomized Response*, Journal of the American Statistical Association 70/1975.
[10] J. J. A. Moors: *Optimization of the Unrelated Question Randomized Response Model*, Journal of the American Statistical Association, 1971, s. 627—629.
[11] J. M. Shimizu, G. S. Banham: *Randomized Response Technigue in a National Survey*, Journal of the American Statistical Association, 61/1978.
[12] R. S. Smart, C. B. Liban: *Alcohol Consumtion as Estimated by the Informant Method; A. Household Survey and Sales Data*, Journal of Studies on Alcohol, 43/1982.
[13] A. Święcicki: *Wyniki badań nad spożyciem alkoholu w Polsce*, OBOP, W-wa 1962.
[14] B. J. Volicer, L. Volicer: *Randomized Response Technigue of Estimating Alcohol Use and Noncompliance in Hypertensives*, Journal of Studies on Alcohol, 43/1982.
[15] S. L. Warner: *Randomized Response: A Survey Technigue for Eliminating Evasive Answer Bias*, Journal of the American Statistical Association, 1965, s. 63—69.

Błędy badania reprezentacyjnego

dr hab. Henryk Hillar
Uniwersytet Warszawski

Zagadnienie błędów badania reprezentacyjnego jest zagadnieniem trudnym. Nie łatwo się w tych sprawach zorientować. Literatura przedmiotu — nie zawsze pomaga. Jest ona bowiem albo wycinkowa, albo bardzo rozwlekła; pojęć wprowadzanych jest wiele, najczęściej nie są to pojęcia rozłączone w swej treści, częściowo się nakładają. Stwarza to dodatkowe trudności w oznaczeniach, których rozrzut jest znaczny; wiele wprowadza się rodzajów klasyfikacji. Często nie można przejść myślami z jednej pracy do drugiej. Więcej się mówi i to opisowo o źródłach błędów, uczula na nie i na ich miejsca powstawania[1]) aniżeli podaje się sposoby uwzględniania ich w wynikach liczbowych. Dlatego pragnę te zagadnienia usystematyzować dla siebie i dla dydaktyki, którą prowadziłem do czasu zlikwidowania tak ważnego kierunku, jak ekonometria na Uniwersytecie Warszawskim; jeszcze jeden nie udany eksperyment, bowiem obecnie reaktywuje się ten kierunek i być może przywróci się jako kursowy przedmiot wykłady z metody reprezentacyjnej[2]).

Zagadnienie błędów badania reprezentacyjnego jest dobrym przykładem, ilustrującym pojęcie rozważań akademickich, użyte bez ujemnego zabar-

[1]) W naszej literaturze mamy już bardzo dobrą książkę, traktującą o tych zagadnieniach w sposób zwarty i wyczerpujący. Jest to praca Jana Kordosa.
[2]) Przypadek Uniwersytetu Warszawskiego może być ilustracją, jak konieczne jest istnienie Ministerstwa Edukacji Narodowej, które wprowadzając ramowy program studiów, mogłoby zapobiec niektórym anomaliom.

wienia. Również zagadnienie błędów badania reprezentacjnego jest istotnym wskazaniem na trudności, na jakie napotyka konstruowana od lat sześćdziesiątych uogólniona teoria metody reprezentacyjnej. Wyniki badań reprezentacyjnych są publikowane. Są one dyskutowane z różnych względów; często są do ich wartości liczbowych zgłaszane zastrzeżenia. Dlatego należy podać dokładność szacunków oraz wskazać na źródła błędów[3]).

Pragnąc poznać błędy badania reprezentacyjnego[4]), którego wyniki zostały podane w publikacji, trzeba prześledzić miejsca, w których mogły one powstać. W tym celu trzeba prześledzić drogę od powzięcia zamiaru o badaniu, aż do opublikowania wyników. Tę drogę zwykło się dzielić na trzy etapy:

etap pierwszy: źródłem błędów może być zastosowana teoria metody reprezentacyjnej;

etap drugi: źródłem błędów przeważnie jest planowanie badania reprezentacyjnego, rozumianego jako proces od określenia celu badania, aż do opracowania wyników;

etap trzeci: źródłem błędów może być proces wydawniczy wyników z przeprowadzonego badania. Mogą to być błędy niezamierzone, jak i celowo wprowadzone.

Statystyka zwykło się obciążać odpowiedzialnością za błędy powstałe na etapie pierwszym i drugim. Wśród błędów badania reprezentacyjnego wyodrębnia się dwie grupy: błędy losowe (etap pierwszy) oraz błędy nielosowe (etap drugi). Z błędami losowymi zapoznajemy się od razu na początku studiów metody reprezentacyjnej (MR), a im bliżej dotykamy praktyki, pojawiają się błędy nielosowe. Ten podział jest już jednak historyczny i to pragnę podkreślić w tym referacie. Jednak w toku studiów błędów, w takiej kolejności należy te zagadnienia poznawać, bowiem ustaliły się już pojęcia oraz jest to niezbędne do rozumienia pomiaru błędów i wniosku końcowego, wskazującego na istotność tylko niektórych z określonych błędów.

BŁĘDY LOSOWE

Źródłem błędów losowych jest fakt, że pobrana próba losowa jest tylko jedną z możliwych realizacji. Istnieją dodatnie prawdopodobieństwa wylosowania różnych prób z tej samej populacji i przy tych samych warunkach badania. Istnieje rozkład prawdopodobieństwa próby (estymatora). Zauważmy, że:

$$E(T_n) \overset{\text{def.}}{=} \sum_{s \in S} (t_n)_s P[T_n = (t_n)_s]$$

[3]) Bardzo dobitnie wskazuje na to w swym referacie Jan Kordos.

[4]) Badanie reprezentacyjne jest pojęciem szerszym niż metoda reprezentacyjna, bowiem obejmuje również technikę pomiaru, opracowanie wyników.

gdzie: S — zbiór wszystkich możliwych prób o ustalonej liczebności n w ustalonym schemacie losowania, a przez realizację t_n estymatora T_n rozumiemy wartość dokładną, nie obarczoną błędami pomiarów.

Błędy losowe mierzymy wariancją, obciążeniem i odchyleniem wartości estymatora T_n od wartości określonego parametru. Obciążenie wynika z konstrukcji estymatora oraz teoretycznego rozkładu próby losowej[5]).

W początkach rozwoju MR nacisk był położony na kontrolę i badanie błędów losowych. To podejście jest właściwe statystyce matematycznej, cały bowiem nacisk był położony na likwidację obciążeń estymatorów oraz na konstrukcję schematów losowania, o coraz to mniejszych wariancjach. To podejście nazywam akademickim, ponieważ wyraźnie odbiega od tego co się obserwuje w praktyce. Przechodząc do praktyki, pojawiają się trudności i to nie tylko dlatego, że błędów losowych nie można wyodrębnić z mierzonych błędów, ale również powstają trudności pojęciowe. Tak dobrze znany ze statystyki matematycznej **parametr** staje się niezrozumiały i nieuchwytny. Niezrozumiały, bowiem niczego jeszcze nie mówiono o pomiarach rzeczywistych, w praktyce dokonywanych.

Nie znając parametru, jak zdefiniować i zmierzyć obciążoność estymatora? Bardzo wyraźnie te zagadnienia podkreśla L. Kish. Estymator, oznaczony przez nas jako T_n w praktyce nie występuje, dlatego tak trafnie określa się rozważania z estymatorem T_n (a więc wolnym od błędów pomiarów), jako tylko akademickie.

BŁĘDY NIELOSOWE

Błędami nielosowymi nazywamy te wszystkie, które powstają w toku badania statystycznego, pełnego lub metodą reprezentacyjną, a wynikają z innych źródeł niż teoretycznego rozkładu próby losowej i konstrukcji estymatora. Te błędy zauważono w latach sześćdziesiątych, wskazano na ich rolę w badaniach i wnioskowaniu. Teoria tych błędów nie jest jeszcze dostatecznie opracowana[6]). I tak się składa, że dyskusja o tych właśnie błędach jest rozwlekła i złożona oraz niejednolicie przez autorów ujmowana; nieraz bardzo trudno przejść z jednej pracy do drugiej, ponieważ trudno te błędy oszacować. Istotnym jest uczulanie na nie, wskazywanie na ich źródła i miejsca powstawania, aby już w trakcie badania wyeliminować je względnie ograniczyć. Nie zwalnia to oczywiście z obowiązku ich pomiaru, ale ze względu na trudności tego pomiaru, jest to najwłaściwsze postępowanie, zbliżające pomiar do wartości prawdziwej.

---

[5]) Błędy, których źródłem są: stosowany estymator oraz schemat losowania, nazywa się błędami statystycznymi.

[6]) A może nie można się tego spodziewać ze względu na praktyczność tego tematu, co wymaga oddzielnego traktowania dla każdego przypadku?

KLASYFIKACJA BŁĘDÓW NIELOSOWYCH

Dla wyraźniejszego wyodrębnienia tych błędów klasyfikujemy je w trzech grupach, zgodnie z tokiem badania: **błędy przygotowania badania reprezentacyjnego, błędy popełniane w czasie zbierania materiału, błędy opracowania zebranego materiału.**

Przedstawiona klasyfikacja błędów zwraca uwagę na ich powstawanie, uczula na nie, ale nie jest przydatną do ich mierzenia, bowiem w praktyce nie można tych błędów wyodrębnić, a nawet określić. Pragnąc mierzyć błędy nielosowe musimy je klasyfikować inaczej, klasyfikować w zależności od sposobu pomiaru. Zauważono, że można wśród tych błędów wyodrębnić błędy przypadkowe (mające rozrzut) i systematycznie popełniane[7]). Błędy przypadkowe wpływają na rozrzut wartości estymatora, a systematyczne na odchylenie. Ponieważ rozrzut i odchylenie dla błędów losowych potrafimy już mierzyć (wariancja i obciążenie), więc tę samą metodologię zastosujemy i do błędów nielosowych.

W praktyce, często nie można wyodrębnić błędów losowych od nielosowych, mierzymy je przecież jednocześnie, jako wypadkową wszystkich błędów, a więc mierzymy je łącznie, a sam podział na błędy losowe i nielosowe jest tylko konstrukcją myślową, konstrukcją historyczną.

Ze względu na pomiar, podział błędów na losowe i nielosowe jest nieprzydatny. Dlatego wprowadzono nowy podział błędów, zgodny z techniką pomiaru (a więc tak jak się one ujawniają) oraz wprowadzono kilka nowych pojęć, urealniających rozważania, szczególnie przez uwzględnienie pomiarów.

BŁĘDY STAŁE I BŁĘDY ZMIENNE[8])

Metodologicznie istotny jest podział wszystkich błędów badania reprezentacyjnego na błędy stałe i błędy zmienne.

Błędy stałe — są to błędy niezależne od obserwowanego elementu. Mierzyć je będziemy obciążeniem. Aby jednak odróżnić to obciążenie od już wprowadzonego[9]), nazwiemy je **obciążeniem całkowitym.**

Błędy zmienne — są to błędy zależne od obserwowanego elementu (a więc mają nietrywialny rozkład). Mierzyć je będziemy wariancją, którą znowu dla odróżnienia od już wprowadzonego pojęcia, nazwiemy **wariancją całkowitą.**

---

[7]) Specjalnie używam tych terminów, aby uniknąć słowa losowe. Terminologia anglo-saska jest w lepszej sytuacji, bowiem posiada kilka określeń losowości.

[8]) Właściwie omawianie błędów badania reprezentacyjnego od tego można zacząć, ale ze względu na obciążenia z przeszłości, dla wyjaśnienia narastających pojęć istniejących w literaturze, trzeba było powyższy rys podać.

[9]) Pojęć stosowanych w MR, a zaczerpniętych ze statystyki matematycznej i rachunku prawdopodobieństwa. W istocie, pojęcia są nie zmienione, należy tylko zwrócić uwagę na rozkład statystyki, który jest tutaj stosowany; raz jest to rozkład teoretyczny (T_n), innym razem jest to rozkład zrealizowany w badaniu (X_n).

Można by pominąć określania błędów stałych oraz zmiennych i mówić bezpośrednio o obciążeniu całkowitym i wariancji całkowitej. I to było by najściślejsze[10]), ponieważ część błędów stałych też zależy od rozkładu. Wprowadzono te pojęcia tylko dla odróżnienia błędów od metod ich pomiaru, analogicznie, jak jest przy błędach losowych i nielosowych.

PODSTAWOWE POJĘCIA

Istotne warunki badania. Jest to zbiór warunków, których wynikiem ma być zbliżenie do poznania prawdziwych wartości. Warunki te obejmują: procedury wyboru losowego i estymacji, technikę przeprowadzenia obserwacji, zatrudnienie i szkolenie ankieterów, metody symbolizacji i opracowania tablic wynikowych. Określają teoretycznie oczekiwane wyniki obserwacji i obciążenia. Jednak zawsze opis tych warunków pozostaje niepełny, np. o błędach zmiennych nie można niczego powiedzieć bez pobrania próby losowej.

Wartość w populacji generalnej. Oznaczamy ją wskaźnikiem *p*: $\bar{\mu}_p$. Tę wartość otrzymamy, gdy próba losowa będzie zawierać wszystkie elementy populacji generalnej, poddane tym samym istotnym warunkom badania. W zasadzie można przyjąć, że ta wartość jest wolna od błędów zmiennych.

Wartość prawdziwa. Oznaczamy ją wskaźnikiem *pr*: $\bar{\mu}_{pr}$. Jest to wartość w populacji generalnej wolna od błędów wszelkiego rodzaju. Jej znajomość jest celem planowania badania reprezentacyjnego.

Estymand. Oznaczamy go przez $\bar{\mu}$. Jest to wartość w populacji generalnej wolna od obciążeń, które mogłyby być spowodowane konstrukcją estymatora i zastosowanego schematu losowania. Wartość ta służy do zdefiniowania nie obciążoności estymatora. Zależy ona od wyboru losowego, estymacji i istotnych warunków badania.

Obciążenie całkowite. Formalnie jest mierzone różnicą między wartością oczekiwaną estymatora a wartością prawdziwą:

$$I.(\bar{X}) - \bar{\mu}_{pr} = BIAS$$

Oznaczam tutaj celowo estymator przez $\bar{X}$, a nie przez T, w celu zaznaczenia odmienności rozkładów; rozkład estymatora $\bar{X}$ jest rozkładem rzeczywistym, a więc obarczonym błędami badania[11]). Istotnym problemem badania reprezentacyjnego jest pomiar tego obciążenia. Pragnąc wyodrębnić źródła obciążeń korzystamy z tożsamości, np.:

$$E(\bar{X}) - \bar{\mu}_{pr} \equiv [E(\bar{X}) - \bar{\mu}] + [\bar{\mu} - \bar{\mu}_{pr}]$$

---

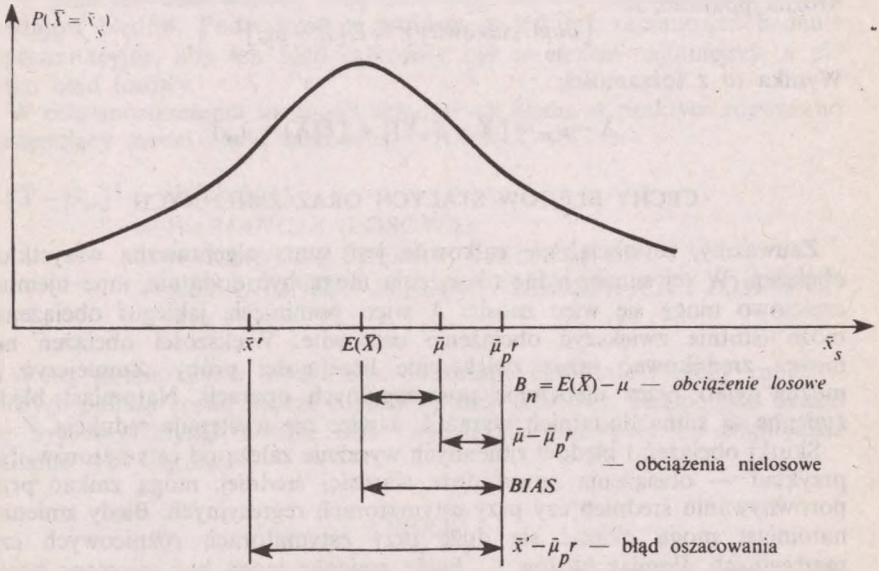
[10]) Na zaburzenie ścisłości wskazane zostanie przy omówieniu terminu błędy systematyczne.

[11]) Bo chociaż zamiast T_n jest tutaj $\bar{X}$, to w różnicy wielkości $\bar{X}$ i $\bar{\mu}$ znoszą się obciążenia nielosowe.

gdzie: pierwszy składnik — to wielkość obciążeń losowych, a drugi — to obciążenia nielosowe. Obciążenia losowe można rozłożyć dalej, korzystając z tożsamości:

$$E(\bar{X})-\bar{\mu}\equiv[E(\bar{X})-\bar{\mu}_p]+[\bar{\mu}_p-\bar{\mu}]$$

gdzie: pierwsze, to obciążenia zgodne, znikające ze wzrostem próby, a drugie, to obciążenie które występuje i przy badaniu pełnym.



Przedstawiona analiza obciążeń pozwala na wyjaśnienie terminu — **błędy systematyczne.** Tym określeniem można nazwać obciążenia nielosowe oraz tę część obciążeń losowych, które nie są zgodne (bowiem obciążenia zgodne zależą od rozkładu, a więc nie są systematyczne). Ponieważ, jak z praktyki wynika, udział obciążeń losowych jest pomijalny w mierzonym obciążeniu całkowitym, to jeszcze można nazwać obciążenie całkowite — błędem systematycznym. Uważam jednak, że rozciągnięcie tego określenia na błędy nielosowe jest zbyt mylące, chociaż praktyka pokazuje, że składowa wariancji, wynikająca z błędów nielosowych, jest pomijalna. Błąd systematyczny nie zależy od wielkości próby; jego występowania nie należy wyolbrzymiać wówczas, gdy potrafimy go oszacować.

Wariancja całkowita. Mierzy błędy zmienne i jest równa:

$$E[(\bar{X})-E(\bar{X})]^2=VE^2$$

W przypadku występowania tylko błędów losowych, to:

$$VE^2 \equiv D^2(\bar{X}) = D^2(T_n)$$

Tę właśnie wariancję mierzymy, wykorzystując dane uzyskane z próby.

Błąd całkowity. Z definicji wynika, że błąd całkowity jest równy:

$$Błąd\ całkowity = \sqrt{VE^2 + BIAS^2}$$

Pojęcie to zastępuje **błąd standardowy,** który jest jego uogólnieniem. Można pokazać, że

$$[Błąd\ całkowity]^2 = E[\bar{X} - \bar{\mu}_{pr}]^2$$

Wynika to z tożsamości:

$$\bar{X} - \bar{\mu}_{pr} = [\bar{X} - E(\bar{X})] + [E(\bar{X}) - \bar{\mu}_{pr}]$$

CECHY BŁĘDÓW STAŁYCH ORAZ ZMIENNYCH

Zauważmy, że obciążenie całkowite jest sumą algebraiczną wszystkich obciążeń. W tej sumie, jedne obciążenia mogą być dodatnie, inne ujemne, częściowo mogą się więc znosić. A więc, pominięcie jakiegoś obciążenia może istotnie zwiększyć obciążenie całkowite. Większości obciążeń nie można zredukować przez zwiększenie liczebności próby. Zmniejszyć je można tylko przez ulepszenie poszczególnych operacji. Natomiast błędy zmienne są sumą dodatnich wyrażeń, a więc nie występuje redukcja.

Skutki obciążeń i błędów zmiennych wyraźnie zależą od estymatorów. Na przykład — obciążenia nawet duże wartości średniej, mogą znikać przy porównywaniu średnich czy przy estymatorach regresyjnych. Błędy zmienne natomiast mogą okazać się duże przy estymatorach różnicowych czy regresyjnych. Pomiar błędów — błędy zmienne mogą być mierzone przez notowanie odchyleń, występujących między jednostkami w próbie. Ich pomiar wymaga, aby właściwie zaplanowano oddzielenie źródeł zmienności. Pomiar obciążeń wymaga stosowania innych metod — metod zewnętrznych do właściwości badania.

SZACOWANIE NIEDOKŁADNOŚCI WYNIKÓW BADANIA REPREZENTACYJNEGO

Wyróżniamy trzy sposoby szacowania niedokładności wyników:

1) najczęściej wykorzystuje się dane z innych źródeł, z którymi porównuje się otrzymane wyniki. Jednak, aby szacunek był prawdziwy, to muszą być spełnione dwa warunki:
   — musi istnieć niezależne źródło informacji,
   — źródło to musi być dokładne;

2) inną metodą jest badanie zgodności pewnych relacji;
3) jedyną właściwą metodą szacowania niedokładności wyników badania reprezentacyjnego jest przeprowadzenie badania kontrolnego. Jest to również badanie reprezentacyjne, którego jednak celem jest uchwycenie źródeł niedokładności przeprowadzonego badania.

MODEL BŁĘDU CAŁKOWITEGO

W praktyce obserwujemy błąd całkowity, będący wypadkowym z obu rodzajów błędów. Podstawowym problemem jest tak zaplanować badanie reprezentacyjne, aby ten błąd całkowity był w efekcie najmniejszy, a nie tylko błąd losowy.

W celu spostrzeżenia istotności składowych błędu, w praktyce rozważano następujący model błędu całkowitego. Niech $\bar{X} \sim \Theta$

$$
\begin{aligned}
E[\bar{X} - \Theta_{pr}]^2 &= VE^2 + BIAS^2 \\
&= WARIANCJA \text{ (LOSOWA)} \\
&+ WARIANCJA\ BŁĘDÓW\ NIELOSOWYCH \\
&+ KOWARIANCJA\ BŁĘDÓW\ NIELOSOWYCH\ I\ LOSOWYCH \\
&+ BIAS^2.
\end{aligned}
$$

W tej formie mamy wydzielone obciążenie całkowite i błędy zmienne, których pomiar został jeszcze rozbity na trzy składowe. Ten podział okazał się być wygodnym, bowiem takie wydzielenie pozwoliło na empiryczne badanie tych błędów.

WNIOSKI

1. Jeśli przeprowadzane jest badanie pełne, to wariancja i kowariancja błędów losowych i nielosowych są równe zeru.

2. Zauważono, że kowariancja błędów losowych i nielosowych jest na ogół mała i może być pominięta.

Ogólnie można powiedzieć, że ze wszystkich błędów badania reprezentacyjnego, najbardziej znaczące są wariancja i obciążenia nielosowe. Wynika z tego, że wyznaczenie liczebności próby w oparciu tylko o wariancję jest prawidłowe. Konieczna jest jednak znajomość obciążenia nielosowego (błędu systematycznego), aby skorygować wynik. Jeżeli udałoby się wyznaczyć liczebność próby w oparciu o błędy zmienne (to znaczy, uwzględniając wariancję całkowitą oraz obciążenia losowe zgodne), to dla zaznaczenia tego faktu, tak wyznaczoną liczebność nazwalibyśmy efektywną liczebnością próby.

Warto zauważyć, że to wynikające z praktyki spostrzeżenie dobrze koresponduje z klasyfikacją błędów badania reprezentacyjnego, a mianowicie

dwie kategorie (podane przez Jana Kordosa): kategoria pierwsza — stanowią ją błędy, wynikające z tego, że to co obserwujemy odbiega od tego co zamierzamy pomierzyć oraz kategoria druga — stanowią ją błędy powstałe w procesie uogólniania wyników z próby na całą populację.

Ta klasyfikacja jest bardzo trafną i zwartą i być może tutaj otrzymujemy zgodność klasyfikacji źródeł błędów oraz sposobów ich pomiaru i uwzględniania ich w wynikach. Sam jestem tej myśli.

LITERATURA

1. W. E. Deming: *Some Theory of Sampling*, A Wiley, 1950.
2. Leslie Kish: *Survey Sampling*, A Wiley, 1965.
3. Jan Kordos: *Dokładność danych w badaniach społecznych*, GUS, Warszawa 1987.
4. Jan Kordos: *Konieczność uwzględnienia błędów losowych i nielosowych przy planowaniu badań reprezentacyjnych i wykorzystaniu ich wyników*, referat, na str. 98
5. M. N. Murthy: *Sampling Theory and Methods*, Calcutta, 1967.
6. Norman L. Jonson and Harry Smith Jr.: *New Developments in Survey Sampling*, A Wiley 1969.
7. Bill Williams: *A Sampler on Sampling*, A Wiley 1978.

O błędach nielosowych w badaniu dzietności kobiet w ramach Narodowego Spisu Powszechnego 1970

doc. dr hab. Jan Paradysz
Akademia Ekonomiczna w Poznaniu

Badanie dzietności kobiet w ramach narodowych spisów powszechnych było od dawna postulatem demografów. Takie propozycje badań zgłaszano jeszcze w okresie międzywojennym przed spisem 1931. Jednakże liczne zaostrzenia ze strony Głównego Urzędu Statystycznego, głównie natury organizacyjnej, finansowej i kadrowej, powodowały, że ten słuszny postulat demografów zrealizowano dopiero w 1970 r. Niestety, wyniki badań opublikowano w postaci zaledwie 18 tablic wynikowych oraz kilku opracowań analitycznych. Utrudniony dostęp do danych pierwotnych, w postaci kwestionariuszy spisowych i taśm magnetycznych[1]), nie pozwolił

[1]) Żadne z tych źródeł już nie istnieje. Formularze spisowe do badania dzietności kobiet zostały komisyjnie zniszczone około 1980 r.

szerszemu kręgowi demografów spoza Głównego Urzędu Statystycznego na pogłębioną analizę tego bardzo cennego materiału, dotyczącego reprodukcji ludności Polski zarówno w ujęciu wzdłużnym, jak i poprzecznym[2]). Zapewne z tego powodu opublikowane wyniki nie były przedmiotem należytego zainteresowania specjalistów i nie doczekały się właściwej oceny. O ile wiadomo, nie było też prac oceniających reprezentatywność tego badania, ze szczególnym uwzględnieniem błędów nielosowych. W niniejszym referacie pragnę przedstawić próbę zbadania obciążenia błędem statystycznym jednej z najważniejszych cech w tym badaniu, jaką jest staż małżeński.

Już pierwszy rzut oka na tabl. 7 w opublikowanym pierwszym tomie *Dzietności kobiet* (1971, s. 80 i 82) pozwala dostrzec bardzo niską liczbę 169 tys. kobiet o stażu małżeńskim 0 lat. Skądinąd wiadomo, że w 1970 r. zawarto 280,3 tys. małżeństw, w tym 259,1 tys. przez panny. Zatem w tej grupie stażu małżeńskiego nie doszacowano prawie 100 tysięcy. Czy tak było w istocie?

Zanim przejdę do analizy różnic pomiędzy liczbami kobiet kiedykolwiek pozostających w związku małżeńskim „prawdziwymi" i wykazanymi przez NSP'70, należy wyjaśnić co się kryje w cytowanej publikacji GUS, pod określeniem „długość trwania małżeństwa". W jaki sposób obliczano staż małżeński? Otóż w „Formularzu do badania dzietności kobiet" nie ma daty ślubu, jest tylko rok kalendarzowy zawarcia pierwszego małżeństwa. Ponieważ NSP'70 nie ujmował stanu na koniec roku, lecz na 8 grudnia 1970 r., to nie można było obliczyć dla każdej kobiety stażu małżeńskiego w latach ukończonych. To co nazwano długością trwania małżeństwa jest w istocie stażem osiągniętym w ciągu 1970 r. przez poszczególne kohorty małżeństw. Zatem zerowy rok trwania małżeństwa odpowiada kohorcie małżeństw 1970, 1 — 1969, 2 — 1968, 3 — 1969 itd. Uwzględniając sezonowość małżeństw — por. Rocznik Demograficzny 1971, s. 76 — pomiędzy momentem spisu a końcem 1970 r. mogło być zawarte około 21 tys. pierwszych małżeństw, czyli w 1970 r. do momentu spisu zawarto ponad 238 tys. małżeństw, w których nupturientka była panną.

Przechodząc do analizy liczb zestawionych w tabl. 1 należy zwrócić uwagę, że nie uwzględniono w niej następujących kategorii kobiet zamężnych:

a) zmarłe przed momentem spisu,
b) saldo migracji zagranicznych, które w latach 1958—1970 było ujemne,
c) wodowy i rozwiedzione, które przed spisem 1970 powtórnie wyszły za mąż.

Pominięcie pierwszej z tych kategorii nie ma większego znaczenia, analizowane tu kohorty małżeństw zawierają głównie kobiety młode, poniżej

---

[2]) Wysoka wartość tego materiału źródłowego polegała między innymi na tym, że dotyczył on także okresu sprzed 1960 r., dla którego, praktycznie biorąc, nie ma innych źródeł statystycznych.

40 roku życia, w momencie spisu. Abstrahowanie od salda migracji zagranicznych jest koniecznością, gdyż bardzo skąpa i wątpliwa ich statystyka nie pozwala na dostatecznie rozsądne szacunki spadku liczebności kohort małżeńskich z tego tytułu.

Konieczne jest także pominięcie w omawianych kohortach małżeństw tych kobiet, które powtórnie wyszły za mąż. W cytowanej publikacji GUS żadna z tablic, zawierająca długość stażu małżeńskiego, nie uwzględnia kobiet powtórnie wychodzących za mąż. Brak jest też pogłębionych badań na ten temat[3]). Oczywiście ani to źródło różnic pomiędzy $M1(t)$ i $K(t)$, ani też dwa poprzednie nie mają poważniejszego znaczenia w przypadku dwóch najmłodszych kohort małżeńskich 1969 r. i 1970 r. Bez obawy popełnienia większego błędu można powiedzieć, że w tych dwóch kohortach małżeństw odsetki nie przebadanych mężatek wynoszą odpowiednio 9 i 29%. Przy szacowaniu braków w pozostałych kohortach małżeńskich należy jednak zachować dużą ostrożność, do czego skłania kształtowanie się wskaźników $w(t) = K(t) : M1(t)$. Dla kohort 1960 i 1965 $w(t)$ są większe od jedności, natomiast dla sąsiednich kohort wskaźniki natężenia są wyraźnie niższe. Taki układ jest charakterystyczny dla błędów „pamięci" w badaniach anamnestycznych. Można jednak wyrazić wątpliwości, czy są to istotnie błędy „pamięci" popełniane przez respondentów, wszak ślub jest tak dużym wydarzeniem w życiu człowieka, a zapewne jeszcze większym w życiu kobiety, że po kilku zaledwie latach nie zapomina się jego daty. Zatem usprawiedliwione będzie wymienienie kilku innych alternatyw, a raczej możliwości, gdyż nie jest wykluczone, że w badanej zbiorowości występują jednocześnie. Zatem, po drugie, nie jest wykluczone podawanie alternatywnie bądź ślubu kościelnego, bądź cywilnego z preponderencją tego, którego rok kończył się „okrągłą" liczbą. Po trzecie, nie jest wykluczona mniejsza lub większa nierzetelność rachmistrza spisowego. Natomiast, co się tyczy najmłodszych kohort małżeńskich, bardzo duże braki młodych kobiet mogły być spowodowane dwoma czynnikami: duża mobilność kobiet w związku z małżeństwem (zmiany adresu po wyjściu za mąż) oraz pominięcie mężatek, mieszkających w gospodarstwach zbiorowych, jak: hotele robotnicze, domy studenckie, domy rencistów itp.

Na koniec należy odpowiedzieć na pytanie, czy tak znaczne obciążenie próby z badanej populacji, pod względem tak ważnej cechy, jak staż małżeński, nie dyskwalifikuje tych materiałów statystycznych jako źródła danych dla badań demograficznych. Moim zdaniem, nie dyskwalifikuje, chociaż nieco je ogranicza, a w każdym razie nakazuje dużą ostrożność przy korzystaniu z tych danych. Ponadto, jeśli chodzi o wykorzystanie tego materiału we wzdłużnej analizie demograficznej, to najmłodsze kohorty

[3]) Nieliczne prace, dotyczące powtórnych związków małżeńskich odnoszą się do ujęcia przekrojowego, por. na przykład L. Bolesławski; *Tablica zmian stanu cywilnego kobiet 1970–1972*. Wiadomości Statystyczne nr 9, 1974 s. 20—22.

TABL. 1. PORÓWNANIE LICZB KOBIET OBJĘTYCH BADANIEM DZIETNOŚCI W RAMACH NSP'70, KTÓRE TYLKO RAZ ZAWARŁY ZWIĄZEK MAŁŻEŃSKI, Z LICZBĄ MAŁŻEŃSTW (w tysiącach osób)

y	Rok t	Kobiety objęte spisem			Małżeństwa panien	$\frac{K(t)}{M1(t)}$
		mężatki	reszta	razem		
		$Km(t)$	$R(t)$	$K(t)$	$M1(t)$	
0	1970	160,8	8,2	169,0	238,3[1]	0,709
1	1969	205,2	21,2	226,4	250,1	0,905
2	1968	208,2	25,4	233,6	238,3	0,980
3	1967	190,9	26,2	217,1	219,6	0,989
4	1966	176,3	27,1	203,4	208,0	0,978
5	1965	166,3	30,0	196,3	183,2	1,072
6	1964	170,4	30,0	200,4	212,3	0,944
7	1963	165,3	31,0	196,3	202,3	0,970
8	1962	173,0	29,9	202,9	210,9	0,962
9	1961	171,9	30,6	202,5	219,2	0,924
10	1960	200,9	33,7	234,6	227,9	1,029
11	1959	192,8	31,3	224,1	259,2	0,865
12	1958	204,1	32,1	236,2	248,4	0,951

Oznaczenia: y — długość trwania małżeństwa (por. zastrzeżenia w tekście), t — rok zawarcia związku małżeńskiego.

Uwaga. 1. Liczbę małżeństw w 1970 r. pomniejszono o szacunek związków małżeńskich zawartych w okresie 9—31 grudnia 1970 r. Było ich około 20,8 tys.

Źródło: zestawienie i obliczenia własne na podstawie różnych edycji Statystyki ludności, Roczników Demograficznych 1945—1966 i 1967—1968 oraz *Ostateczne wyniki NSP'70. Dzietność Kobiet. Polska*, część 1, GUS, Warszawa 1971 s. 80 i 82.

małżeńskie mają niewielkie znaczenie ze względu na krótki staż małżeński. Natomiast w przekrojowej analizie demograficznej zniekształcenie rzeczonej struktury ma duże znaczenie, ale można to naprawić. W tym celu wystarczy oszacować „prawdziwe" wskaźniki natężenia $\hat{w}(t)$, a następnie, na ich podstawie, obliczyć nowe liczby kobiet $\hat{K}(t)$ dla $t=1959, 1960, 1961, 1964, 1965, 1966, 1969$ i 1970. Dla wyrównania $w(t)$ pominięto dwie najbardziej obciążone kohorty 1969 i 1970, a dla okresu 1958—1968 wypróbowano trzy funkcje: liniową, potęgową i Törnquista pierwszego typu[4]). Uzyskano następujące wyniki:

a) $$\hat{w}(\hat{t}) = 0{,}9260 + 0{,}0072\ \hat{t} \qquad R = 0{,}406$$

[4]) Odnośnie metody estymacji tej funkcji patrz T. Stanisz: *Funkcje jednej zmiennej w badaniach ekonomicznych*. PWN, Warszawa 1986 s. 188—189.

b) $$\hat{w}(\hat{t}) = 0{,}9191 \cdot \sqrt[30{,}62]{\hat{t}} \qquad R = 0{,}433$$

c) $$\hat{w}(\hat{t}) = \frac{0{,}986667\ \hat{t}}{\hat{t} + 0{,}075034}, \qquad R = 0{,}353$$

gdzie wszędzie $\hat{t} = 1, 2, ..., 11$ dla lat 1958, 1959, ..., 1968.

Obok równań podano współczynniki korelacji wielorakiej pomiędzy wskaźnikiem $w(\hat{t})$ a zmienną czasową $\hat{t}$. W każdym przypadku współczynnik ten jest niski. Mimo to, uważam, że modele tendencji rozwojowej b) i c) dobrze nadają się do wyrównania $w(t)$ w okresach 1959—1960 i 1964—1966 oraz ich ekstrapolacji na lata 1969 i 1970. Model a) nie nadaje się do ekstrapolacji na ostatnie wskazane lata, gdyż $w(t)$ przekracza wówczas 1. Warto podkreślić, że dodatkową zaletą funkcji Törnquista jest możliwość oszacowania górnej asymptoty, w naszym przypadku jest ona równa 0,986667. Funkcje b i c mają podobny kształt, obie są wypukłe do góry i charakteryzują się malejącym tempem przyrostów.

Oszacowane modele tendencji rozwojowej wskaźników $w(t)$ wykorzystano do obliczenia skorygowanych liczb kobiet, składających się na ich strukturę według stażu małżeńskiego w momencie NSP'70. Skorygowane liczebności kobiet obliczono według wzoru:

$$\hat{K}(t) = w(\hat{t}) \cdot M1(t)$$

TABL. 2. SKORYGOWANE LICZEBNOŚCI KOBIET WEDŁUG ROKU ZAWARCIA ZWIĄZKU MAŁŻEŃSKIEGO I STANU NA DZIEŃ 8 XII 1970 R., KTÓRE KIEDYKOLWIEK POZOSTAWAŁY W ZWIĄZKU MAŁŻEŃSKIM

Rok t	$\hat{w}(t)$ według funkcji			Liczby kobiet — $\hat{K}(t)$		
	a	b	c	a	b	c
1970		0,9994	0,9810	—	238,2	233,8
1969	—	0,9968	0,9805	—	249,3	245,2
1966	0,9836	0,9875	0,9785	204,6	205,4	203,5
1965	0,9764	0,9837	0,9775	178,9	180,2	179,1
1964	0,9692	0,9794	0,9762	205,8	207,9	207,2
1961	0,9476	0,9617	0,9685	207,7	210,8	212,3
1960	0,9404	0,9527	0,9626	214,3	217,1	219,4
1959	0,9332	0,9401	0,9510	241,9	243,7	246,5

Źródło: obliczenia własne na podstawie danych w tabl. 1.

Jak to wykazują dane tabl. 2, model a) daje wyniki bardziej odbiegające od dwóch pozostałych i, w tym przypadku, nie nadaje się do ekstrapolacji na lata 1969 i 1970. Model potęgowo-pierwiastkowy i Törnquista wykazują większe rozbieżności przy ekstrapolacji, gdzie różnica pomiędzy wynikami dochodzi do 5,5 tys. mężatek. Oceniając ten skrajny wynik merytorycznie można go potraktować jako swego rodzaju przedział ufności, gdyż model b) daje dla 1970 r. wartość przeszacowaną, a model c) nie doszacowaną. Podobnie można byłoby zinterpretować wynik obliczeń dla 1969 r. Co się zaś tyczy uzyskanych wyników wyrównywania obu skupień (1960 i 1965), to zbieżność wyników, jakie dają oba modele, jest zadowalające.

METODA REPREZENTACYJNA W PRAKTYCE STATYSTYCZNEJ

Zagadnienia „braku odpowiedzi" w badaniach gospodarstw domowych

mgr Antoni Kubiczek
Główny Urząd Statystyczny

Jednym z istotniejszych problemów w badaniach statystycznych jest brak informacji ze wszystkich objętych badaniem jednostek. Brak ten występuje w szczególności w badaniach społecznych.

Brak informacji z niektórych jednostek, prowadzić może do niekompletnej informacji daleko odbiegającej od stanu rzeczywistego. Dzieje się tak zwłaszcza wtedy, kiedy występuje silna koncentracja „braku odpowiedzi" w określonych grupach jednostek, wyodrębnionych z punktu widzenia, np. wieku, płci, miejsca zamieszkania czy innych istotnych cech, zależnych od stosownego zachowania się jednostek.

Przykładowo, badając uczestnictwo ludności w kulturze, wiadomo że najbardziej aktywnymi kulturalnie są osoby w wieku 18—30 lat, z wykształceniem wyższym i mieszkańcy miast. W przypadku więc silnej koncentracji „brak odpowiedzi" w tych właśnie grupach wyniki badania są zaniżane w stosunku do stanu rzeczywistego; odwrotnie — w przypadku koncentracji „braku odpowiedzi" w rocznikach starszych, z wykształceniem niższym i mieszkańców wsi — wyniki badania są zawyżane.

W swoim referacie pragnąłbym przedstawić niektóre tylko aspekty związane z przyczynami występowania „braku odpowiedzi", metody ich eliminacji oraz zaprezentować zjawisko „braku odpowiedzi" w badaniach budżetów gospodarstw domowych, prowadzonych przez GUS i na ich tle przedstawić kilka wniosków, dotyczących braku odpowiedzi w badaniach reprezentacyjnych.

Należy dodać, że literatura omawiająca zagadnienie „braku odpowiedzi" nie jest obszerna, a zwłaszcza polska. Jednym z propagatorów tej tematyki

jest Richard Platek [1]. Jego prace, w wielu przypadkach pionierskie, stanowią bogate źródło dociekań naukowych. W literaturze polskiej jedynym autorem, który omawia zjawisko „braku odpowiedzi" i analizuje jego wpływ na precyzję i dokładność badań statystycznych, jest Jan Kordos [2].

„BRAK ODPOWIEDZI" — PRZYCZYNY I KONSEKWENCJE

W badaniach reprezentacyjnych, obserwacją objęta jest losowo dobrana próbka (n) gospodarstw domowych (osób). Uzyskane informacje z takiej próby są ocenami parametrów populacji generalnej (N). W wielu przypadkach niecała wylosowana próba podlega badaniu, bowiem część jednostek — z założenia — nie bierze udziału w badaniu. Spośród natomiast jednostek podlegających badaniu, tylko część składa wymagane informacje (wykres).

Najczęściej używaną miarą, charakteryzującą przebieg badania, jest wskaźnik realizacji badania, definiowany jako iloraz liczby jednostek zbadanych i liczby podlegających badaniu:

$$\hat{R}=\frac{n_r}{n_e}$$

Wskaźnik „braku odpowiedzi" wynika z dopełnienia do jedności:

$$(N\hat{R})=1-\hat{R}$$

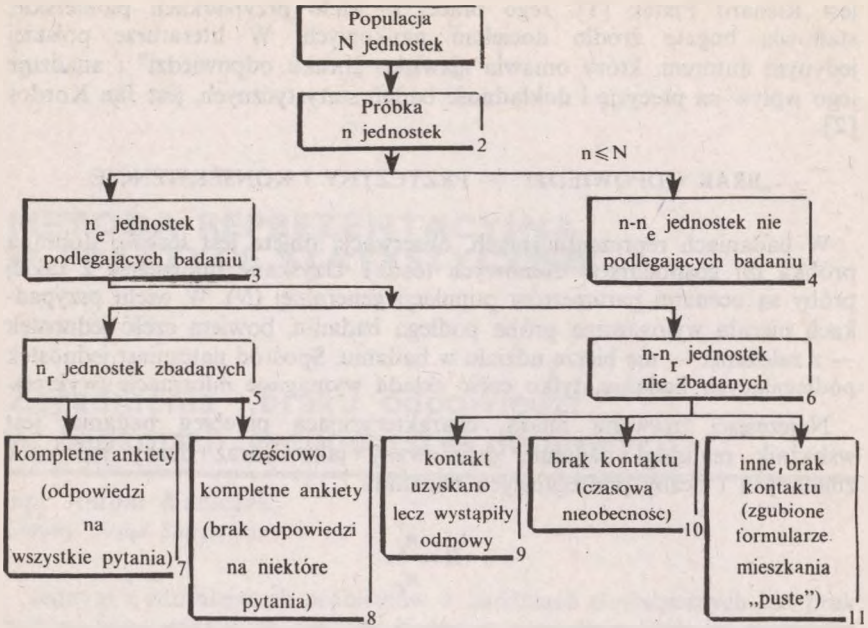
Mogą być jednak stosowane i inne miary, takie jak:

(a) wskaźnik kontaktu $=\dfrac{(5)+(9)}{(5)+(6)}$;

(b) wskaźnik kompletności $=\dfrac{(5)}{(2)}$;

(c) wskaźnik efektywności kontaktu $=\dfrac{(5)}{(5)+(9)}$ itd.

Powodem braku odpowiedzi mogą być przyczyny zarówno subiektywne, jak i obiektywne. Do przyczyn subiektywnych zaliczyć należy wszystkie te, które są wyrazem postaw jednostek wobec samego aktu badawczego, względnie tematu badania; natomiast do przyczyn obiektywnych — te, które spowodowane zostały nieuzyskaniem kontaktu z wylosowaną jednostką. A zatem przyczyny obiektywne, w bardzo wielu przypadkach, zależą od samego badacza, od jego umiejętności, doświadczenia i wiedzy. Dlatego też im poświęcę najwięcej uwagi.



Źródło: [1].

Jakie mogą być przyczyny zarówno obiektywne, jak i subiektywne, braku odpowiedzi:

Operat losowania. Postulatem podstawowym we wszystkich badaniach jest, aby operat losowania był pełny i aktualny. Zdarza się często, że czas, jaki upłynął od sporządzenia operatu do aktu badawczego, wynosi 2—3 lata. Operat losowania ulega zatem deaktualizacji (brak mieszkań — wyburzenia, pożary itd. względnie zmiana miejsca zamieszkania, zgony itd.).

Termin badania. Wybór odpowiedniego terminu badania jest jednym z bardzo istotnych czynników, który ma wpływ na wielkość zjawiska „braku odpowiedzi". Przykładami mogą być np. okres wakacyjny dla badania zagadnień młodzieży, godziny przedpołudniowe dla badania osób aktywnych zawodowo, okres żniw dla badania gospodarstw chłopskich. Znaczący wpływ na wysokość odsetka „braku odpowiedzi" ma również sytuacja społeczno-polityczna, zwłaszcza w badaniach społecznych.

Tematyka badania. Z wielkim trudem znaleźć można temat, który byłby życzliwie przyjmowany przez jednostki badane. Znacznie łatwiej wymienić tematy drażliwe, które z niechęcią przyjmowane są przez badanych. Owa „niechęć" w wielu przypadkach koncentruje się w określonych grupach jednostek, wyróżnionych z punktu widzenia cech społeczno-demograficznych, miejsca zamieszkania itd. Wynika ona z ogólnego poziomu

kulturalnego określonej grupy, jej postaw życiowych, światopoglądu, obyczajów, tradycji itd.

Formularz badawczy. Ważne znaczenie, zwłaszcza na „przyjęcie" badania przez jednostki badane, ma formularz. Układ pytań, ich sformułowanie, forma odpowiedzi są nie tylko ważne z punktu widzenia ich opracowania, ale i wpływa na „komfort wywiadu". W przypadku formularza nadmiernie rozbudowanego może wystąpić u badanego zniechęcenie, zniecierpliwienie, co prowadzić może do przerwania badania. Sformułowanie pytań powinno być proste, zrozumiałe i w żadnym wypadku nie powinno wartościować postawy czy zachowania badanego.

Ankieter. Od odpowiedniej postawy ankietera, jego taktu, doświadczenia, umiejętności nawiązania kontaktu z jednostką badaną zależy w poważnym stopniu powodzenie badania. Stąd dobór ankieterów i ich **szkolenie** jest równie ważnym zadaniem badacza co temat badania. A jest to warunek chyba najtrudniejszy do spełnienia.

Przyczyn „braku odpowiedzi" jest znacznie więcej, ich hierarchia jest bardzo różna w poszczególnych badaniach. Badacz powinien w pierwszej kolejności starać się wyeliminować te przyczyny, które są w zasięgu jego możliwości, zależne od niego. Stąd też, przed przystąpieniem do każdego badania, badacz koniecznie powinien zasięgnąć opinii specjalistów o celu badania, jego narzędziach, formie i metodzie. „Zarozumialstwo" badacza leży u podstaw niepowodzeń wielu badań.

Jakie są wpływy „braku odpowiedzi" na uzyskane wyniki badania? Najogólniej można powiedzieć, iż w przypadku znacznego rozproszenia odsetka braku odpowiedzi, według podstawowych cech badanych jednostek, wpływ może być nieznaczny, nieistotny. W sytuacji jednak koncentracji braku odpowiedzi w określonych grupach wpływ jest wyraźnie większy, a niekiedy prowadzić może do stanu, w którym wyniki są bezużyteczne lub po prostu nieprawdziwe.

Jak zatem ustrzec się od nadmiernego odsetka braku odpowiedzi? Jakie środki powinno się przedsięwziąć, aby zminimalizować odsetek braku odpowiedzi, a tym samym uzyskać wyniki reprezentatywne, z określoną precyzją i dokładnością? Do nich należą niewątpliwie:

1. Ponawianie wizyt. Metoda ta polega na wielokrotnym ponawianiu wizyt w przypadku nieobecności jednostki wylosowanej. Wydłuża ona okres badania, a tym samym opóźnia uzyskanie informacji. Powiększa ona również koszt badania, tym niemniej jest ona w wielu przypadkach bardzo skuteczna. Rzecz jasna, termin badania jest zazwyczaj określony, tak więc ponawianie ogranicza się z reguły do 2—3 wizyt. Jeżeli istnieje możliwość ustalenia terminu obecności jednostki badanej (np. informacje od sąsiadów, pozostałych domowników), skuteczność tej metody jest znacznie większa.

2. Wywiady zastępcze. W przypadku tematów, dotyczących faktów i zdarzeń metoda ta zdaje w pełni egzamin. Wątpliwości powstają

w przypadku tematów subiektywnych, tzn. dotyczących opinii, sądów i odczuć badanego. Wtedy metoda ta, raczej nie powinna mieć zastosowania.

3. Podmiany. Metoda sprowadza się do zastępowania jednostki wylosowanej, w przypadku jej nieobecności lub odmowy, inną jednostką o podobnych cechach np. w tym samym wieku, gospodarstwo o podobnym składzie demograficznym, analogicznym poziomie dochodów itd. Tego typu podmiany nazywa się również **podmianami celowymi**. Mogą również wystąpić **podmiany przypadkowe,** tzn. gospodarstwo sąsiadujące z gospodarstwem wylosowanym względnie inny domownik. Do podmian przypadkowych podchodzić należy jednak z dużą ostrożnością. Najczęściej jednostki te bada się pod względem zgodności struktury społecznej z jednostkami wylosowanymi, i dopiero w przypadku zgodności włącza się je do ogólnego zbioru badanych jednostek.

4. Wybór kwotowy, próbki kwotowe. Metoda ta polega na wyborze jednostek badanych według określonych cech, a uzyskane wyniki są ważone w oparciu o wagi z innych badań (najczęściej ze spisów powszechnych). Zastosowanie tej metody zależne jest od pełnej unifikacji stosowanych w danym badaniu klasyfikacji np. ze spisu powszechnego. Próbki kwotowe i postępowanie przy ich zastosowaniu, to odmiany tzw. standaryzacji wyników. Różnica polega jednak na tym, że **ważenie wyników** jest używane w przypadku, gdy „brak odpowiedzi" wyraźnie zniekształca próbkę wylosowaną, automatycznie wyważoną.

5. Użycie zachęt. Zastosowanie tej metody jest bardzo kosztowne, wiąże się ona z wynagradzaniem uczestników badania premią pieniężną względnie udziałem np. w losowaniu atrakcyjnych nagród rzeczowych. Metoda ta ma wielu zwolenników i przeciwników. Niewątpliwie w przypadku długotrwałych badań, jak np. badań budżetów rodzinnych, premia pieniężna jest formą rekompensaty za wysiłek włożony w badanie, choć jej wysokość powinna być neutralna, tzn. nie powinna skłaniać do nadmiernego entuzjazmu (wyniki mogą być „przesłodzone").

W praktyce stosowane są i inne metody, pozwalające wyeliminować wpływ „braku odpowiedzi" na wyniki badania, na etapie opracowania wyników. Duży przegląd takich metod został zaprezentowany w pracy J. Kordosa [2].

Powstaje jednak zasadnicze pytanie: kiedy, w jakich sytuacjach „brak odpowiedzi" wpływa istotnie lub nieistotnie na wyniki badania, tzn. kiedy wpływ ten jest dopuszczalny, a kiedy nie. Jednoznacznej odpowiedzi na powyższe pytanie nie ma. Niekiedy odsetek odmów na poziomie 3—4% jest dopuszczalny, a powyżej 5% czyni wyniki badania bezużytecznymi; w innych zaś poziom 20—30% braku odpowiedzi jest w pełni zadowalający. Na określenie wpływu braku odpowiedzi mają bowiem znaczenie takie zagadnienia, jak: stopień koncentracji „braku odpowiedzi", wielkość próby, rozproszenie badanych cech w populacji generalnej, schemat losowania itd.

Nie można jednak z góry przesądzić, kiedy wpływ ten jest znaczący, a kiedy mniej, bez dokładnej znajomości populacji „braku odpowiedzi" i jej struktury, z punktu widzenia podstawowych cech społeczno-demograficznych. Dla ustalenia stanowiska w tej sprawie konieczne są specjalne badania „braku odpowiedzi", tzn. rozpoznanie — kto nie wziął udziału w badaniu, jego wiek, płeć, miejsce zamieszkania czy też inne cechy, ważne w danym badaniu.

Znajomość tych struktur pozwala nie tylko na rozstrzygnięcie, czy uzyskane wyniki mogą być istotnie skażone „brakiem odpowiedzi", ale pozwala niekiedy na wyeliminowanie jej wpływu na wyniki, stosując odpowiednie techniki estymacji.

„BRAK ODPOWIEDZI" W BADANIACH BUDŻETÓW GOSPODARSTW DOMOWYCH

Schemat losowania próby do badań budżetów gospodarstw domowych jest dwustopniowy. Na stopniu pierwszym losowane są tzw. terenowe punkty badań (rejony statystyczne), na drugim zaś, w fazie pierwszej, 150 mieszkań w terenowych punktach badań wylosowanych na pierwszym stopniu, a w fazie drugiej odpowiedniej liczby gospodarstw domowych uczestniczących w badaniach budżetów rodzinnych. Omawiając zagadnienie „braku odpowiedzi" w badaniach budżetów, nie sposób pominąć implikacje, wynikające ze schematu losowania.

W fazie pierwszej, drugiego stopnia, przeprowadza się tzw. ankietyzację we wszystkich 150 mieszkaniach w każdym terenowym punkcie badań. Ankietyzacja dostarcza podstawowych informacji o strukturze społeczno-demograficznej badanej populacji, koniecznych do wylosowania, w fazie drugiej, gospodarstw do badań budżetów.

„Brak odpowiedzi" a ankietyzacja. W 1985 r. do badania wylosowano łącznie 135 tys. mieszkań, a po wyeliminowaniu mieszkań zamieszkałych przez gospodarstwa domowe, które w myśl przyjętych założeń nie były objęte badaniem (gospodarstwa domowe pracowników MON, MSW i obywateli państw obcych) — 132 tys. mieszkań. Badanie przeprowadzono w 117,8 tys. mieszkań, w których zamieszkiwało łącznie 121,1 tys. gospodarstw domowych. Ankietyzacji nie przeprowadzono zatem w 14,2 tys. mieszkań, tj. 10,6% ogółu mieszkań, podlegających badaniu.

Przyczynami nieprzeprowadzenia wywiadu były:
— odmowa uczestnictwa w badaniu (3,7% wylosowanych mieszkań),
— brak mieszkań „formalnie" wylosowanych do próby względnie mieszkania niezamieszkałe (3,5% mieszkań),
— brak kontaktu z gospodarstwem domowym — nieobecność czasowa (3,4% mieszkań).

Niestety nie znana była struktura gospodarstw domowych nie objętych ankietyzacją ani z punktu widzenia grupy społeczno-ekonomicznej, ani

struktury demograficznej. Warto wspomnieć, że ankietyzacja gospodarstw domowych połączona jest zwykle z badaniem właściwym. Omawiana wyżej ankietyzacja połączona została z badaniem na temat „sytuacja bytowa gospodarstw domowych w 1985 r. (DS-8)".

Uzyskane informacje społeczno-demograficzne posłużyły do sporządzenia operatu losowania w drugiej fazie, drugiego stopnia losowania gospodarstw do badań budżetów.

„Brak odpowiedzi" w badaniach budżetów gospodarstw domowych. Zastosowana metoda rotacyjna w badaniach budżetów rodzinnych zakładała: a) wymianę kwartalną próby, b) odnowienie 1/3 kwartalnej próby każdego roku. Odnowienie miało na celu zapobiegnięcie ekonomizacji budżetu domowego (poprzez długotrwałe prowadzenie zapisów) oraz zmniejszenie odsetka odmów uczestnictwa w badaniu.

Z uwagi na powyższy tok postępowania próbę gospodarstw domowych, prowadzących budżety rodzinne, można było podzielić na: a) gospodarstwa przystępujące do badania po raz pierwszy, b) gospodarstwa domowe uczestniczące już w badaniach w roku (latach) poprzednim. W związku z odmiennym „zachowaniem" się wyżej wymienionych gospodarstw, omówione zostaną oddzielnie.

Gospodarstwa przystępujące do badania budżetów po raz pierwszy.

Odsetek gospodarstw nie podejmujących badania, wyniósł w 1986 r. 31,5%, przy czym naistotniejszą przyczyną były odmowy (16,2%), znacznie mniejszymi udziałami charakteryzowały się pozostałe przyczyny: przewlekła choroba lub rozpad gospodarstwa domowego (5,5%), utrudniony kontakt z gospodarstwem (4,7%), zmiana miejsca zamieszkania (1,6%) oraz inne przyczyny (zgon, wyburzenia itd. — 3,5%).

TABL. 1. NIEPODJĘCIE BADANIA BUDŻETÓW RODZINNYCH WEDŁUG PRZYCZYN I GRUP SPOŁECZNO-EKONOMICZNYCH GOSPODARSTW DOMOWYCH

Gospodarstwa domowe	Ogółem	Przyczyny niepodjęcia badania				
		odmowy	zmiana miejsca zamieszkania	przewlekła choroba, rozpad gospodarstwa	utrudniony kontakt z gospodarstwem	inne przyczyny
	w % ogółu gospodarstw danej grupy					
Ogółem	**31,5**	**16,2**	**1,6**	**5,5**	**4,7**	**3,5**
Pracownicze	31,1	18,4	1,9	2,7	4,5	3,6
Robotniczo-chłopskie	22,5	16,3	0,4	1,5	2,5	1,8
Chłopskie	28,6	18,1	0,8	4,5	3,5	1,7
Emerytów i rencistów	36,9	11,1	1,7	13,1	6,5	4,5

Najwyższy odsetek niepodjęcia badań wystąpił w gospodarstwach emerytów i rencistów, przy czym najpoważniejszą przyczyną była choroba, która uniemożliwia prowadzenie zapisów przez dłuższy okres. Odmowy w tej grupie gospodarstw również w znacznej mierze wynikały z tytułu wieku badanych osób. Niepodjęcie badania budżetów w przeważającej mierze wystąpiło w gospodarstwach małych (1- i 2-osobowych).

TABL. 2. NIEPODJĘCIE BADANIA BUDŻETÓW RODZINNYCH WEDŁUG LICZBY OSÓB W GOSPODARSTWIE DOMOWYM

Gospodarstwa domowe	Ogółem	Gospodarstwa o liczbie osób					
		1	2	3	4	5	6 i więcej
	w % ogółu gospodarstw danej grupy						
Ogółem	**31,5**	**43,8**	**33,2**	**31,9**	**27,7**	**24,1**	**22,9**
Pracownicze	31,1	44,4	36,2	33,3	27,9	26,1	23,2
Robotniczo-chłopskie	22.5	44,4	26,4	23,9	26,5	18,9	19,3
Chłopskie	28.6	44,9	28,4	28,9	27,3	22,5	27,7
Emerytów i rencistów	36,9	43,7	32,4	31,5	29,1	24,3	32,2

Struktura wewnętrzna poszczególnych grup społeczno-ekonomicznych oraz dane w tabl. 2, wyjaśniają powody zróżnicowania odsetka niepodjęcia badania, w poszczególnych grupach gospodarstw. Przewaga gospodarstw licznych (5- i 6-osobowych) w grupie gospodarstw robotniczo-chłopskich (przeciętna wielkość gospodarstwa — 4,56) oraz względnie najniższe odsetki niepodjęcia badania w gospodarstwach dużych, wyjaśnia najniższy odsetek „ogółem" niepodjęcia badania w tej grupie społecznej. W gospodarstwach emerytów i rencistów odwrotnie; gospodarstwa są z reguły małe (przeciętna wielkość gospodarstwa — 1,86) oraz względnie najwyższe odsetki niepodjęcia w tych gospodarstwach (1- i 2-osobowych), stąd najwyższy odsetek „ogółem" niepodjęcia badania.

Z punktu widzenia poziomu uzyskiwanego dochodu przez gospodarstwa domowe, niepodjęcie badania koncentrowało się w niskich i wysokich grupach dochodowych. Biorąc pod uwagę termin badania, najwyższe odsetki niepodjęcia badania notuje się w III kwartale, a więc w okresie urlopów i żniw.

Gospodarstwa przerywające badania. Ta grupa gospodarstw uczestniczyła już w badaniach budżetów w okresie (okresach) wcześniejszym, stąd odsetki niepodjęcia badania są ponad dwukrotnie mniejsze.

W 1986 r. odsetek ten wyniósł 13,1%, przy czym najpoważniejszą przyczyną pozostały odmowy (6,3%), znacznie rzadziej przewlekła choroba lub rozpad gospodarstwa domowego (3,2%), utrudniony kontakt (0,9%), zmiana miejsca zamieszkania (1,5%) względnie inne przyczyny (1,1%).

TABL. 3. GOSPODARSTWA DOMOWE PRZERYWAJĄCE BADANIA WEDŁUG PRZYCZYN I GRUP SPOŁECZNO-EKONOMICZNYCH

Gospodarstwa domowe	Ogółem	Przyczyny przerwania badania				
		odmowy	zmiana miejsca zamieszkania	przewlekła choroba, rozpad gospodarstwa	utrudniony kontakt z gospodarstwem	inne przyczyny
	w % ogółu gospodarstw danej grupy					
Ogółem	**13,1**	**6,3**	**1,5**	**3,2**	**0,9**	**1,1**
Pracownicze	14,0	7,8	1,8	2,0	0,9	1,4
Robotniczo-chłopskie	8,9	6,1	0,6	1,0	0,3	0,9
Chłopskie	10,6	6,3	0,6	2,4	0,4	0,9
Emerytów i rencistów	13,6	3,4	1,5	6,6	1,2	0,9

Również pod względem grupy społeczno-ekonomicznej oraz liczby osób w gospodarstwie i poziomu dochodów, grupa ta wykazywała sporą zbieżność z grupą gospodarstw, przystępujących do badania po raz pierwszy.

METODA ELIMINACJI „BRAKU ODPOWIEDZI" W BADANIACH BUDŻETÓW

W badaniach budżetów gospodarstw domowych przyjęta została metoda podmian celowych, tzn. w miejsce gospodarstwa domowego nie podejmującego udziału w badaniu wprowadza się gospodarstwo domowe z tej samej grupy społeczno-ekonomicznej, o podobnej strukturze demograficznej oraz o zbliżonym poziomie dochodów na 1 osobę. Zastosowanie tej metody ułatwia układ operatu losowania, w którym gospodarstwa domowe zostały uszeregowane według grup społeczno-ekonomicznych, następnie według liczby osób w rodzinie i według rosnącego poziomu dochodów (w gospodarstwach powiązanych z rolnictwem indywidualnym, według rosnącego areału gospodarstwa rolnego).

Przyjęcie tej metody w sposób istotny ogranicza wpływ „braku odpowiedzi" na wyniki badania, lecz nie eliminuje go całkowicie. Posiadane informacje o gospodarstwach nie podejmujących badania, choć w stosunku do innych badań są znaczne, tym niemniej nie wystarczają do pogłębionej analizy tej grupy gospodarstw.

Konieczne są specjalne badania, które by dostarczyły wielu innych informacji, jak np. grupy zawodowe, grupy środowiskowe, kręgi kulturowe. Szczególnie głębokiej analizy wymagają gospodarstwa odmawiające udziału w badaniach. Jakie są istotne przyczyny odmów, jaki jest rodowód tej grupy.

Istnieje dość stereotypowe przeświadczenie, że są to, najogólniej mówiąc, osoby z marginesu społecznego, nadużywające alkoholu, osoby z rzadka podejmujące pracę stałą itd. To przekonanie nie jest poparte żadnym rzetelnym badaniem. Jeżeli jednak tak jest w istocie, to tym bardziej grupa ta wymaga dokładnego zbadania.

Badanie odmów jest badaniem bardzo trudnym i złożonym zarówno z powodów wyżej podkreślonych, jak również ze względu na brak, jak dotąd, doświadczenia w tego typu badaniach. Ten pierwszy krok jest jednak konieczny.

PODSUMOWANIE

1. Wpływ „braku odpowiedzi" na wyniki badania, zwłaszcza badania reprezentacyjnego, jest bezsporny. Wpływ ten może być nieistotny, w granicach założonej tolerancji, może jednak być również znaczący, zniekształcający jego wyniki, a niekiedy prowadzić może do stanu bezużyteczności wyników.

2. Przyczyny „braku odpowiedzi" powinny być brane pod uwagę przez badacza na wszystkich etapach badania zarówno w okresie planowania badania, realizacji terenowej, jak i opracowania wyników. Na wszystkich tych etapach należy przedsięwziąć wszelkie możliwe kroki w celu maksymalnego ograniczenia skali „braku odpowiedzi". W tym celu postulować należy, aby każde badanie, przed przystąpieniem do jego realizacji, możliwie szeroko skonsultować w gronie specjalistów.

3. Należałoby przyjąć, jako obowiązującą praktykę w badaniach statystycznych (zarówno pełnych, jak i reprezentacyjnych), zbieranie informacji o jednostkach nie podejmujących badania. W każdej publikacji, omawiającej wyniki badania, powinien się znaleźć choćby niewielki akapit traktujący o „braku odpowiedzi" (odsetek jednostek nie podejmujących badania, kroki jakie badacz przedsięwziął celem ograniczenia wpływu tego zjawiska na wyniki badania oraz dopuszczalny możliwy błąd, jaki użytkownik wyników powinien w analizie uwzględnić).

4. W badaniach o szczególnie szerokim zasięgu społecznym i będących podstawą wielu analiz i decyzji (jak np. badania budżetów gospodarstw domowych), należałoby rozwinąć badania specjalne poświęcone gospodarstwom domowym odmawiającym udziału w badaniach. Wynikiem takich badań powinna być odpowiedź na pytania: kogo nie mamy w badanej zbiorowości, jaka jest skala tego zjawiska, na ile brak określonych grup gospodarstw (osób) zniekształca wyniki badania? itd.

5. Problematyka „braku odpowiedzi" powinna się znaleźć w centrum zainteresowania naukowców i praktyków, zajmujących się badaniami statystycznymi, a w szczególności reprezentacyjnymi. Efektem ich dociekań powinny być metody eliminacji skażenia wyników „brakiem odpowiedzi".

LITERATURA

[1] R. Platek: *A survey practitioner's notion of nonresponse.* Statistiska Centralbyran, Promemorior fran u/ustm, 26
R. Platek, G. B. Gray: *Some aspects of nonresponse adjustment,* Survey Methodology, vo. 11, 1985
[2] J. Kordos: *Dokładność danych w badaniach społecznych,* Biblioteka Wiadomości Statystycznych, GUS, Warszawa 1987, tom 35
[3] *Budżety gospodarstw domowych w 1986 r.,* GUS, Warszawa 1987

Jednorazowe badania ankietowe w Zintegrowanym Systemie Badań Gospodarstw Domowych

mgr Wiesław Łagodziński
Główny Urząd Statystyczny

Zintegrowany System Badań Gospodarstw Domowych (ZSBGD) został pomyślany, jako program scalający różne badania o tematyce społecznej, realizowane na reprezentacyjnej próbie gospodarstw domowych. Założenia tego systemu, opracowane na podstawie dyskusji w czasie 30 posiedzeń roboczych Zespołu Prezesa do Spraw ZSBGD, zawarto w specjalnej publikacji wydanej w serii „Biblioteka Wiadomości Statystycznych"[1]). Prace rozpoczęte w IV kwartale 1982 r. trwają do chwili obecnej.

Równolegle z opracowywaniem założeń metodologicznych i organizacyjnych ZSBGD wprowadzono do realizacji program jednorazowych badań ankietowych. W tym samym czasie zmieniono także system badań budżetów gospodarstw domowych, przechodząc z ciągłych badań rocznych na badania kwartalne, prowadzone metodą rotacyjną[2]).

Założenia Systemu i program badań na lata 1983—1986 zawierały nie tylko opis całej koncepcji i jego genezę, ocenę aktualnego stanu badań i aspekty metodologiczne, ale także charakterystykę systemu i schematu losowania „próby-matki".

---

[1]) *Problemy integracji statystycznych badań gospodarstw domowych,* Warszawa 1987, GUS, „Biblioteka Wiadomości Statystycznych", tom 34, Bibliografia ss. 70—71.

[2]) Szczegółową charakterystykę organizacji i metodologii badań budżetów gospodarstw domowych metodą rotacyjną, patrz: *Metodyka i organizacja badań budżetów gospodarstw domowych* (z praktyki GUS), Warszawa 1986, GUS, Zeszyty Metodyczne. Zasady Projektowania i organizacji badań statystycznych Nr 62. Rozdz. 4 *Metoda rotacyjna badań i ich organizacja* (ss. 31—42). Patrz też bibliografia przedmiotu (ss. 135—136).

Zamieszczono ponadto 19 ekspertyz tematycznych na temat: potrzeb i możliwości prowadzenia różnego typu badań w ZSBGD oraz perspektyw rozwoju Systemu do 1990 r. i po 1990 r.

Wystąpienie moje poświęcone jest nie tyle prezentacji założeń Systemu ile charakterystyce jednorazowych badań ankietowych, które w ramach tego Systemu udało się zrealizować.

Celowość podania tego rodzaju informacji wynika z przekonania, iż rodzaj, ilość, zakres tematyczny i przydatność społeczno-polityczna uzyskanych wyników oraz także i znaczny społeczny zasięg rozpowszechniania publikacji, opracowanych na podstawie tych badań — każe postawić na porządku dziennym problem ich reprezentatywności, a także jakości i dokładności danych oraz precyzji wyników.

Jeśli chodzi o założenia metodologiczne ZSBGD to trudno dokonać syntezy tak obszernego opracowania, liczącego ponad 300 stron. Dla przypomnienia tylko, posługując się nie tylko samym opracowaniem, ale także opracowaniami syntezującymi[3]), przypomnijmy najważniejsze przesłanki, jakie legły u podstaw ZSBGD.

Z założeń Systemu wynika, że integrację badań gospodarstw domowych można osiągnąć poprzez: zastosowanie w badaniach ujednoliconych, zestandaryzowanych pojęć, klasyfikacji i definicji. Standaryzacja owa musiałaby obejmować zarówno badania budżetów, badania ankietowe, jak i masowe badania społeczne (NSP, spisy rolnicze, mikrospisy itd.) oraz badania dotyczące siły roboczej.

Prace nad ujednoliceniem pojęć i klasyfikacji trwają, chociaż już w obecnej fazie dysponujemy ujednoliconą klasyfikacją grup społeczno-zawodowych gospodarstw domowych oraz zestawem podstawowych definicji, dotyczących badań budżetowych. W sposób ciągły także trwają prace nad ujednoliceniem podstawowych klasyfikacji stosowanych w kolejnych badaniach.

Nie jest to ani proces łatwy, ani bezkolizyjny, ani też nie zawsze kończyły się usiłowania te naszym sukcesem. Takie operacyjne uściślenia i ujednolicenia pojęć, przy tak znacznym wachlarzu tematycznym badań oraz tak znacznej liczbie autorów, współautorów i współuczestników badań — są trudne. Osiągnięte jednak wyniki należy uznać za znaczny postęp, a zintegrowanie tego programu w jeden system organizacyjny uważać należy za sukces statystyki państwowej.

Podstawową zasadą metodyczną, przy doborze próbek do badań w badaniach ZSBGD, jest wykorzystanie „próby-matki", co powinno umożliwić badanie współzależności zjawisk i procesów, zachodzących w skali mikro w gospodarstwach domowych.

---

[3]) Jan Kordos: *Badania społeczne w Zintegrowanym Systemie Badań Gospodarstw Domowych Głównego Urzędu Statystycznego*, Warszawa 1988. Referat na konferencję IGS „Badania społeczne i polityka społeczna".

„Próbę-matkę" utworzono na bazie istniejącego w Polsce podziału terytorium kraju na rejony statystyczne i obwody spisowe. Po NSP'1978 utworzony został rejestr rejonów i obwodów, zapisany na taśmie magnetycznej. Rejestr ów został wykorzystany do utworzenia operatu losowania „próby-matki" do ZSBGD. Losowanie było dwustopniowe. Jednostką losowania I stopnia był terenowy punkt badań (tpb, rejon statystyczny, liczący minimum 250 gospodarstw). Dla wylosowania próby tpb zastosowano schemat losowania warstwowego z prawdopodobieństwem wyboru proporcjonalnym do szacunkowej liczby mieszkań w tpb; dążono przy tym do uzyskania próby możliwie bliskiej próbie automatycznie wyważonej. Wylosowana próba została podzielona na 4 podpróby.

Losowanie II stopnia przeprowadzono dwufazowo. W fazie pierwszej wylosowane zostały podpróbki mieszkań do badań demograficznych i społecznych objętych ZSBGD oraz przygotowany został operat losowania próby II stopnia do badań budżetów rodzinnych. Szczegółowy opis procedur losowania do badań oraz rachunkowe uzasadnienie schematu losowania zawiera wspomniane opracowanie B. Lednickiego. Aktualnie trwają przygotowania do zmiany „próby-matki" i powtórnego losowania, w marcu 1989 r.

Trzecim warunkiem osiągnięcia znaczącej i ważącej integracji badań jest wykorzystanie do ich realizacji personelu o odpowiednich kwalifikacjach, stale dokształcanego zarówno w umiejętnościach warsztatowych, jak i w „podbudowie" teoretycznej i metodologicznej prowadzonych badań.

Ostatni problem dotyczy lokalizacji badań ankietowych w systemie organizacyjnym i programie merytorycznym badań budżetowych. Z reguły w badaniach jednorazowych wykorzystuje się, w zależności od celu badania, określoną liczbę podpróbek kwartalnych, a czasem także wszystkie podpróbki w badaniu rocznym. W praktyce oznacza to, że możemy dysponować w badaniach próbą gospodarstw domowych od 2,7 tys. do 30,0 tys. bądź dowolną kombinacją z 4 podpróbek kwartalnych, mieszczącą się w tym przedziale liczbowym. Natomiast w przypadku badania członków gospodarstw domowych, objętych badaniami budżetowymi — w skali roku — badaniami możemy objąć około 80,0 tys. osób we wszystkich kategoriach wieku, a w tym około 46 tys. osób w wieku 15 i więcej lat.

CHARAKTERYSTYKA JEDNORAZOWYCH BADAŃ ANKIETOWYCH ZREALIZOWANYCH W RAMACH PROGRAMU ZSBGD

Liczba zrealizowanych badań

Wyszczególnienie	1982—1988	1982	1983	1984	1985	1986	1987	1988
Badania ankietowe . . .	22	1	2	2	5	3	4	5
Badania warunków bytu ludności	3	1	—	—	1	1	—	—

W latach 1982—1988 zrealizowano 22 jednorazowe badania ankietowe: w 1982 r. — 1, w 1983 r. — 2, w 1984 r. — 2, w 1985 r. — 5, w 1986 r. — 3, w 1987 r. — 4 i w 1988 r. — 5.

W latach 1982—1988 przeprowadzono 3 badania warunków bytu ludności: w 1982 r. (DS-1), w 1985 r. (DS-8) i w 1986 r. (DS-12).

Badania te były związane z przygotowaniem operatu losowania do badań budżetów gospodarstw domowych.

W pierwszym (DS-1, 1982 r.) **badaniu** objęto 60 tys. gospodarstw, w drugim i trzecim ponad 120 tys. gospodarstw. W badaniach tych zastosowano po raz pierwszy mierniki subiektywne, dotyczące oceny sytuacji materialnej gospodarstw domowych; badano także dochody według ich źródeł, warunki mieszkaniowe oraz wyposażenie gospodarstw w przedmioty trwałego użytkowania.

W drugim badaniu (DS-8, 1985 r.) wprowadzono informacje o cechach społeczno-demograficznych członków gospodarstw, sytuacji zdrowotnej i ocenie funkcjonowania służby zdrowia, paleniu papierosów oraz o ocenie sytuacji materialnej i dochodowej. Ponadto badano ocenę jakości i wiarygodności zebranego materiału ankietowego.

W trzecim badaniu (DS-12, 1986 r.) skoncentrowano się głównie na problematyce mieszkaniowej, ale badano także skład gospodarstw domowych oraz ich dochody. Wyniki wszystkich tych badań opublikowano w specjalnych publikacjach tematycznych[4]).

Druga grupa badań (jak dotychczas najliczniejsza, bo obejmująca 11 badań), to badania poświęcone poszczególnym dziedzinom warunków socjalno-bytowych lub określonym zjawiskom (procesom) społecznym.

W 1983 r. zrealizowano badanie wspólnie z SGPiS poświęcone problematyce migracji (DS-3)[5]).

W 1984 r. zrealizowano całoroczne, rotacyjne badanie dobowego budżetu czasu (DS-5). Badaniem objęto 45 tys. osób w wieku 18 lat i więcej, które notowały wykorzystanie czasu w ciągu jednej doby. Badanie prowadzono przez cały rok, a każdego dnia w badaniu uczestniczyło około 1500 osób. Badanie to stanowiło kontynuację badań prowadzonych przez GUS w 1969 r. i w 1976 r. Badania te należą do najczęściej publikowanych i cytowanych spośród badań społecznych GUS, zrealizowanych w ramach ZSBGD. Dotychczas bowiem opublikowano 16 opracowań, dotyczących

---

[4]) *Sytuacja materialna gospodarstw domowych w 1982 r. (w świetle badania ankietowego)*, Warszawa 1984, GUS, Seria: Opracowania Analityczne.
Sytuacja bytowa gospodarstw domowych w 1985 r., Warszawa 1987, GUS Seria: Materiały Statystyczne, nr 42.
Warunki mieszkaniowe gospodarstw domowych oraz wydatki na mieszkanie w 1986 r., Warszawa 1987, GUS, Seria: Materiały Statystyczne, Nr 51.

[5]) Jest to, jak dotychczas, jedyne badanie zrealizowane w ramach ZSBGD, z którego GUS nie tylko nie opublikował żadnych materiałów, ale z którego nie posiada nawet materiałów źródłowych.

dobowego bilansu czasu, 8 opracowań, dotyczących bilansu czasu w wybranych częściach dnia oraz kilkadziesiąt innych opracowań, w których badania te wykorzystano bezpośrednio. Z wydanych opracowań zacytować należy dwa najważniejsze — jedno tabelaryczne i jedno analityczne[6]).

W 1985 r. zrealizowano jedno z dwu badań związanych z turystyką i wypoczynkiem (DS-7, *Wypoczynek 1984*). Przeprowadzono je na początku 1985 r., obejmując nim wszystkie gospodarstwa domowe, uczestniczące w badaniach budżetów rodzinnych w roku poprzednim. Łącznie przebadano 21 tys. gospodarstw. Było to badanie typowo diagnostyczne, zrealizowane na zamówienie Rady do Spraw Rodziny. Zebrano w nim informacje o wypoczynku i formach jego wykorzystania. Wyniki wstępne opublikowano w 1985 r., a ostateczne w 1986 r.[7]).

W 1985 r. zrealizowano pierwsze w ZSBGD (a trzecie w GUS) ankietowe badanie uczestnictwa w kulturze. Badanie zrealizowano w lipcu, obejmując nim 10,6 tys. osób w wieku 15 lat i więcej. Tematyka badań obejmowała uczestnictwo we wszystkich dziedzinach kultury, problematykę zagospodarowania czasu wolnego i zainteresowań osobistych oraz wyposażenia kulturalnego gospodarstw domowych. Wstępne wyniki badań opublikowano w 1986 r., a pełne w 1987 r.[8]).

W 1986 r. zrealizowano drugie badanie, dotyczące turystyki (DS-11, *Uczestnictwo w turystyce*). Badaniem objęto 24,4 tys. osób w wieku 15 lat i więcej, wchodzących w skład 10,8 tys. gospodarstw, uczestniczących w badaniach budżetowych w I półroczu 1986 r. Tematyka badania dotyczyła uczestnictwa we wszystkich dziedzinach i formach turystyki oraz uwarunkowań tego uczestnictwa. W badaniu tym zaistniały dwa nowe elementy różniące je od innych badań (poza badaniem uczestnictwa w kulturze — DS-9 z 1985 r.); **po pierwsze** było to badanie zrealizowane odpłatnie na zlecenie GKKFiT[9]), a **po drugie** w badaniu tym (już po raz drugi) niezależnie zbierano dane „indywidualne" oraz dane o gospodarstwie domowym. Wyniki opublikowano w 1987 r. w syntetycznym opracowaniu tabelarycznym[10]), natomiast komplet tabulogramów (obejmujący ponad 2,5 tys. stron) przekazano do dyspozycji zamawiającego.

Także w 1986 r. (w IV kwartale) na zlecenie (odpłatne) zrealizowano badanie potrzeb żywnościowych. Wyniki badania opublikowano

---

[6]) *Dobowy budżet czasu mieszkańców Polski w 1984 r.*, Część I i II Warszawa 1985, GUS, Seria: Opracowania Statystyczne.
Analiza budżetu czasu ludności Polski w latach 1976 i 1984, Warszawa 1987, GUS, Seria: Studia i Prace, nr 12.

[7]) *Wypoczynek ludności w 1984 roku*, Warszawa 1986, GUS, Seria: Materiały Statystyczne, nr 38.

[8]) *Uczestnictwo w kulturze w 1985 roku*, Warszawa 1987, GUS, Seria: Materiały Statystyczne, nr 45.

[9]) Do zagadnienia odpłatnych badań powrócimy w dalszej części wystąpienia.

[10]) *Uczestnictwo w turystyce ludności Polski w 1986 roku*, Warszawa 1987, GUS, Seria: Materiały Statystyczne, nr 46.

w 1987 r.[11]). Zleceniodawca (SGPiS) oprócz wyników badań otrzymał do dyspozycji zbiór danych.

W I kwartale 1987 r., zrealizowano kolejne badanie zlecone (Instytut Kardiologii), dotyczące zdrowia rodziny (DS-14). Badaniem objęto 21,4 tys. gospodarstw domowych. Miało ono charakter specjalistyczny i spowodowało sporo trudności realizacyjnych. Dotychczas opublikowano tylko wstępne wyniki w formie „notatki sygnalnej". Niestety, wielokrotnie przyjmowane i ustalone terminy wydania publikacji analitycznej nie zostały dotrzymane. Ostatnie ustalenia określają termin wydania tej publikacji na połowę 1990 r.

Ostatnie badania z grupy, jak to nazwaliśmy badań „problemowych" lub „dziedzinowych", to badania zrealizowane już w 1988 r. W czerwcu i lipcu na zlecenie Instytutu Statystyki i Demografii SGPiS zrealizowano autorskie badanie, dr Ewy Frątczak, *pt. Droga życiowa* (DS-20). Badanie dotyczyło biografii rodzinnej, migracyjnej i zawodowej osób w wieku 45 i więcej lat, wchodzących w skład gospodarstw domowych objętych badaniami budżetowymi w II kwartale 1988 r. Opracowanie badań jest w toku. Zakończenie i opublikowanie wyników przewiduje się w maju 1989 r. W czerwcu 1988 r. zrealizowano, na zamówienie ZUS, badanie dotyczące zasiłków rodzinnych (DS-22). Badanie o wąskospecjalistycznym zakresie i przystosowane do bardzo arbitralnej koncepcji zleceniodawcy. Badanie opracowano, a wynik „przejął" zleceniodawca do swego wyłącznego użytku. Trwają przygotowania do wydania publikacji tekstowo-tabelarycznej. Ukaże się ona w II połowie 1989 r.

We wrześniu 1988 r., w ramach podprogramu 03 *Jakość życia a warunki bytu*[12]), zrealizowano autorskie badanie dra J. Rutkowskiego (ZBSE GUS i PAN). Badanie to jest w fazie realizacji terenowej. Opracowanie materiałów na emc i publikację pierwszych wyników przewiduje się w czerwcu 1989 r.

Trzecia grupa jednorazowych badań ankietowych realizowanych w programie ZSBGD to jednoroczne badania ankietowe określonych grup lub kategorii społecznych. Badania te nazywamy w skrócie „środowiskowymi". W latach 1983—1987 zrealizowano 5 takich badań, a aktualnie trwają przygotowania do szóstego, które będzie zrealizowane w I kwartale 1989 r. W serii tej zrealizowano, w 1983 r. (DS-2) i 1987 r. (DS-15), badania sytuacji społeczno-zawodowej kobiet. W pierwszym badaniu wykorzystano podpróbkę II kwartału (około 5,0 tys. kobiet w wieku 18—59 lat)[13]). Drugie badanie

---

[11]) *Potrzeby żywnościowe gospodarstw domowych*, Warszawa 1987, GUS, Seria: Materiały Statystyczne, nr 49.

[12]) Doc. dr hab. Jan Kordos, jako Prezes Rady Głównej Polskiego Towarzystwa Statystycznego, jest koordynatorem II° programu 03 w Centralnym Programie Badań Podstawowych 09.09 „Polityka społeczna w PRL". Koordynatorem I stopnia jest prof. dr hab. Stanisław Czajka — Dyrektor Instytutu Polityki Społecznej na Uniwersytecie Warszawskim. Badanie nt. jakości życia jest jednym z 14 tematów realizowanych w programie problemu 03 (*Jakość życia a warunki bytu*) w latach 1987—1990.

[13]) *Sytuacja społeczno-zawodowa kobiet w 1983 r.*, Warszawa 1984 r. GUS, Seria: Materiały Statystyczne, nr 21.

przeprowadzono w II kwartale 1987 r. (objęło ono około 4,8 tys. kobiet w wieku 15—59 lat). Problematyka obu badań była bardzo podobna, dotyczyła ona sytuacji rodzinnej, organizacji życia domowego, obowiązków domowych, pracy zawodowej, kształcenia itd. Wyniki tego badania zawiera publikacja, która ukaże się w I kw. 1989 r.[14]).

W 1984 r. (DS-4) i w 1987 r. (DS-19) przeprowadzono badania sytuacji socjalno-bytowej młodych małżeństw. Badaniami objęto te małżeństwa, w których jedno z małżonków nie ukończyło 30, a drugie 35 roku życia. Badaniami objęto około 3,2 tys. małżeństw w 1984 r. oraz 2,5 tys. w 1987 r. Tematyka badań dotyczyła cech społeczno-demograficznych małżeństwa, sytuacji materialnej i mieszkaniowej, wyposażenia gospodarstw domowych, wykorzystania czasu wolnego, dążeń, potrzeb i aspiracji. Wyniki pierwszego badania już opublikowano[15]), natomiast wyniki drugiego są w druku.

W tej serii badań zrealizowano także dwa badania, dotyczące grup o szczególnie zagrożonej pozycji socjalno-bytowej. W I kwartale zrealizowano badanie, dotyczące sytuacji bytowej ludzi starszych (DS-6), objęto badaniem osoby w wieku 60 lat i więcej. Zbadano łącznie 9,5 tys. osób, wchodzących w skład gospodarstw objętych badaniami budżetowymi. Badanie dotyczyło nie tylko sytuacji bytowej, ale także stanu zdrowia, uzyskiwanej pomocy, warunków mieszkaniowych oraz subiektywnej oceny przez badanych sytuacji życiowej. Wyniki badania opublikowano w 1985 r.[16]). Następne badanie planowane jest w 1989 r.

Drugie ze wspomnianych badań, to zrealizowane w I kwartale 1986 r. badanie sytuacji bytowej rodzin wielodzietnych i niepełnych z dziećmi (DS-10). Zbadano 2,2 tys. rodzin wielodzietnych oraz około 1,1 tys. rodzin niepełnych z dziećmi. Tematyka dotyczyła wszystkich aspektów warunków bytowych, a w tym szczególnie zagrożeń socjalnych i obciążeń życiowych. Wyniki badania opublikowano w 1987 r.[17]).

Ostatnia, 3 grupa, to badania problemowo-środowiskowe, w skrócie nazywane kombinowanymi, w których wyodrębnione, według ściśle określonych kryteriów, grupy badanych; obserwowane ze względu na określone sytuacje, zachowania lub reakcje.

W II połowie 1987 r. zrealizowano, na zamówienie AWF, autorskie badanie doc. dra T. Łobożewicza nt. *Turystyki i rekreacji osób w starszym wieku* (DS-16). Łącznie zbadano około 2,7 tys. osób w wieku 60 lat i więcej. W badaniu wykorzystano podpróbkę gospodarstw domowych, uczest-

---

[14]) *Sytuacja społeczno-zawodowa kobiet w 1987 roku, Warszawa 1988 r.* GUS, Seria: Materiały Statystyczne.

[15]) *Sytuacja socjalno-bytowa młodych małżeństw w 1984 r.* Warszawa 1984 r., GUS, Seria: Materiały Statystyczne, nr 26.

[16]) *Sytuacja bytowa ludzi starszych w 1985 roku*, Warszawa 1985, GUS, Seria: Statystyka Polski, nr 36.

[17]) *Sytuacja bytowa rodzin wielodzietnych i niepełnych z dziećmi* Warszawa 1987 r., GUS, Seria: Materiały Statystyczne, nr 41.

niczących w badaniach budżetowych w II kwartale 1987 r. Publikacja zawierająca wyniki z tych badań jest w druku.

W IV kwartale zrealizowano autorskie badanie W. Łagodzińskiego (w ramach podprogramu 03 [12]) warunków życia i potrzeb młodzieży. Łącznie badaniem objęto około 4,2 tys. osób w wieku 15—29 lat, które bądź pozostały w rodzinach, bądź prowadzą samodzielne gospodarstwa domowe. Tematyka badania dotyczyła warunków i potrzeb mieszkaniowych, kształcenia i potrzeb kształceniowych, pracy zawodowej, uczestnictwa w kulturze i turystyce, przynależności organizacyjnej oraz subiektywnych ocen progów i barier awansu życiowego młodzieży. Publikacja, zawierająca wyniki tego badania, jest w fazie przygotowania do druku.

I wreszcie ostatnie badanie z tej serii, to badanie zrealizowane w I kwartale 1988 r. nt. *Faza rozwoju rodziny*, przygotowane przez zespół pracowników SGPiS pod kierunkiem doc. dr. hab. A.Kurzynowskiego. Badanie to jest w fazie opracowania maszynowego, a w I kwartale 1989 r. przewiduje się opublikowanie raportu z badań.

OCENA EFEKTÓW REALIZACJI BADAŃ ANKIETOWYCH W ZSBGD W LATACH 1983—1987

Pierwszy okres realizacji praktycznej programu jednorazowych badań ankietowych w oparciu o próbkę gospodarstw domowych miał spełnić i spełnił podstawowe cele rozwoju systemu badań.

Realizując owe 22 badania wypełniono dotkliwą lukę informacyjną w zakresie problematyki warunków socjalno-bytowych dostarczając organom władzy, nauce i opinii publicznej znaczną ilość informacji. Pozwoliło to z kolei na znaczne osłabienie nacisku, przede wszystkim na GUS.

Pozwoliło to także na eksperymentalne wdrożenie do praktyki tej części ZSBGD, która zakładała wypełnienie luki metodologicznej, programowej i organizacyjnej w zakresie badań społecznych. Wypełniła się jednak przede wszystkim luka informacyjna w zakresie badań warunków socjalno-bytowych, łącząca z jednej strony badania spisowe, a z drugiej badania budżetowe.

Tempo, w jakim realizowano program tych badań, a przede wszystkim ich ilość spowodowała, że w znacznym stopniu od programu realizacji, opracowania i prezentacji „odstał" program refleksji teoretycznej i analizy statystycznej wyników prowadzonych badań, GUS nie jest jednak placówką mającą komfort realizacyjny placówek naukowych, a przy tym brak mu wystarczającego wsparcia teoretycznego ze strony tych placówek. To wszystko spowodowało, iż pilną stała się sprawa pogłębionej oceny reprezentacyjności i precyzji badań oraz analizy porównawczej wyników.

Praktycznie 1989 r. zamknie ów okres burzliwej realizacji prawie każdego zamówienia (nieomal natychmiast), szybkiego opracowania wyników oraz ich rozpowszechniania (przygotowywana w ciągu 2—4 miesięcy informacja wstępna, opracowanie wyników w 6 miesięcy po badaniu i wydanie publikacji finalnej w 12—15 miesięcy po badaniu). Realizacja tego programu pozwoliła także wypracować system organizacyjny przygotowania, pilotowania, realizacji i opracowywania maszynowego badań. Ogromną rolę, w zakresie organizacji, odegrało Kierownictwo GUS, które stwarzając warunki organizacyjne i materialne PTS umożliwiło rozwój systemu badań „częściowo odpłatnych".

Forma organizacyjna Biura Badań i Analiz Statystycznych przy Radzie Głównej PTS stała się jedyną szansą realizacji 7 z 9 badań zrealizowanych w ramach ZSBGD, w latach 1987—1988.

Obok problemów, wynikających z braku teoretycznej „podbudowy" i statystycznej, pogłębionej analizy, a w tym przede wszystkim analizy reprezentatywności, brak uporządkowania w tematyce badań budzi pewien niepokój oraz pojawianie się problemów przypadkowych.

Ze spraw wymagających dalszej pracy należy wymienić sprawę doskonalenia terenowego aparatu realizacyjnego oraz doskonalenie procedur przygotowania narzędzi do badań.

W pierwszym przypadku jest to sprawa rozwoju szkolenia centralnego, prowadzenia szkolenia regionalnego ankieterów oraz przygotowania odpowiednich materiałów szkoleniowych, a w tym przede wszystkim materiałów audiowizualnych.

W drugiej natomiast sprawie — szansą jest informatyzacja procesu przygotowania formularzy do badań (chociażby przez wdrożenie do praktyki systemu BLAiSE).

Ważną i mającą duże znaczenie praktyczne jest kwestia ustabilizowania systemu konsultacji poprzedzających przygotowywane badanie. W dotychczasowych badaniach najczęściej konsultantami były instytucje i organizacje rządowe i społeczne. Za mały jest udział (tak w każdym razie widzę to z perspektywy ostatnich 3—4 lat) przedstawicieli placówek naukowych. Na usprawiedliwienie dodać trzeba, że większość placówek nie chce w takim tempie uczestniczyć w przygotowaniu badań, ponieważ nie daje im szans na rzeczową analizę omawianych materiałów. Nie jest to jednak problem poważny, a już bieżący rok wskazuje na nawiązanie rzeczowych i bardzo pożytecznych form współpracy z wieloma ważnymi dla nas placówkami (między innymi przez udział PTS w realizacji CPBP 09.09).

W reasumpcji stwierdzić należy, że realizacja programu jednorazowych badań ankietowych w ramach programu ZSBGD, wsparta w odpowiedni sposób pracami o charakterze teoretycznym, może się utrwalić jako jeden z najbardziej efektywnych systemów badań społecznych w Polsce.

BIBLIOGRAFIA

Ważniejsze opracowania, dotyczące Zintegrowanego Systemu Badań Gospodarstw Domowych oraz program badań budżetów rodzinnych.

1. *Problemy integracji statystycznych badań gospodarstw domowych*, Warszawa 1987. Biblioteka Wiadomości Statystycznych, tom 34.
2. *Metodyka i organizacja badań budżetów gospodarstw domowych (z praktyki GUS)*, Warszawa 1986. GUS, Zeszyty Metodyczne. Zasady projektowania i organizacji badań statystycznych nr 62.
3. *Podstawowe definicje i klasyfikacje stosowane w statystyce gospodarstw domowych*, Warszawa 1988. GUS, Zeszyty metodyczne. Klasyfikacje, nomenklatury i wskaźniki nr 73.
4. *Instrukcja w sprawie organizacji i sprawozdawczości z badań budżetów gospodarstw domowych metodą rotacyjną w latach 1986–1990*. Warszawa 1985. Główny Urząd Statystyczny. Departament Badań Społecznych.
5. Jan Kordos: *Dokładność danych w badaniach społecznych*. Warszawa 1987. GUS. Biblioteka Wiadomości Statystycznych, tom 35. „Wiadomości Statystyczne".
6. Jan Kordos: *Metoda rotacyjna w badaniach budżetów gospodarstw domowych* (nr 9, 1982).
7. Jan Kordos: *Możliwości zbudowania ZSBGD w Polsce* (nr 9, 1983).
8. Jan Kordos: *Metodologiczne aspekty ZSBGD* (nr 3, 1984).
9. Jan Kordos: *Organizacyjne aspekty ZSBGD* (nr 7, 1986).
10. Elżbieta Krzyżanowska: *Koncepcja badania siły roboczej w ramach ZSBGD* (nr 12, 1985).
11. Jan Rutkowski: *Podstawowe pojęcia statystyki społecznej* (nr 10, 1984).

Problematyka badania budżetów gospodarstw domowych metodą rotacyjną oraz innych badań społecznych w świetle doświadczeń Wojewódzkiego Urzędu Statystycznego w Gdańsku

mgr Stanisława Ryckowska
Wojewódzki Urząd Statystyczny w Gdańsku

METODA BADAŃ

Badania budżetów gospodarstw domowych metodą rotacyjną, z kwartalnym okresem wymian rodzin w roku, prowadzone są od 1982 r.

Wprowadzenie do badań budżetów rotacji gospodarstw domowych, pozwoliło, przy zachowaniu w zasadzie tych samych środków finansowych, na ponad dwukrotne zwiększenie liczby badanych rodzin w porównaniu do

stosowanej dotąd metody ciągłej, w której gospodarstwa domowe uczestniczą w badaniu przez cały rok lub dłużej.

W badaniach prowadzonych metodą rotacyjną, podobnie jak w badaniu metodą ciągłą, uczestniczą 4 następujące społeczno-ekonomiczne grupy gospodarstw domowych: pracowników, chłopów, robotniko-chłopów, emerytów i rencistów.

DOBÓR GOSPODARSTW DO BADAŃ

Stosowana jest metoda losowego doboru rodzin do badań. Operat losowania jednostek stanowi wykaz rejonów statystycznych oraz zamieszkanych w nich mieszkań, opracowany dla potrzeb Narodowego Spisu Powszechnego. Rozpoczęcie w 1982 r. badań budżetów metodą rotacyjną poprzedzone zostało pracami przygotowawczymi, związanymi z wymianą próby gospodarstw domowych. Całość prac przygotowawczych można by podzielić na 4 etapy.

I etap. Losowanie przez GUS rejonów statystycznych; WUS sporządzają wykazy zamieszkanych mieszkań w wylosowanych rejonach.

II etap. WUS podają liczbę spisanych mieszkań w poszczególnych rejonach; GUS losuje w rejonie po 150 mieszkań, w których przeprowadzona będzie ankietyzacja rodzin, czyli wywiady wstępne.

III etap. WUS przeprowadzają ankietyzację rodzin w wylosowanych mieszkaniach, następnie po odpowiednim powarstwowaniu i pogrupowaniu ankiet sporządzają dla każdego rejonu wykaz gospodarstw domowych przygotowanych do badania.

IV etap. WUS podają liczbę zankietyzowanych gospodarstw domowych w poszczególnych terenowych punktach badań (tpb); GUS losuje jednostki do badania budżetów.

W każdym tpb GUS losuje tzw. podpróbkę stałą, liczącą 4 rodziny w kwartale (4×4 kw. $= 16$ rodzin) i podpróbkę zmienną, składającą się z 2 rodzin (4 kwartały $\times$ 2 rodziny $\times$ 4 lata $=$ 32 rodziny). Na cały czteroletni cykl badań wylosowanych jest więc 48 gospodarstw domowych. Pozostałe gospodarstwa domowe, znajdujące się na sporządzonym wykazie rodzin zankietyzowanych, stanowią rezerwę w przypadku, gdy któreś z gospodarstw wylosowanych nie wyrazi chęci przystąpienia do badań. Liczba rodzin rezerwowych wydawałaby się dość znaczna i wystarczająca dla zapewnienia ciągłości badań. Jednak w tych tpb, gdzie zankietyzowana została, z różnych względów, mniejsza liczba niż 150 gospodarstw, występują trudności zarówno w doborze rodzin do badania, jak i zachowaniu struktury wylosowanych rodzin, gdyż liczebność danej grupy gospodarstw domowych jest niewielka.

W praktyce obserwuje się ponadto znaczną liczbę odmów rodzin uczestniczących w badaniach. Z tych względów zachodziłaby konieczność zwiększenia liczby mieszkań losowanych, w ramach prac przygotowawczych do badań, ze 150 do np. 200 mieszkań.

ORGANIZACJA BADANIA

Na terenie poszczególnych województw badania budżetów gospodarstw domowych prowadzą wojewódzkie urzędy statystyczne, a ściślej wchodzące w ich skład oddziały badań społecznych. Nadzór merytoryczny nad przebiegiem tych badań sprawuje Departament Badań Społecznych GUS.

Bezpośredni kontakt z rodzinami wylosowanymi do badań mają instruktorzy. Są nimi etatowi pracownicy wojewódzkich urzędów statystycznych i dodatkowo od 1986 r., w ramach rozszerzenia badań, osoby zatrudnione na podstawie umów, pełniące funkcję instruktorów.

Instruktorzy etatowi mają przydzielone dwa tpb, współpracują w ciągu kwartału z 12 rodzinami, po 6 rodzin w każdym tpb. Oznacza to, że instruktor organizuje w ciągu roku badanie i sprawuje nadzór nad 48 rodzinami.

Osoby pełniące funkcję instruktorów prowadzą badania w dodatkowych wiejskich tpb, utworzonych w ramach rozszerzania badań i zwiększenia liczby gospodarstw domowych, związanych z rolnictwem. Osoba pełniąca funkcję instruktora organizuje i prowadzi badania w jednym tpb.

W województwie gdańskim badania budżetów gospodarstw domowych prowadzone są w 42 tpb. Zatrudnionych jest 17 instruktorów etatowych, którzy prowadzą badania w 34 tpb (każdy po 2 tpb) oraz 8 osób pełniących funkcję instruktora (każdy po 1 tpb).

Ilość punktów badań oraz liczbę rodzin objętych badaniem w poszczególnych miastach i na terenie gmin przedstawia poniższe zestawienie:

Wyszczególnienie	Ilość tpb	Liczba rodzin	
		w kwartale	w roku
Województwo	**42**	**252**	**1008**
Miasta	26	156	624
Gdańsk	13	78	312
Gdynia	6	36	144
Pruszcz	1	6	24
Starogard	2	12	48
Tczew	2	12	48
Wejherowo	2	12	48
Gminy	16	96	384
Gniew	1	6	24
Kartuzy	1	6	24
Karsin	1	6	24
Kolbudy	1	6	24
Kościerzyna	1	6	24
Pelplin	1	6	24

(dok.)

Wyszczególnienie	Ilość tpb	Liczba rodzin	
		w kwartale	w roku
Gminy (dok.)			
Pruszcz	2	12	48
Puck	3	18	72
Przywidz	1	6	24
Somonino	1	6	24
Sulęczyno	1	6	24
Stara Kiszewa	1	6	24
Żukowo	1	6	24

Badania obejmują 252 rodziny w kwartale. Wyniki roczne uzyskuje się natomiast od 1008 rodzin. W 26 punktach miejskich badania obejmują 624 rodziny w ciągu roku i w 16 punktach wiejskich — 384 rodziny.

W porównaniu do ogólnej liczby gospodarstw domowych badanych w kraju — 29664 rodziny, województwo nasze pod względem liczby badanych gospodarstw znajduje się na trzecim miejscu, po województwie katowickim, w którym badanych jest 2736 rodzin i stołecznym warszawskim — 1776 rodzin.

ZAGADNIENIE ODMÓW W BADANIACH BUDŻETÓW RODZINNYCH

Badanie budżetów gospodarstw domowych jest jednym z najtrudniejszych badań prowadzonych przy pomocy metody reprezentacyjnej.

Trudności w prowadzeniu badań wynikają przede wszystkim z kilku przyczyn:

— gospodarstwa domowe udzielają informacji na zasadzie dobrowolności,
— badanie dotyczy spraw drażliwych i niechętnie ujawnianych, tj. dochodów i wydatków,
— wymagane jest od rodzin prowadzenie systematycznych zapisów.

Z tych przyczyn wynika, że w badaniu budżetów gospodarstw domowych występuje problem odmów przystąpienia do badania przez wylosowaną rodzinę. Ponadto, częste w ostatnim okresie, podwyżki cen na rynku podważają w niektórych rodzinach sens prowadzenia takich badań, jak też utrudniają rejestrowanie wydatków w aktualnych cenach.

Z dotychczasowej, siedmioletniej praktyki prowadzenia badań metodą rotacyjną w naszym województwie wynika, że spory odsetek gospodarstw domowych, z różnych przyczyn, nie podejmuje badań.

Stosunkowo najczęściej nie podejmują badania gospodarstwa najmniejsze — jednoosobowe. Dotyczy to wszystkich społeczno-ekonomicznych grup gospodarstw domowych objętych badaniem.

Na ogólną liczbę 252 wylosowanych gospodarstw domowych do badań w kwartale, około 80 gospodarstw jest zamienianych przez rodziny z listy rezerwowej. Oznacza to, że **w każdym kwartale**, w zasadzie, **co trzecia rodzina wylosowana** do badań budżetów **odmawia przystąpienia do badań.**

Ten lub zbliżony odsetek odmów gospodarstw wylosowanych do badań budżetów, obserwuje się na przestrzeni wszystkich lat badań prowadzonych metodą rotacyjną, tj. od 1982 r.

FUNKCJA INSTRUKTORA A PROBLEMY KADROWE

W badaniu budżetów gospodarstw domowych metodą rotacyjną, pełnienie funkcji instruktora nie jest łatwe. W sytuacji, gdy rodziny niechętnie uczestniczą w badaniu, a losowy dobór rodzin do badania powoduje rozrzut adresów zamieszkania, werbowanie na każdy kwartał rodzin do prowadzenia badań stanowi dla instruktora najtrudniejszy problem.

Z drugiej zaś strony, pracownik zatrudniony na stanowisku instruktora badań budżetów rodzinnych winien mieć szczególne predyspozycje do łatwego nawiązywania kontaktu z ludźmi, być dobrym psychologiem i statystykiem. Od jego umiejętnej współpracy z rodziną. możliwości zdobycia zaufania oraz wnikliwej analizy prowadzonych zapisów budżetowych, w niemałym stopniu zależy jakość uzyskanych informacji o badanych rodzinach.

W ostatnim okresie szczególnie duże trudności kadrowe występują w miastach Gdańsku i w Gdyni oraz w tpb, prowadzonych systemem pracy zleconej.

Nowo zatrudnieni instruktorzy, po poznaniu w praktyce charakteru pracy, a szczególnie „niewdzięcznej" roli na etapie werbowania rodzin do badania, gdzie spotykają się z częstymi odmowami, a niekiedy przykrościami, zniechęcają się do dalszej pracy i odchodzą.

Wśród zatrudnionej kadry instruktorów w naszym województwie, spora część (8 osób, na 17 instruktorów etatowych) pracuje na tym stanowisku od rozpoczęcia badań metodą rotacyjną, tj. od 1982 r. Osoby te przeszły okres „próby", zdobyły doświadczenia w wykonywanej pracy, a także przyzwyczaiły się do charakteru pracy, jakim jest praca terenowa i praca w domu, bez codziennego dojazdu do zakładu pracy.

BADANIA ANKIETOWE

Wprowadzenie rotacyjnej metody badań budżetów rodzinnych stworzyło możliwość przeprowadzania, na odrębnej w każdym kwartale podpróbce rodzin, dodatkowych ankiet o różnej tematyce społecznej, rozszerzającej zakres informacji o gospodarstwach domowych. Te dodatkowe badania

ankietowe przeprowadzane są w ramach wdrażania ogólnokrajowego Zintegrowanego Systemu Badań Gospodarstw Domowych (ZSBGD).

Zgodnie z założeniami pierwszego etapu wdrażania systemu, który obejmował lata 1982—1985, zrealizowano tzw. program „minimum" badań. Były to badania: społeczno-zawodowej sytuacji kobiet, przyczyn migracji, sytuacji socjalno-bytowej młodych małżeństw, budżetu czasu dorosłych mieszkańców kraju, sytuacji bytowej ludzi starszych, form wypoczynku, uczestnictwa w kulturze, uczestnictwa w turystyce, a także sytuacji bytowej gospodarstw domowych.

Drugi etap wdrażania zintegrowanego systemu, obejmuje lata 1986—1990. Przeprowadzone dotąd, tj. do 1988 r. badania ankietowe dotyczyły następującej tematyki: stanu zdrowia rodzin i ich potrzeb żywnościowych, sytuacji społeczno-zawodowej kobiet, warunków życia i potrzeb młodzieży, turystyki i rekreacji ludzi w starszym wieku, faz rozwoju rodziny i jej potrzeb oraz drogi życiowej. Ponadto, na podpróbce rodzin badanych w III kwartale 1988 r. przewiduje się badania nt. jakości życia rodzin.

W prowadzonych badaniach ankietowych, na próbie rodzin objętych badaniem budżetów, obserwuje się pewien niekorzystny wpływ tych dodatkowych badań na uczestnictwo rodzin w badaniu budżetów. Rodziny, które bez większego zainteresowania i przekonania (często wręcz na usilne prośby instruktora) przyjęły badania budżetów, dodatkowo wypytywane w formie ankiet na różne tematy życia osobistego, rodzinnego, niekiedy dotyczące spraw drażliwych, zniechęcają się do badań, odmawiając podjęcia badań budżetów w analogicznych kwartałach następnych lat. Z tego względu, **prowadzenie badań ankietowych na innej, nowej próbie rodzin, nie objętej badaniem budżetów, byłoby, jak się wydaje, lepszą formą prowadzenia badań dodatkowych.**

*
* *

Przedstawione ważniejsze problemy realizacji w terenie badania budżetów gospodarstw domowych i badań społecznych, wskazują, że badania te są badaniami trudnymi. Wymagają dużego wysiłku organizacyjnego, odpowiednio przygotowanego personelu nadzoru, stałej, fachowej kadry instruktorów. Prowadzenie tych badań równolegle przez ten sam personel (instruktorów, inspektorów), szczególnie w okresach spiętrzeń pracy, powoduje znaczne przeciążenia pracowników, a w konsekwencji stwarza większe możliwości pomyłek i niedokładności w obu badaniach.

Rotacyjna metoda badań jest o tyle dogodna, że pozwala na przeprowadzanie na odrębnej, kwartalnej podpróbce rodzin, dodatkowych badań ankietowych. Jednakże prowadzenie badań społecznych na nowej, innej podpróbce rodzin (wylosowanych specjalnie do tego rodzaju badań) nie obciążałoby rodzin uczestniczących w badaniu budżetów. Zbieranie zbyt

wielu informacji od tych samych rodzin, wpływa bowiem na zniechęcenie pewnej części rodzin i ich rezygnację z uczestniczenia w badaniu budżetów w kolejnych okresach.

Wprowadzenie rotacji gospodarstw domowych w badaniu budżetów pozwala (w zasadzie przy takich samych nakładach) na znaczne zwiększenie liczby rodzin badanych w stosunku do metody ciągłej. Jednak, mimo skrócenia w tej metodzie okresu prowadzenia zapisów przez rodzinę z 12 do 3 miesięcy w roku, występuje nadal problem odmów uczestniczenia w badaniach wylosowanych rodzin. Mimo stosowania pewnych zasad, przy dokonywaniu zmian rodzin, jest to zjawisko niekorzystne.

Z praktyki prowadzonych badań wynika, że spora liczba rodzin prowadziłaby zapisy budżetu domowego przez dłuższy aniżeli kwartalny okres (często ze względu na przyzwyczajenie się i „wciągnięcie" do prowadzenia zapisów). Z kolei, gdy chodzi o podjęcie badań w analogicznym kwartale następnego roku — a więc po rocznej przerwie, wiele rodzin nie wyraża już zgody na ponowne podjęcie badań.

Występujące trudności w cokwartalnym doborze rodzin do badania budżetów, duża liczba odmów rodzin z tzw. podpróbki stałej, która winna uczestniczyć w analogicznych kwartałach 4-letniego cyklu badań, a tym samym zmniejszania się badanej populacji rodzin służącej do porównań danych w czasie — przemawiałyby za przyjęciem dłuższego okresu wymian rodzin.

Badanie budżetów gospodarstw domowych w stosunkowo krótkim, kwartalnym okresie, nie odzwierciedla pełnego poziomu i struktury wydatków (przychodów), jakie mają miejsce w ciągu roku. Biorąc pod uwagę występującą sezonowość wydatków (przychodów) w budżetach gospodarstw domowych należałoby — doskonaląc metodę badania budżetów rodzinnych, rozważyć słuszność przyjętej zasady uczestniczenia wylosowanych rodzin zawsze w tych samych analogicznych kwartałach kolejnych lat.

Inny problem, to doskonalenie sposobu opracowywania zebranych informacji. Chodzi tu o znacznie większe zastosowanie maszyn elektronicznych, gdzie podstawowym dokumentem byłoby odpowiednio przygotowane (sprawdzone i zasymbolizowane w zakresie przychodów i wydatków) książeczki budżetowe rodzin. Dotychczasowy sposób opracowywania budżetów gospodarstw domowych jest bowiem zbyt pracochłonny. Wymaga przenoszenia często tych samych danych na kilku formularzach (W-02, W-03, ark. analitycznym do W-03, a także częściowo na W-04K i W-04R).

BIBLIOGRAFIA

1. J. Kordos: *Metoda rotacyjna w badaniach budżetów gospodarstw domowych w Polsce*, Wiadomości Statystyczne nr 9, 1982 r.
2. J. Kordos: *Organizacyjne aspekty zintegrowanego systemu badań gospodarstw domowych*, Wiadomości Statystyczne nr 7, 1986 r.

3. M. Daszyńska: *Zintegrowany system badań gospodarstw domowych (założenia metodyczne i organizacyjne, etapy wdrażania)*, Wiadomości Statystyczne nr 1, 1988 r.
4. D. Śleszyńska: *Trudne badania budżetów gospodarstw domowych*, Wiadomości Statystyczne nr 6, 1984 r.
5. *Metoda i organizacja badań budżetów gospodarstw domowych*, Zeszyty Metodyczne nr 62, GUS, Warszawa 1986 r.

Zastosowanie metody reprezentacyjnej w statystyce rolniczej

mgr inż. Florian Ruszkowski
Główny Urząd Statystyczny

Statystyka rolnicza w zasadniczy sposób różni się od innych działów statystyki, co warunkuje system stosowanych metod w ocenie zjawisk statystyczno-ekonomicznych w gospodarce żywnościowej.

Do warunków obiektywnych, mających bardzo duży wpływ na kształtowanie się poziomu ocen produkcji rolniczej, należy zaliczyć przede wszystkim:
— wielkość podmiotów gospodarowania (gospodarka nieuspołeczniona, gospodarka państwowa, gospodarka spółdzielcza),
— bazę produkcyjno-ekonomiczną (baza inwestycyjna, umaszynowienie, wyposażenie w środki produkcji, zasoby siły żywej, zasady ekonomiczno-finansowe gospodarowania itp.),
— znaczną zmienność przyrodniczo-środowiskową i glebowo-terytorialną, mającą silny związek z produkcyjnością gospodarstw,
— ścisłe uzależnienie poziomu produkcji roślinnej od trudnego do przewidzenia przebiegu warunków agrometeorologicznych,
— niekorzystną strukturę agrarną indywidualnych gospodarstw rolnych,
— wielość ocenianych ziemiopłodów (gatunków, odmian i sposobów użytkowania),
— niedoskonałość metod sprawdzalnych.

Do warunków subiektywnych m.in. należy zaliczyć:
— duże zróżnicowanie gospodarstw pod względem kultury rolnej,
— nieufność lub tendencyjność niektórych producentów rolnych do podawania nieprecyzyjnych informacji statystycznych zarówno z zakresu produkcji roślinnej, jak i zwierzęcej.

W gospodarce uspołecznionej podstawą pozyskiwania informacji statystycznych jest sprawozdawczość z gospodarstw. Trudności metodyczne

pogłębiają się, gdy ocenia się zjawiska rolnicze, występujące w gospodarce nieuspołecznionej. Gospodarka ta stanowi przeciętnie w skali kraju około 80% użytków rolnych.

Przy rozpatrywaniu metod stosowanych w statystyce produkcji rolniczej w toku należy wyróżnić metody stosowane w ocenie produkcji roślinnej i produkcji zwierzęcej.

PRZEGLĄD STOSOWANYCH METOD W STATYSTYCE ROLNICZEJ

W zakresie oceny produkcji roślinnej w toku wykorzystywanych jest w praktyce równolegle kilka metod.

Należy wymienić następujące metody:

▲ ekspertyzy rzeczoznawców i specjalistów; oceny dokonywane przez specjalistów (gminnych, wojewódzkich), na podstawie gruntownej znajomości terenu, czynników plonotwórczych, analizy porównawczej. Zaletą jej jest niski koszt i mała pracochłonność, a także prosta organizacja. Ocena może być jednakże mało precyzyjna i tendencyjna, gdy dokonujący jej ma małe doświadczenie praktyczne, bądź pracę wykonuje niesolidnie;

▲ metoda rozliczeń bilansowych, która oparta jest na sprawdzeniu prawidłowości szacunków, przy wykorzystaniu zasady bilansowej rozliczania produkcji, uwzględniającej podstawowe elementy przychodów i rozchodów produkcji podstawowych ziemiopłodów;

▲ metoda oceny zbiorów na podstawie wyników kombajnowania, która polega na ocenie wielkości zbiorów roślin, dokonywanych przy zastosowaniu kombajnów. Metoda ta jest dość pracochłonna, a jej wyniki zależą od podejścia kombajnistów do „akcji";

▲ metoda pomiarów biometrycznych, polegająca na pomiarze podstawowych elementów plonu. Główną zaletą tej metody jest, że poszczególne składniki struktury plonu są mierzone, a nie ustalane wizualnie. Ostateczny natomiast efekt jest zależny od: właściwego wyboru gospodarstwa, określenia odpowiedniej techniki oceny dla właściwej fazy fenologicznej roślin i solidnego przeprowadzenia pomiarów. Do metod biometrycznych należy także zaliczyć próbne zbiory, tzn. zbiory dokonywane na określonej małej powierzchni (np. 1 m^2), przed właściwym sprzętem badanego ziemiopłodu.

Niezwykle ważnymi źródłami danych w statystyce rolniczej są: spisy rolnicze oraz informacje korespondentów rolniczych i ogrodniczych.

Przeprowadzany corocznie, na przełomie czerwca i lipca, pełny spis rolniczy dostarcza informacji o użytkowaniu gruntów, powierzchni zasiewów i pogłowia zwierząt. Wśród dodatkowych tematów spisu można wymienić: spis maszyn i urządzeń rolniczych, budynków inwentarskich, zaopatrzenia rolnictwa w wodę.

Informacje korespondentów, datujące się od przełomu lat 1918 i 1919 zapewniają dane, dotyczące: produkcji polowej, produkcji zwierzęcej, cen w obrocie wolnorynkowym, kosztów usług, płac robotników sezonowych.

METODA BADANIA REPREZENTACYJNEGO

Wśród stosowanych metod w ocenie produkcji rolniczej szczególne znaczenie ma metoda reprezentacyjna. W ocenie produkcji rolniczej wszelkie badania na gruncie, bezpośrednio w gospodarstwach, stanowią podstawę jakichkolwiek informacji na ten temat. Nie ma oficjalnej sprawozdawczości podpisanej przez indywidualnego użytkownika gospodarstwa rolnego. Wszelkie oceny rolnicze szacunkowe, czy też powstałe na drodze statystyki matematycznej (modeli ekonometrycznych) dokonywane aktualnie w Departamencie Rolnictwa, Leśnictwa i Ochrony Środowiska, np. dla podstawowych ziemiopłodów zbóż i rzepaku, odbiegają często w zasadniczy sposób od plonów rzeczywistych. Należy dodać, iż metoda ekonometryczna uwzględnia główne czynniki plonotwórcze, warunki atmosferyczne (temperatura i wilgotność ujęte w syntetyczny wskaźnik Sielaninowa), nawożenie mineralne, procent upraw intensywnych, bonitacja gleb. Dane te nie wyczerpują innych czynników, mających wpływ na plonowanie, jak również ich zmienność, a przede wszystkim wystąpienia suszy, nadmiernego uwilgotnienia, jak również, związanych z tymi zjawiskami, rozwoju chorób i szkodników, utrudnieniami w sprzęcie i stratami w przechowywaniu ziemiopłodów.

Trzeba też wziąć pod uwagę nie zbadane jeszcze w sposób dostateczny możliwości kompensacyjno-regeneracyjne środowiska roślinnego. Możliwości te nie raz, zaledwie w ciągu kilku dni, mimo znamiennych, charakterystycznych symptomów, wskazujących na określony poziom plonowania, zaskoczyły wieloletnich, doświadczonych zarówno naukowców, jak i praktyków rolniczych, zupełnie innym wynikiem plonowania niż spodziewany był kilka dni wcześniej.

Również w ocenie produkcji zwierzęcej występuje duża zmienność produkcji — zwana sezonowością, a wynikająca przede wszystkim z systemu żywienia: letni—zimowy. Ważny wpływ na zmienność w pogłowiu i produkcyjności zwierząt mają relacje ekonomiczno-cenowe, zaopatrzenie w środki produkcji, siłę roboczą itp.

W przedstawionej sytuacji idealnie byłoby, gdyby badać całą populację rolniczą. Z różnych jednak względów, przede wszystkim finansowych oraz organizacyjnych możliwości dokonania badań we wszystkich indywidualnych gospodarstwach rolnych (blisko 3 mln gospodarstw o powierzchni ogólnej powyżej 0,5 ha) nie ma. Praktycznie, ze względu na tematykę badania, dokonywany jest we wszystkich gospodarstwach, raz w roku, powszechny spis rolny, który zapewnia podstawowe informacje o użytkowa-

niu gruntów, powierzchni zasiewów tylko dla głównych ziemiopłodów oraz podstawowe informacje o pogłowiu zwierząt gospodarskich.

Inne badania rolnicze są i mogą być tylko dokonywane metodą reprezentacyjną. Stosując do badania oceny produkcji rolniczej w toku tę metodę skraca się czas badania i opracowywania wyników, jak również zmniejsza jego koszt. Ponadto w sytuacji, gdy przewidywany poziom plonowania roślin, w zależności od przebiegu warunków atmosferycznych, zmienia się z dnia na dzień, czas pełnego opracowania materiału i jego prezentacji użytkownikom jest rzeczą podstawową w statystyce rolniczej.

W tym miejscu należy dodać, że zapotrzebowanie na precyzyjne, aktualne informacje o poziomie produkcji rolniczej jest bardzo duże, jak również znaczenie ich w podejmowaniu decyzji na różnym szczeblu zarządzania (krajowym, wojewódzkim) m.in. określania importu i eksportu płodów rolnych, jak również kształtowania cen itp., co podkreśla jeszcze rolę i znaczenie metody badań reprezentacyjnych w rolnictwie.

RYS HISTORYCZNY BADAŃ REPREZENTACYJNYCH W ROLNICTWIE

Badania reprezentacyjne w rolnictwie mają już dość długą historię. Próby zastosowania metody reprezentacyjnej w badaniu stanu pogłowia zwierząt i drobiu w indywidualnych gospodarstwach chłopskich rozpoczął GUS już w 1956 r.; od 1957 r. do 1962 r. kwartalne, reprezentacyjne spisy pogłowia prowadził w 420 wybranych wsiach. Dobór wsi do próbki był celowy.

W 1959 r. został przeprowadzony, po raz pierwszy, reprezentacyjny spis grudniowy pogłowia zwierząt. Liczebność próby wynosiła 5% zbiorowości generalnej gospodarstw indywidualnych. Do próby wybierano co 20 gospodarstwo z pełnego formularza spisu czerwcowego. Do 1966 r. spisy pogłowia prowadzono metodą wyczerpującą, a w 1967 r. przeprowadzono spis reprezentacyjny, którym objęto 10% zbiorowości generalnej gospodarstw.

W 1970 r. metodę reprezentacyjną zastosowano do określenia plonów głównych ziemiopłodów rolnych. Próba losowa, w każdej objętej badaniem plonów jednostce administracyjnej (gminie lub mieście posiadającym z gminą wspólną radę narodową), była tej samej wielkości i wynosiła 30 gospodarstw indywidualnych. Losowanie przeprowadzono oddzielnie dla każdej gminy metodą warstwową, natomiast jako warstwy przyjmowano następujące grupy obszarowe gospodarstw indywidualnych, uwzględniane w czerwcowych spisach rolniczych:

I — gospodarstwa o powierzchni ogólnej od 0,50 do 1,99 ha,
II — gospodarstwa o powierzchni ogólnej od 2,00 do 4,99 ha,
III — gospodarstwa o powierzchni ogólnej od 5,00 do 6,99 ha,
IV — gospodarstwa o powierzchni ogólnej od 7,00 do 9,99 ha,
V — gospodarstwa o powierzchni ogólnej 10,00 ha i więcej.

Gospodarstwa losowano proporcjonalnie do ogólnej powierzchni gruntów w każdej z warstw, przy pomocy tablic liczb żelaznych, opracowanych przez prof. Hugona Steinhausa.

Z innych ważniejszych badań wykonywanych metodą reprezentacyjną, można wymienić badanie:

— indywidualnych gospodarstw rolnych korzystających z usług kółek rolniczych w 1969 r. Wielkość próby była określona w oparciu o liczbę gospodarstw w województwie: I grupa województw o małej liczebności gospodarstw — 20% ogółu, II grupa o średniej — 10%, III grupa — 5% gospodarstw;
— zużycia nawozów sztucznych i wapna nawozowego pod poszczególne uprawy (1969 r.). Próba wynosiła 5% gospodarstw indywidualnych;
— budownictwa w indywidualnych gospodarstwach rolnych (1969 r.).

W badaniu budownictwa wyróżniono 5 grup województw:

I grupa: województwa, w których liczebność populacji generalnej wynosiła $N=2000$; w grupie tej próba stanowiła 25% populacji; próbę zrealizowano, losując pięć 5% podprób, stosowano losowanie systematyczne;

II grupa: województwa, w których liczebność populacji wynosiła $N=4000$; próba stanowiła 20% populacji; zrealizowano ją przez losowanie czterech 5% podprób;

III grupa: województwa, w których populacja generalna wynosiła około 15000; w tej grupie województw próba stanowiła 10% populacji generalnej i zrealizowano ją, losując pięć 2% podprób;

IV grupa: województwa, w których populacja budynków $N=25000$; próba stanowiła 8% populacji, a zrealizowano ją, przez wylosowanie czterech 2% podprób;

V grupa: województwa, w których populacja budynków wynosiła około 35000; próba stanowiła 5% populacji, a zrealizowano ją, losując pięć 1% podprób.

PODSTAWOWE ZASADY METODYCZNE BADAŃ REPREZENTACYJNYCH W ROLNICTWIE

Podstawowymi badaniami w rolnictwie, przeprowadzonymi metodą reprezentacyjną, są: oceny plonów oraz spisy pogłowia zwierząt gospodarskich; oba badania prowadzone są na tej samej próbie gospodarstw indywidualnych. Do 1983 r. wynosiła ona 5% ogółu gospodarstw o powierzchni ogólnej powyżej 0,5 ha. W 1983 r. ze względów organizacyjno-technicznych, zwłaszcza możliwości wykonawczych, o czym będzie mowa w dalszej części artykułu, próba została ograniczona do 3%.

Operatem losowania są wykazy gospodarstw indywidualnych wykorzystywane w pełnym spisie rolnym. Losowanie próby jest indywidualne, wykonywane w każdej gminie oddzielnie, po uprzednim uporządkowaniu gospodarstw według grup obszarowych. Losuje się próbę, przy zastosowaniu liczb żelaznych z losowanym początkiem.

Poprzez wykorzystanie liczb żelaznych i zjawiska ekwipartycji losowanie upodabnia się do losowania procentowego, warstwowego. Trudno o uniwersalny schemat losowania warstwowego, ponieważ w różnych gminach jest różna liczba gospodarstw w poszczególnych grupach obszarowych, co nie zapewnia procentowego, proporcjonalnego ich udziału w próbie.

Praktycznie technika losowania próby do badań jest następująca: gospodarstwa szereguje się po spisie rolnym w obrębie gminy bądź gminy i miasta o wspólnej radzie narodowej, od najmniejszego do największego pod względem powierzchni ogólnej. Następnie dzieli się populację na warstwy po 100 kolejnych gospodarstw. Z każdej warstwy losowane są 3 gospodarstwa. Taki sposób losowania gospodarstw do badań zapewnia zachowanie rzeczywistej struktury obszarowej gospodarstw w gminie.

Ustalony jest następujący tryb prac związanych z przeprowadzaniem badania: grupowanie, numeracja i losowanie gospodarstw. Gminny organizator rolniczych badań statystycznych sprawdza w terenie zgodność ze stanem faktycznym danych, odnoszących się do poszczególnych gospodarstw objętych wykazami. Wprowadza niezbędne poprawki, a także zobowiązany jest uzyskać zgodę użytkowników gospodarstw na przeprowadzenie w gospodarstwie kilkakrotnych badań rolniczych. W przypadku, gdy użytkownik gospodarstwa odmówi zgody na przeprowadzenie badania i wszystkie próby nakłonienia go do uczestniczenia jego gospodarstwa w badaniu plonów zawiodą, lub wylosowane gospodarstwo faktycznie należy do innej grupy obszarowej niż podano w wykazie, gminny organizator rolniczych badań statystycznych musi zastąpić to gospodarstwo innym. W ten sposób otrzymuje się ostateczne listy gospodarstw przeznaczonych do badania reprezentacyjnego.

W badaniu plonów zaleca się, by liczba gospodarstw objętych badaniem przez jedną osobę nie była większa niż 10.

Uogólnianie danych z badania reprezentacyjnego na pełną zbiorowość, w odniesieniu do oceny plonów, jest dokonywane przy zastosowaniu estymacji prostej. Wyniki reprezentacyjnych spisów pogłowia zwierząt są uogólniane estymatorem złożonym, ilorazowym. Obliczany jest stosunek wyników spisu pełnego do wyników z próby i następnie tym mnożnikiem oblicza się ogólny stan pogłowia. Stan poszczególnych grup wiekowych zwierząt gospodarskich wylicza się natomiast według struktury, z próby reprezentacyjnej.

Założeniem metodycznym jest, aby losowanie nowej próby było przeprowadzane co 5 lat.

Ocena produkcji roślinnej

Reprezentacyjnym badaniem plonów objęte są następujące uprawy: pszenica ozima i jara, żyto, jęczmień ozimy i jary, owies, pszenżyto, mieszanki zbożowe na ziarno, gryka, proso i inne zbożowe (łącznie),

kukurydza na ziarno, kukurydza na zielonkę, strączkowe jadalne na suche ziarno, ziemniaki, buraki cukrowe, rzepak i rzepik, len (słoma nie odziarniona), mak, słonecznik, soja, gorczyca i inne oleiste (łącznie), produkcja z łąk w przeliczeniu na siano w poszczególnych trzech pokosach.

Badanie pełne (3%) dokonywane jest dwa razy w roku, w następujących terminach i obejmuje:
— w sierpniu, powierzchnię uprawy i zbiory przewidywane,
— w listopadzie, powierzchnię uprawy i zbiory uzyskane.

W latach poprzednich zakres sprawozdania z badań reprezentacyjnych (R-72) obejmował także orientacyjny bilans czterech zbóż z mieszankami oraz bilans ziemniaków, który był sprawdzianem rzetelności podawanych przez rolników informacji.

Dane uzyskiwane są w formie wywiadów z użytkownikiem gospodarstwa rolnego. Użytkownik informuje o poziomie spodziewanych i uzyskanych zbiorów poszczególnych ziemiopłodów oraz powierzchni ich zasiewów, według czerwcowego spisu rolnego.

Organizator gminny zestawia wyniki na formularzu zbiorczym pomocniczym dla gminy. Formularz ten w boczku zawiera miejsce na imienny wykaz gospodarstw (nazwisko i imię użytkownika oraz miejsce zamieszkania), a w części właściwej rubryki na zamieszczenie danych o powierzchni zasiewów i zbiorach poszczególnych ziemiopłodów. Dane o powierzchni i zbiorach dla gminy są sumowane, a przeciętny plon obliczany jest jako iloraz sumy zbiorów i sumy powierzchni.

Układ formularzy jest opracowany w ten sposób, że zapewnia wahadłowe ich wykorzystywanie, w dwóch szacunkach plonów. Dane o przeciętnych plonach na szczeblu gminy przenoszone są na formularz podstawowy (R-62 — szacunek plonów w gminie, mieście) i wykorzystywane są, jako ważne źródło o poziomie plonowania, na podstawie którego rzeczoznawca gminny określa plon ostateczny, poszczególnych ziemiopłodów.

W niektórych WUS, np. w województwie sieradzkim czy lubelskim dane z badań reprezentacyjnych plonów obliczane są techniką komputerową, co w poważnym stopniu odciąża gminną służbę rolną od sporządzania dość pracochłonnych zestawień, jak również eliminuje błędy rachunkowe, popełnione przy opracowywaniu wyników badań metodą ręczną.

Problemem jest przekazywanie opracowanych systemem epd materiałów do gmin, gdyż dane z badań są istotnym źródłem informacji do oceny plonów w poszczególnych szacunkach w gminie, dla rzeczoznawcy gminnego PIPR.

Dla sprawdzenia wyników badań reprezentacyjnych przeprowadzono analizę statystyczno-matematyczną danych o plonach w województwie kieleckim. Województwo to wybrano do analizy ze względu na: największą w kraju liczbę indywidualnych gospodarstw rolnych wylosowanych do

badań, a także dużą liczbę gospodarstw — reprezentantów w poszczególnych gminach, znaczne terytorialne zróżnicowanie warunków fizjograficzno-glebowych oraz poziom kultury rolnej.

Z analizy materiałów wynika, iż w skali województwa zarówno względny błąd standardowy plonu przewidywanego, jak i względny błąd standardowy plonu uzyskanego dla większości badanych upraw jest nieistotny statystycznie, z wyjątkiem nie odziarnionej słomy lnu. Wysokie są współczynniki korelacji między danymi o plonie przewidywanym a uzyskanym — od 0,972 dla ziemniaków do 1,000 rzepaku, lnu oraz łąk (I, II i III pokos). Błędy przeciętne dla województwa, wyliczone z gmin zarówno dla plonu przewidywanego, jak i uzyskanego, są wyższe od wojewódzkich, ale także rzadko przekraczają kilka procent.

We wszystkich gminach (72) współczynniki korelacji trafności oceny plonów były wysokie, nie niższe niż 0,900. Natomiast wyższe były względne błędy standardowe plonu przewidywanego i uzyskanego i sięgały powyżej 50%.

Wyniki analizy precyzji badań reprezentacyjnych plonów w poszczególnych gminach były różne. Np. w gminie Masłów względne błędy standardowe plonu przewidywanego i uzyskanego są nieistotne statystycznie. Współczynnik korelacji wahał się od 1,000 (mieszanki zbożowe, ziemniaki i łąki — I, II, III pokos) do 0,980 (jęczmień jary).

Metoda reprezentacyjna znajduje również zastosowanie w pierwszym szacunku plonów. Badanie dokonywane jest na części 3% próby, tj. wybieranych jest 10 gospodarstw do badań, licząc co dziesiąte z listy porządkowej gospodarstw wylosowanych do badania reprezentacyjnego plonów w drugim i trzecim szacunku. Na tej próbie przeprowadzane są pomiary biometryczne głównych elementów plonu podstawowych zbóż: żyta i pszenicy ozimej. Określane są: zagęszczenie kłosów na 1 m^2, przeciętna liczba ziarn w kłosie i masa 1000 ziaren oraz straty, występujące w okresie wegetacji związane z rozwojem chorób i szkodników, terminem i sposobem sprzętu. Wyniki pomiarów biometrycznych są zestawiane na formularzach R-71, w układzie poszczególnych gospodarstw.

Średni plon żyta lub pszenicy ozimej obliczany jest według wzoru:

$$P_{netto} = \frac{K \times Z \times C}{10000} - S$$

gdzie: P — plon ziarna w dt/ha,
K — liczba kłosów na 1 m^2,
Z — liczba ziaren w kłosie,
C — masa 1000 ziaren w g,
S — straty w dt.

Wyliczony w ten sposób plon przenoszony jest na formularz podstawowy R-62, gdzie stanowi ważny punkt odniesienia, do określenia przez rzeczoznawcę gminnego ostatecznego plonu (żyta, pszenicy ozimej) w skali gminy.

W statystyce produkcji roślinnej do pomiarów biometrycznych, jako metody oceny plonów, przywiązuje się bardzo dużą wagę, jako do metody mierzalnej, opartej na konkretnych pomiarach elementów plonu na plantacji. Opracowane są, przez były Inspektorat Główny PIPR, metody oceny podstawowych ziemiopłodów rolnych oraz ogrodniczych, łąk, a w przygotowaniu jest metodyka oceny plonów upraw pastewnych (Zeszyty Metodyczne GUS nr 50 i 63).

Ocena produkcji zwierzęcej

Metoda badań reprezentacyjnych stosowana jest także do oceny zmian w pogłowiu i produkcyjności zwierząt gospodarskich. Zakres kwartalnego, reprezentacyjnego spisu pogłowia zwierząt gospodarskich (R-19) obejmuje 6 podstawowych działów, dotyczących pogłowia zwierząt i produkcyjności zwierząt: owce, konie, bydło, mleczność krów i zagospodarowanie mleka, trzoda chlewna oraz drób.

W bardziej szczegółowym ujęciu poszczególne działy zawierają także inne pozycje, niezbędne dla określenia bilansu obrotu stada, takie jak: stan pogłowia w gospodarstwie na początku kwartału: strona przychodowa — ogółem, w tym z urodzenia oraz z zakupu i innych źródeł; strona rozchodowa — padnięcia, ubój gospodarczy wraz z darowizną i wymianą sąsiedzką, sprzedaż do punktu skupu, sprzedaż pozostała i rozchody inne; stan pogłowia w końcu kwartału. Dodatkowo, poszczególne gatunki zwierząt są przedstawione w podstawowych grupach wiekowych.

Kwartalne, reprezentacyjne spisy pgłowia przeprowadzane są trzy razy w roku, na przełomie miesięcy: grudzień/styczeń, marzec/kwiecień, wrzesień/październik i obejmują próbę wynoszącą 3% ogółu indywidualnych gospodarstw rolnych o powierzchni ogólnej gospodarstwa powyżej 0,50 ha. Dane o stanie pogłowia zwierząt gospodarskich za II kwartał zapewnia pełny spis rolniczy dokonywany na przełomie czerwca i lipca.

Kwartalny, reprezentacyjny spis zwierząt gospodarskich pozwala na śledzenie postępu w rozwoju produkcji zwierzęcej, analizę elementów warunkujących intensyfikację chowu, prognozowanie rozwoju pogłowia, jak również produkcji żywca, opracowanie bilansu obrotu stada, mleka itp. Dane ze spisów kwartalnych są wykorzystywane zarówno na szczeblu wojewódzkim, jak i na szczeblu krajowym. Liczebność próby nie pozwala na pełne uogólnienie na szczeblu gminy; stosowany jest estymator ilorazowy. Jednakże i na tym szczeblu może być wykorzystywany do określenia pewnych tendencji rozwojowych, opracowywania trendów.

Metoda reprezentacyjna znajduje także zastosowanie w badaniach kontrolnych, np. do sprawdzania rzetelności wyników pełnego spisu rolniczego. Wykonywane są następujące badania kontrolne:

1) spis kontrolny zwierząt w 100 obwodach w kraju, wylosowanych w sposób losowy,
2) spis interwencyjny, mający na celu kontrolę rachmistrzów, wykonywany jest w trakcie trwania spisu podstawowego i obejmuje 20% gmin,
3) kontrolne pomiary powierzchni, obejmujące 10 gospodarstw w każdym województwie (pomiary przeprowadzają służby geodezyjne przy współpracy pracowników merytorycznych WUS).

ZALETY I WADY BADAŃ REPREZENTACYJNYCH W ROLNICTWIE

Podstawowymi zaletami metody reprezentacyjnej są: mniejszy koszt badań, mniejsza pracochłonność, szybkie opracowywanie materiałów i operatywna prezentacja wyników, zapewnienie wielokrotnych (bieżących) danych o rozwoju produkcji rolniczej, co jest niezwykle ważne w statystyce i informacji rolniczej, ze względu na dużą zmienność tej produkcji, trudnej do przewidzenia w momencie prognozowania.

Ponadto, badania reprezentacyjne pozwalają na wprowadzenie konkretnych, niemożliwych do przeprowadzenia w badaniu pełnym, metod mierzalnych, np. wymienionych wcześniej pomiarów biometrycznych, które są badaniami najbardziej pewnymi i obiektywnymi.

Wadą badań reprezentacyjnych w rolnictwie jest to, że wśród kryteriów losowania próby nie uwzględnia się wielu uwarunkowań, bardzo ważnych z punktu widzenia charakterystyki przyrodniczo-rolniczej, ze względu na trudności w ich wyeksponowaniu.

W stosowanej 3% próbie uwzględnia się strukturę obszarową gospodarstw, a nie bierze się pod uwagę warunków przyrodniczo-glebowych poszczególnych gospodarstw oraz ich poziomu kultury rolnej, decydujących o poziomie plonowania, jak również intensywności produkcji zwierzęcej. Dla rzadziej uprawianych roślin rolniczych występują niejednokrotnie przypadki pominięcia ich w próbie reprezentacyjnej.

W reprezentacyjnych badaniach rolniczych nie występuje, tak jak w badaniach społecznych, zjawisko braku odpowiedzi, natomiast znaczną uciążliwość stanowi dość intensywny proces wypadania indywidualnych gospodarstw rolnych z próby, wskutek np. zmiany użytkownika, łączenia gospodarstw itp., co w dużej mierze obniża reprezentatywność próby oraz zniekształca strukturę badanej zbiorowości.

I wreszcie, najpoważniejszy problem stanowi sposób przeprowadzenia badań reprezentacyjnych. Specyfika statystyki rolniczej, zwłaszcza ocena plonów wymaga od przeprowadzających badanie dużej wiedzy zarówno teoretycznej, jak i praktycznej z zakresu zwłaszcza biologii roślin, zasad

agrotechniki i odmianoznawstwa oraz dużej solidności i rzetelności. Dotyczy to zwłaszcza przeprowadzania pomiarów biometrycznych. GUS nie posiada swoich służb, a opiera się w badaniach na osobach postronnych, zwłaszcza w doborze rachmistrzów, przeprowadzających kwartalne spisy zwierząt gospodarskich.

W badaniu reprezentacyjnym plonów wykonawcą jest gminna służba rolna, której przygotowanie merytoryczne do przeprowadzenia ocen jest większe. Sytuacja się pogorszyła od momentu, gdy nastąpił podział gminnej służby rolnej na administracyjną, zatrudnioną w urzędach gmin i doradczą, zatrudnioną w wojewódzkich ośrodkach postępu w rolnictwie.

Trzeba podkreślić, iż mimo tych mankamentów zastosowanie metody reprezentacyjnej w badaniach rolniczych jest przydatne, z uwagi na znaczną precyzję wyników i wysoki wskaźnik korelacji, z ocenami i wynikami osiągniętymi w badaniach pełnych, jak i w sprawozdawczości gospodarki uspołecznionej.

LITERATURA

[1] *Badania statystyczne metodą reprezentacyjną w krajach socjalistycznych*, Biblioteka Wiadomości Statystycznych, Warszawa 1971 r.

[2] E. Cypelt: *Organizacja i metody badań reprezentacyjnych Państwowej Inspekcji Produkcji Rolniczej*, Wiadomości Statystyczne nr 1, Warszawa 1988 r.

[3] J. Kordos: *Możliwości szerszego stosowania metody reprezentacyjnej w badaniach statystycznych krajów – członków RWPG*, Przegląd Statystyczny, R XVIII, zeszyt 1, Warszawa 1970 r.

[4] J. Kordos: *Zastosowanie metody reprezentacyjnej w statystyce rolniczej*, Przegląd Statystyczny, R XV, zeszyt 3, Warszawa 1968 r.

[5] *Metodologia badań reprezentacyjnych w GUS*, Biblioteka Wiadomości Statystycznych, Warszawa 1979 r.

[6] F. Ruszkowski: *Informacje korespondentów GUS — barometrem sytuacji w rolnictwie*, Plon nr 44/45, Warszawa 1984 r.

[7] F. Ruszkowski: *Podstawowe metody stosowane w ocenie produkcji roślinnej w toku* — cz. I., Służba Rolna nr 9, Poznań 1983 r.

[8] F. Ruszkowski: *Podstawowe metody stosowane w ocenie produkcji roślinnej* — cz. II — Pomiary biometryczne, Służba Rolna nr 10/11, Poznań 1983 r.

[9] F. Ruszkowski: *Pomiary biometryczne podstawowymi metodami oceny produkcji roślinnej w toku*, Wiadomości Statystyczne nr 5, Warszawa 1986 r.

[10] F. Ruszkowski: *10 lat funkcjonowania Państwowej Inspekcji Produkcji Rolniczej*, Wiadomości Statystyczne nr 11, Warszawa 1986 r.

[11] F. Ruszkowski: *Statystyka produkcji roślinnej* (skrypt), PTE, Łódź 1984 r.

[12] F. Wileński: *Spisy rolne jako źródło informacji o rozwoju rolnictwa*, Wiadomości Statystyczne nr 7, Warszawa 1988 r.

[13] Zeszyt Metodyczny — *Definicje z zakresu oceny produkcji roślinnej w toku*, GUS, Warszawa 1982 r.

[14] Zeszyt Metodyczny — *Zasady dokonywania oceny produkcji roślinnej w toku w badaniach statystycznych*. Część I. Rolnictwo, GUS, Warszawa 1984 r.

[15] Zeszyt Metodyczny — *Zasady dokonywania oceny produkcji roślinnej w toku w badaniach statystycznych*. Część II. Ogrodnictwo, GUS, Warszawa 1986 r.

SŁOWO KOŃCOWE

Zarówno prezentowane referaty, jak również ożywiona dyskusja dowiodły, że tematyka konferencji jest niezwykle istotna dla dalszego rozwoju badań statystycznych. Dotyczy to nie tylko Głównego Urzędu Statystycznego, ale także uczelni wyższych i instytutów naukowo-badawczych, które coraz częściej korzystają z tej metody badawczej. Chodzi tu zarówno o prawidłowe zastosowanie metody reprezentacyjnej w praktyce, jak i o właściwe nauczanie tego przedmiotu na uczelniach wyższych. Trzeba się zgodzić z tymi głosami w dyskusji, a szczególnie chodzi mi o obszerną wypowiedź doc. dra Stanisława Wierzchosławskiego, którą można by uważać za pewnego rodzaju podsumowanie przebiegu konferencji, wyrażającą niepokój z powodu nieuwzględniania metody reprezentacyjnej w programie nauczania na niektórych kierunkach studiów. Metoda ta zyskała uznanie nie tylko w badaniach statystycznych, ale także w wielu badaniach naukowych, której zastosowanie będzie z pewnością dalej wzrastało. Niezbędne są prace badawcze nad efektywnym jej zastosowaniem w praktyce. Jak wspomniałem w „Słowie wstępnym", wniosek taki zgłoszono już na międzynarodowym seminarium, poświęconym metodzie reprezentacyjnej.

Z dyskusji, prowadzonej na tej konferencji, wynika również, że metodologiczne problemy badań reprezentacyjnych nie są szerzej znane ogółowi statystyków. Znane są zwykle postępowania związane z losowaniem próbki (operaty losowania, schematy losowania, metody estymacji i metody oceny precyzji), ale zaniedbana jest problematyka błędów nielosowych i zagadnienia metodologiczne, dotyczące wstępnej fazy przygotowania badania. Na ogół nie traktuje się łącznie błędów losowych i nielosowych w publikacjach wyników z badań reprezentacyjnych, a często ogranicza się jedynie do omówienia rozmiarów błędów losowych, gdy wspomina się w publikacji o błędach badania. Jak słusznie podkreśla doc. dr St. Wierzchosławski, nie poświęca się dostatecznej uwagi w zastosowaniu metody reprezentacyjnej do badania zjawisk zmieniających się w czasie (np. badania longitudinalne). Na tej konferencji nie omawialiśmy tych zagadnień, a przecież często mamy do czynienia z badaniem zbiorowości w pewnych odcinkach czasowych (badania powtarzalne) lub z obserwacjami ciągłymi (np. badania budżetów gospodarstw domowych). Dla badań tego rodzaju niezwykle ważna jest budowa „próby-matki". Dobrze się stało, że tym

problemom poświęcono specjalny referat oraz poruszono te zagadnienia w dyskusji. Pozostało jeszcze wiele znaków zapytania i dlatego niezbędne są dalsze prace w tym kierunku. Dotyczy to w szczególności budowanego Zintegrowanego Systemu Badań Gospodarstw Domowych, który wdrażany jest w Polsce od 1982 r.

Poważnym ograniczeniem w szerszym zastosowaniu metody reprezentacyjnej w praktyce, jest tzw. problem małych obszarów, dla których wyniki badań reprezentacyjnych wykazują zwykle niewielką precyzję. W ostatnim okresie rozwijane są specjalne techniki badawcze, które mogą być pomocne w tego rodzaju badaniach. W Polsce dotychczas nie zajmowaliśmy się tymi problemami. Ze względu na wagę tych zagadnień należałoby w przyszłości poświęcić im specjalną uwagę, a może także zorganizować ogólnopolską konferencję naukową.

W czasie dyskusji dużo miejsca poświęcono problemowi braku odpowiedzi (non-response). Zagadnień tych dotyczył referat R. Platka oraz A. Kubiczka. Problemy te są niezwykle ważne w każdym badaniu reprezentacyjnym, a szczególnie w badaniach społecznych, w których występuje wiele czynników nie sprzyjających podejmowaniu badań. Powstaje więc problem minimalizacji błędów tego rodzaju oraz zastosowania odpowiednich technik, które w tych badaniach okazać się mogą bardziej efektywne.

Dyskutowano tu również celowość powołania przy Polskim Towarzystwie Statystycznym sekcji badań reprezentacyjnych. Wydaje się, że inicjatywę tę należy poprzeć i zgłosić na następnej sesji plenarnej Rady Głównej PTS, aby załatwić ją od strony formalnej. Może potraktować ten problem nieco szerzej i powołać sekcję metodologii i organizacji badań statystycznych, która obejmie także zagadnienia metody reprezentacyjnej, tj. całość problematyki badań statystycznych. Towarzystwo może wspierać działania takiej sekcji, gdyż jest to w pełni zgodne z jego statutem. Umożliwiłaby ona integrację teoretyków i praktyków, zajmujących się badaniami reprezentacyjnymi oraz przyczyniłaby się do bardziej efektywnego zastosowania tej metody w praktyce.

W zakończeniu chciałbym podziękować tym wszystkim, którzy przyczynili się do zorganizowania obecnej konferencji naukowej, a w szczególności prof. dr. hab. Mirosławowi Krzysztofiakowi oraz pracownikom naukowym Uniwersytetu Gdańskiego.

Jan Kordos

Dotychczas w serii
BIBLIOTEKA WIADOMOŚCI STATYSTYCZNYCH
ukazały się następujące pozycje:

1967 r.

Tom 1. Wybrane problemy statystyki w Polsce
Tom 2. Zagadnienia statystyki rolniczej
Tom 3. Statystyka regionalna. Aktualny stan i problemy rozwoju
Tom 4. Problemy demograficzne Polski Ludowej

1968 r.

Tom 5. Bilans gospodarki narodowej

1969 r.

Tom 6. Problemy demograficzne Ziem Zachodnich i Północnych PRL
Tom 7. Zastosowanie metod matematycznych w statystyce
Tom 8. Problemy demograficzne Kielecczyzny
Tom 9. Mierniki rozwoju regionów

1970 r.

Tom 10. Założenia programowe i organizacyjne Narodowego Spisu Powszechnego w 1970 r.
Tom 11. Wybrane problemy prognoz statystycznych
Tom 12. Aktualne problemy statystyki regionalnej w krajach europejskich
Tom 13. Teoretyczne i metodologiczne problemy statystyki społecznej

1971 r.

Tom 14. Badania statystyczne metodą reprezentacyjną w krajach socjalistycznych
Tom 15. Wybrane problemy metodologiczne badań reprezentacyjnych
Tom 16. Rola i zadania statystyki państwowej w planowaniu i zarządzaniu gospodarką narodową w krajach RWPG

1972 r.

Tom 17. Aktualne problemy statystyki państwowej
Tom 18. Eksperymentalne badania budżetów rodzinnych metodą rotacyjną

1973 r.

Tom 19. Problemy demograficzne województwa zielonogórskiego
Tom 20. Podstawowe problemy rozwoju statystyki regionalnej
Tom 21. Stan i perspektywy rozwoju statystyki w Polsce
Tom 22. Rozwój regionalny Polski w świetle badania dochodu narodowego

1974 r.

Tom 23. Problemy mierników poziomu życia ludności
Tom 24. Aktualne problemy demograficzne kraju

1976 r.

Tom 25. Statystyka i ekonometria w Polsce Ludowej
Tom 26. Zagadnienia metodologiczne statystyki społeczno-demograficznej

1978 r.

Tom 27. Rola młodzieży w życiu społeczno-gospodarczym kraju
Tom 28. Narodowy Spis Powszechny 7 grudnia 1978 r.
Metodologia i organizacja
Tom 29. Metodologia badań reprezentacyjnych w GUS
Prace Komisji Matematycznej

1979 r.

Tom 30. Narodowy Spis Powszechny 1978 jako źródło informacji o migracjach
Tom 31. Statystyka i ekonometria w Polsce Ludowej (wyd. drugie — zmienione)

1981 r.

Tom 32. Tematyka i organizacja spisów powszechnych w Polsce

1983 r.

Tom 33. Problemy demograficzne województwa lubelskiego

1987 r.

Tom 34. Problemy integracji statystycznych badań gospodarstw domowych
Tom 35. Dokładność danych w badaniach społecznych (autor Jan Kordos)

1989 r.

Tom 36. Problemy badań statystycznych metodą reprezentacyjną

Cena zł 600,–