

Rozdział V

DOBÓR PRÓB LUDNOŚCI

W rozdziale tym przedstawiamy zasadnicze informacje o doborze prób ludności do naszych badań. Zaczynamy od opisu operatu losowania, jakim posłużyliśmy się w badaniach "Struktura społeczna I" i "Struktura społeczna II", a następnie przedstawiamy konkretne zasady doboru prób w każdym z tych badań. Kolejną pozycją jest badanie "Struktura społeczna - Łódź, 1988", w którym dobór próby był pomyślany jako możliwie zbliżony do tego, jaki stosowano w badaniach wcześniej prowadzonych w tym samym mieście (Wesołowski, 1970; Słomczyński i Wesołowski, 1973; Janicka, 1987). Rozdział ten kończą uwagi o doborze próby do badania kontrolnego wobec badania "Struktura społeczna II". Ze względu na szczególny charakter badania kontrolnego, próba tego badania jest podpróbą zrealizowanej części próby badania zasadniczego.

CHARAKTERYSTYKA OPERATU LOSOWANIA

W obu badaniach - "Struktura społeczna I" i "Struktura społeczna II" - dobierano próbę ludności korzystając z operatu losowania, jakim dysponuje Centrum Badania Opinii Społecznej. Z tego też względu jakość wylosowanych prób zależy od poprawności konstrukcji operatu losowania tej instytucji. Operat ów - nazwany ANLUD - utworzony został przez Centrum Badania Opinii Społecznej w 1986 roku. Celem jego konstrukcji było stworzenie możliwości probabilistycznego doboru prób do części badań realizowanych przez Centrum.

Konstrukcję operatu losowania oparto na materiałach Głównego Urzędu Statystycznego. Wykorzystano przy tym podział terytorialny kraju stosowany podczas prowadzenia spisów powszechnych. Elementarnymi jednostkami tego podziału są *obwody spisowe*. Pojedynczy obwód obejmuje w miarę jednorodną i spójną część podziału administracyjnego i odpowiada wyodrębniającemu się w sposób naturalny skupisku ludności. Przykładowo, obwodem spisowym może być mała wieś, część dużej wsi, kilka domów przy jednej ulicy w małym mieście, sektor bloków mieszkalnych, pojedynczy duży budynek. Zwykle obwody spisowe obejmują od kilkunastu do kilkuset osób.

ANLUD jest losową próbą osób w wieku powyżej 12 lat. Losowania osób dokonano dwustopniowo. Pierwsza faza objęła *warstwowanie* ludności Polski w wieku powyżej 12 lat. W każdym województwie wyróżniono cztery warstwy w zależności od wielkości miejscowości zamieszkania. Posłużono się do tego celu podziałem na wsie oraz miasta małe (do 20 tysięcy mieszkańców), średnie (od 20 do 100 tysięcy mieszkańców) i

duże (ponad 100 tysięcy mieszkańców). W sumie utworzono 196 warstw, a następnie określono liczbę ludności należącej do każdej warstwy i ustalono liczbę obwodów spisowych wchodzących do operatu losowania. Konkretnie obwody w wymaganej liczbie wylosowano korzystając z zestawień obwodów dla poszczególnych województw. W sumie, w drodze losowania dobrano 815 obwodów spisowych.

Druga faza losowania polegała na doborze - w ramach każdego obwodu spisowego - określonej liczby *gospodarstw domowych*, których członkowie w wieku powyżej 12 lat mieliby wejść w skład przygotowywanego operatu. Losowanie przeprowadzono przy ustaleniu jednolitej dla wszystkich obwodów frakcji osób, określonej w taki sposób, aby operat losowania zawierał 130 tysięcy osób. Procedura doboru gospodarstw domowych zakończyła drugą fazę losowania jednostek obserwacji do operatu doboru prób.

Ze względu na dostępną dokumentację prowadzoną przez Główny Urząd Statystyczny, zastosowanie bezpośredniego losowania jednostek obserwacji przy konstrukcji opisywanego operatu było praktycznie niemożliwe. W zamian więc zastosowano *losowanie dwustopniowe*, gdzie jednostkami doboru pierwszego stopnia były *obwody spisowe*, zaś jednostkami drugiego stopnia - *gospodarstwa domowe*. W schemacie tego rodzaju, podstawowym źródłem obciążenia losowości jest pierwszy etap: dobór obwodów spisowych. Aby maksymalnie ograniczyć powstanie obciążeń losowości, zastosowano procedurę warstwowania i starano się dobrać możliwie dużą liczbę obwodów spisowych.

Trzeba jednak zwrócić uwagę, że decyzja dotycząca liczby losowanych obwodów spisowych ma charakter arbitralny i zależy od środków przeznaczonych na konstrukcję omawianego operatu. W świetle kryteriów statys-

tycznych, w omawianym przypadku obciążenie losowości jest malejącą funkcją liczby wziętych pod uwagę (wylosowanych) obwodów spisowych. Jednakże, w oparciu o aparat analityczny statystyki matematycznej nie jest praktycznie możliwe ustalenie dostatecznej, czy pożądanej liczby tych obwodów. Zwróćmy uwagę, że ANLUD był pomyślany jako operat doboru prób o różnej strukturze w zakresie cech społeczno-demograficznych.

Liczbę 815 obwodów spisowych można uznać za na tyle dużą, że w większości typowych sytuacji badawczych - w których przedmiotem badań jest próba przekrojowa - błąd statystyczny wprowadzony "warstwowością" losowania może być, w kontekście źródeł innych błędów, całkowicie pominięty. W takim przekonaniu umacnia zróżnicowanie tych obwodów ze względu na ich wielkość i kryteria podziałów terytorialnych.

Wylosowanie gospodarstw domowych w wybranych obwodach spisowych stanowiło ważny krok w opracowywaniu operatu ANLUD. Jednakże zasadnicza część prac polegała na przeprowadzeniu przez Centrum Badań Opinii Społecznej *weryfikacyjnego badania*, którego przedmiotem byli wszyscy członkowie wylosowanych gospodarstw domowych spełniający kryterium wieku, a więc mający powyżej 12 lat. Badanie to - obejmujące około 130 tysięcy osób - zostało zrealizowane techniką wywiadu kwestionariuszowego. Podstawowym celem tego badania było ustalenie poprawnego brzmienia nazwiska i imienia każdej osoby, jej aktualnego adresu oraz podstawowych cech demograficznych: płci i wieku. Dodatkowo pytano także o inne charakterystyki społeczne, które potencjalnie mogły służyć jako kryteria doboru prób - jak osiągnięte wykształcenie i wykonywany zawód. W przypadku obu tych charakterystyk posłużono się narzędziem pomiaru opartym na zestawie pytań uwzględniających istniejącą wiedzę metodologiczną (np. Cichomski, 1974; Pohoski i Słomczyński,

1978; Sawiński, 1986) i testowanych we wcześniejszych badaniach prowadzonych przez Centrum. Pytano także o poziom dochodów na osobę w gospodarstwie domowym, warunki mieszkaniowe, posiadanie wybranych dóbr materialnych, przynależność do organizacji społeczno-politycznych oraz o stosunek do wiary i praktyk religijnych. We wszystkich tych przypadkach także starano się zastosować sprawdzone narzędzia pomiaru (por. np. Daniłowicz i Sztabiński, 1977).

Omawiane badanie weryfikacyjne zostało zrealizowane w terminie od listopada 1985 do lutego 1986 roku. Informacje zawarte w kwestionariuszach zostały następnie zakodowane i przeniesione na nośnik maszyny cyfrowej. Poprawność kodowania była skrupulatnie sprawdzona, a na założonym zbiorze danych przeprowadzono kontrolę logiczną zgodności uzyskanych informacji. Ostatecznie, zbiór danych ANLUD, obejmuje informację o 129040 osobach.

W porównaniu z innymi dostępnymi operatami doboru prób, które potencjalnie mogły mieć zastosowanie w naszych badaniach, ANLUD posiadał szereg zalet:

- po pierwsze, ANLUD został skonstruowany w sposób gwarantujący poprawność pobieranych prób i w tym zakresie był testowany;
- po drugie, dane o jednostkach losowania - w szczególności nazwiska i adresy - były w okresie realizacji naszych badań stosunkowo aktualne;
- po trzecie, operat ten był do chwili naszych badań stosunkowo mało "wyeksploatowany", gdyż dysponent wykorzystywał go jedynie w nielicznych badaniach;
- po czwarte, dzięki odpowiedniemu oprogramowaniu maszyny cyfrowej, próby mogły być sporządzone w krótkim czasie i przy niskich kosztach.

Należy podkreślić, że w Centrum Badania Opinii Społecznej dokonano sprawdzenia poprawności omawianego operatu losowania poprzez porównanie rozkładów podstawowych cech społeczno-demograficznych zbioru ANLUD z analogicznymi rozkładami pochodzącymi z mikrospisu, zrealizowanego przez Główny Urząd Statystyczny w 1984 roku. Testowanie istotności różnic w rozkładach przeprowadzono zarówno dla wszystkich osób znajdujących się w zbiorze ANLUD, jak i dla grup wieku. Nie stwierdzono takich odstępstw, które podważałyby reprezentatywność zawartości operatu doboru prób dla ludności Polski w wieku powyżej 12 lat.

PRÓBA LUDNOŚCI W BADANIU "STRUKTURA SPOŁECZNA I"

Ponieważ w znacznej części problematyka badania "Struktura Społeczna I" koncentruje się wokół sytuacji pracy, stąd też przyjęto, aby badaną zbiorowość ograniczyć do *ludności pracującej*. Dobierając próbę postużyliśmy się definicją ludności pracującej stosowaną przez Główny Urząd Statystyczny, a wykorzystaną również przez Centrum Badania Opinii Społecznej przy konstrukcji operatu ANLUD. Według tej definicji, do *wyróżnionej kategorii czynnych zawodowo*

"zalicza się wszystkie osoby pracujące w gospodarce uspołecznionej i nie uspołecznionej, bez względu na podstawę stosunku pracy bądź innego stosunku prawnego, na podstawie którego świadczona jest praca. Do pracujących zalicza się: (1) w gospodarce uspołecznionej (a) zatrudnionych na podstawie stosunku pracy, (b) osoby wykonujące pracę nakładczą, (c) agentów (...) oraz osoby zatrudnione przez agentów, (d) członków rolniczych spółdzielni produkcyjnych (...), (e) pracujących na działkach przyzagrodowych (...) rolniczych spółdzielni produkcyjnych i (...) państwowych gospodarstw rolnych, (f) osoby zatrudnione przez komitety rodzicielskie (...); (2) w gospodarce nie uspołecznionej (a) w rolnictwie [rolników indywidual-

nych i osoby im pomagajęcej, (b) poza rolnictwem: właścicieli i współwłaścicieli zakładów rzemieślniczych, handlowych i usługowych (łącznie z pomagającymi członkami ich rodzin), rybaków morskich, a także osoby zatrudnione przez właścicieli, osoby zatrudnione w przedsiębiorstwach zagranicznych (...) oraz duchownych wszystkich wyznań" (Zasady, 1986: 14).

Uściślenie pojęcia ludności pracującej – tak, jak jest ono stosowane w naszych badaniach – podano w instrukcji dla ankietów, gdzie wprowadzono również kryterium liczby godzin pracy w miesiącu (por. rozdział III).

Dodatkowo, zakres badanej zbiorowości ograniczono przez kryterium wieku, zastosowane odmiennie dla mężczyzn i kobiet. Do próby losowano mężczyzn urodzonych w latach 1922-1969, czyli mających w 1987 roku od 18 do 65 lat. W przypadku kobiet przedział lat urodzenia wynosił 1927-1969, co dla 1987 roku odpowiada przedziałowi wieku od 18 do 60 lat. Niejednakowe przedziały wieku dla obu płci są związane z różną granicą wieku emerytalnego mężczyzn i kobiet.

Liczebność próby założonej do realizacji ustalono na 2000 osób, na co składały się dwie 1000-osobowe próby dla obu wersji kwestionariusza. Z operatu ANLUD wylosowano jednak dwukrotnie więcej nazwisk i adresów. Z tego zbioru do próby podstawowej zaliczono co drugą osobę, a resztę stanowiła rezerwa, z której korzystano w przypadku trudności dotarcia do właściwego respondenta.

W sumie przeprowadzono wywiady kwestionariuszowe z 1894 osobami. Analiza jakości zrealizowanej próby będzie przedmiotem osobnego studium, w którym dokonamy porównań cech społeczno-demograficznych respondentów w naszym badaniu z informacjami z operatu losowania ANLUD. Ponadto, część porównań będzie polegać na zestawieniu naszych danych z danymi z innych badań, w których zbierano takie same informacje o respondentach. Do tych innych badań należy – w szczególności – badanie

"Co o tym myślisz?" przeprowadzone przez Centrum Badania Opinii Społecznej w 1987 roku. W tym badaniu powtórzono – właśnie w celach porównawczych – baterię pytań o poglądy społeczno-polityczne z naszego badania "Struktura społeczna I".

PRÓBA LUDNOŚCI W BADANIU "STRUKTURA SPOŁECZNA II"

Ponieważ problematyka badania "Struktura społeczna II" była zróżnicowana, w sensie populacji badanie to objęło *dorostą ludność* Polski, ograniczoną jedynie przez kryterium wieku. Dla kobiet i mężczyzn przyjęto jednolicie przedział lat urodzenia 1922-1966, co w 1987 roku odpowiadało zakresowi wieku od 21 do 65 lat. Tak określoną populację można traktować jako zasadniczo równoważną zbiorowości dorosłej ludności Polski, przy pominięciu osób najmłodszych (18-20 lat) i osób najstarszych (powyżej 65 roku życia). W populacji osób dorosłych te dwie pominięte kategorie stanowiły w połowie lat 1980-tych odpowiednio: 6% i 14%.

Względy, dla których nie uwzględniliśmy w omawianym badaniu kategorii osób w wieku 18-20 lat są natury teoretycznej, metodologicznej i praktycznej. Do względów teoretycznych zaliczamy fakt, że lokacja jednostek w strukturze społecznej należących do tej kategorii podlega działaniom fluktuacyjnym związanym z adaptacją do ról społecznych "świata dorosłych" i mogłaby zaciemniać główne rysy obrazu struktury społecznej. Względ metodologiczny polega na tym, że w tej kategorii proporcja jednostek niedostępnych obserwacji – z powodu zmian adresu związanego z odbywaną nauką i służbą wojskową – nie gwarantuje, aby była ona reprezentatywna. Wreszcie, kategoria ta jest intensywnie badana w studiach nad młodzieżą, gdzie problematyka jej usytuowania w strukturze

społecznej również występuje; ze względów praktycznych nie jest zasadne powtarzanie wysiłku innych badaczy.

Nieco inne względy kierowały nami przy pominięciu osób najstarszych. Osoby te, z racji przekroczonego 65 roku życia, na ogół należą do kategorii emerytów i rencistów. Uważamy, iż jest to kategoria ważna w badaniach nad strukturą społeczną, ale wymagająca specjalnego potraktowania ze względu na uzależnienie ekonomiczne i społeczne od różnych instytucji i pomocy rodzinnej. Narzędzia badawcze przygotowane do badania zbiorowości generalnych nie chwytają tych zjawisk i mogą prowadzić do wypaczeń w przedstawianiu sytuacji społecznej tej kategorii. Jeśli w przyszłych opracowaniach będziemy wypowiadać się o umiejscowieniu emerytów i rencistów w strukturze społecznej, to raczej w oparciu o specjalistyczne studia nad "ludźmi starszymi".

W badaniu "Struktura społeczna II" przyjęto, że wielkość próby dorosłej ludności kraju w wieku 21-65 lat ma wynieść 6000 osób. Taka wielkość jest uzasadniona celami badania, w wyniku którego rozkład niektórych zmiennych charakteryzujących strukturę społeczną ma być ekstrapolowany na populację generalną. Ponadto, wielkość ta zapewnia możliwość wykonywania wielu złożonych analiz statystycznych. Dodać też należy, że w ramach przyznaných środków budżetowych większa próba byłaby trudna do realizacji przy – wymaganej w takim badaniu – należytej kontroli jakości danych.

Stosując wspólne zasady losowania jednostek obserwacji dobrano próbę zasadniczą (6000 osób) i próbę rezerwową (3167). Doświadczenia zebrane podczas realizacji badania "Struktura społeczna I" wykazały, że czynnikami silnie skorelowanymi z wykorzystaniem rezerwy są: wielkość miejscowości i jej administracyjny status. W związku z tym, w każdym z

województw wyróżniono pewne klasy miejscowości i dla każdej z nich dobrano próbę rezerwową różnej wielkości. Opiszemy tu zastosowaną procedurę.

W województwie stołecznym jako odrębną klasę potraktowano Warszawę w jej administracyjnych granicach, dobierając rezerwę wielkości aż 150% próby zasadniczej, gdyż tu właśnie dotarcie do respondentów z próby zasadniczej okazuje się najtrudniejsze. Pozostałe miasta i wsie województwa stołecznego stanowiły klasę miejscowości, dla której proporcję tę ustalono na poziomie 70%. W dziesięciu województwach – bydgoskim, toruńskim, włocławskim, siedleckim, szczecińskim, jeleniogórskim, wrocławskim, poznańskim, gdańskim i częstochowskim – wprowadzono trzy klasy miejscowości: miasto wojewódzkie, pozostałe miasta i wszystkie wsie. W klasach tych rezerwę ustalono w proporcjach 70%, 50% i 30%, odpowiednio. Wreszcie, w pozostałych województwach rezerwa wynosiła 50% dla miast i 30% dla wsi.

Podstawowa ocena jakości próby w badaniu "Struktura społeczna II" będzie dokonana w oparciu o porównanie rozkładów podstawowych cech społeczno-demograficznych w jej zasadniczej części zrealizowanej i niezrealizowanej oraz w części rezerwowej. Uzupełnieniem tych analiz będą porównania rozkładów podstawowych cech społeczno-demograficznych w tym badaniu i w innych badaniach socjologicznych, obejmujących również dorosłą ludność Polski.

PRÓBA LUDNOŚCI W BADANIU "STRUKTURA SPOŁECZNA-ŁÓDŹ, 1988"

Badania przeprowadzone w Łodzi w latach 1965, 1967, 1976 i 1980 oparte były na próbach mieszkańców tego miasta, ograniczonych do męż-

czyn w wieku 21-65 lat, aktualnie pozostających w związkach małżeńskich. Operatem losowania prób były spisy wyborców lub spisy mieszkańców Łodzi (por. Słomczyński i Wesołowski, 1970; Janicka, 1987). W naszych obecnych badaniach chcieliśmy możliwie dokładnie powtórzyć sposób losowania próby, aby uzyskać porównywalność z próbami zrealizowanymi w przeszłości. Dlatego też za operat losowania posłużyły adresowe *spisy mieszkańców miasta*. W każdej dzielnicy dokonano losowania systematycznego i wybrano potencjalnych respondentów w liczbie proporcjonalnej do liczby mieszkańców.

Próba założona liczyła 750 osób i w całości ją zrealizowano. Trzeba jednak zaznaczyć, że spisy mieszkańców Łodzi zawierają znaczną proporcję adresów nieaktualnych. Ten sam problem wystąpił też w poprzednich badaniach. Należy przypuszczać, iż obecnie zrealizowana próba jest - jako całość - w równym stopniu reprezentatywna, jak próby poprzednich badań łódzkich. To przypuszczenie mamy zamiar poddać weryfikacji stosując już wypracowane środki porównań z danymi dotyczącymi mieszkańców Łodzi, a pochodzącymi z Miejskiego Urzędu Statystycznego.

PRÓBA LUDNOŚCI W BADANIU KONTROLNYM

Badanie kontrolne oparte było na *podpróbie* zrealizowanej części próby badania "Struktura społeczna II". Ze względu na narzędzie badawcze - ankietę pocztową - zdecydowano się na wyłączenie z badania mieszkańców wsi. Uczyniono tak z dwóch względów: po pierwsze, nawyki do korespondencji listowej na wsi nie są silnie ugruntowane; po drugie, poczta na wsi działa znacznie gorzej niż w mieście. Przyjmując to ograniczenie, dokonano systematycznego losowania osób ze zbioru tych respondentów w

badaniu zasadniczym, którzy są mieszkańcami miast. W sumie wylosowano 2043 osób i do nich wysłano ankietę.

Na ankietę od razu odpowiedziały 858 osoby, po pierwszym monicie - 431 osób, zaś po drugim monicie - 164 osoby. Tak więc, z próby liczącej 2043 przypadków, 590 osób w ogóle nie odpowiedziało na ankietę. Z tej liczby "opornych" wylosowano 100 nazwisk i adresów, aby do nich dotrzeć w inny sposób - mianowicie poprzez ankierów. W 94 przypadkach, ankierom udało się odbyć rozmowę z respondentami lub - poprzez inne osoby - ustalić ich losy.

Zasadnicze analizy w badaniu kontrolnym będą prowadzone oddzielnie na próbie osób, które odpowiedziały na ankietę pocztową i na próbie osób, z którymi rozmawiał ankier. Jeżeli jednak okaże się, iż sposób uzyskiwania informacji - poprzez ankietę lub poprzez ankiera - nie ma istotnego wpływu na odpowiedzi respondentów, to obie próby będzie można łączyć. Takie całościowe analizy będą jednak wymagały ważenia próby osób, z którymi rozmawiał ankier, gdyż spośród potencjalnych kandydatów do takiego kontaktu badawczego wylosowano tylko setkę. Metodologia w zakresie ważenia prób jest dobrze opracowana.