

# Rozdział 1

## Efekty oddziaływania i łączenie danych

### 1.1. Wprowadzenie

Metoda łączenia według prawdopodobieństwa jest jednym ze sposobów na szacowanie wielkości efektów oddziaływania. Trudno jest zatem opisać właściwości tej metody estymacji bez odwoływania się do kontekstu, w jakim znajduje ona zastosowanie. W tym celu zostaną zaprezentowane fundamenty analizy kontrfaktycznej. Stanowi ona ważne narzędzie wspomagające ilościową ewaluację programów. Następnie zostaną omówione metody łączenia danych oraz metodologiczne i statystyczne uwarunkowania wykorzystania łączenia według prawdopodobieństwa. Rozdział zamyka dyskusja, czy wykorzystując metody łączenia według prawdopodobieństwa, można odtworzyć kontrolowane randomizowane eksperymenty.

Za idealną metodę badawczą, wyznaczającą złoty standard nauki, są uznawane w pełni kontrolowane randomizowane eksperymenty. Są one domeną nauk eksperymentalnych takich jak fizyka, biologia czy medycyna. W warunkach laboratoryjnych obserwowane jest zachowanie jednostek przydzielonych do grupy eksperymentalnej i identycznych z nimi jednostek przydzielonych do grupy kontrolnej. Na podstawie porównania reakcji na oddziaływanie zaaplikowane jednostkom w grupie eksperymentalnej z zachowaniem jednostek w grupie kontrolnej są wyciągane wnioski dotyczące efektów oddziaływania. W naukach społecznych przeprowadzenie kontrolowanych eksperymentów na szeroką skalę nie jest możliwe. Najbardziej wiarogodne i uznane wyniki w tej dziedzinie pochodzą z badań, w których stosowano randomizację (LaLonde 1986).

Termin „randomizacja” wprowadzili do statystyki niezależnie Neyman (1935) i Fisher (1935). Fisher opisuje randomizację jako „uza-

sadnioną podstawę” wnioskowania i „podstawy dla poprawności testu [statystycznego]”. Randomizacją nazywamy losowe przydzielenie jednostek do grupy eksperymentalnej i kontrolnej. Zatem nie każdy kontrolowany eksperyment jest randomizowany. Nadzór nad warunkami przeprowadzania eksperymentu i manipulowanie nimi nie musi zapewniać efektu randomizacji. Na przykład podczas prób klinicznych nowatorskie leki są serwowane pacjentom cierpiącym na typ choroby, na który lek ma oddziaływać. Zdanie odwrotne, że każdy randomizowany eksperyment jest kontrolowany, również nie jest prawdziwe. Dobrym przykładem jest znany amerykański program wsparcia osób bezrobotnych *National Supported Work Demonstration*, w którym uczestników w sposób losowy przydzielano do różnych form aktywizacji zawodowej. Ponieważ program był interwencją na rynku pracy, nie wszystkie jego efekty były kontrolowane, np. nie można było oddzielić jednostek poddanych różnym formom aktywizacji w sposób zapewniający brak kontaktu między nimi.

W celu przeprowadzenia skutecznej randomizacji badanie musi zostać odpowiednio zaprojektowane. Jednostki poddane eksperymentowi nie muszą być homogeniczne, próba eksperymentalna i grupa kontrolna nie muszą być próbami losowymi z populacji, ale jednostki muszą być w sposób losowy przydzielone do grupy poddanej oddziaływaniu i grupy kontrolnej (Rosenbaum 2002). Jednak w niektórych przypadkach badacze poszukują uzasadnienia empirycznego dla zjawisk, które nie były głównym przedmiotem zainteresowania badania, wykorzystując zgromadzone dane. W takiej sytuacji metody łączenia pozwalają na przekształcenie danych pochodzących z badań ankietowych w zbiór danych spełniający wszystkie warunki, jakim podlegają dane eksperymentalne. W literaturze w ten sposób przekształcone zbiory określa się mianem danych quasi-eksperymentalnych.

## 1.2. Fundamenty analizy kontryfakcyjnej

Randomizacja jest uważana za metodę wyznaczającą standard oceny wpływu efektu oddziaływania danego czynnika czy wprowadzonej interwencji na rezultat. Losowe przydzielenie oddziaływania oznacza,

że przydział jednostki do grupy eksperymentalnej albo kontrolnej odbywa się w sposób niezależny od wyniku oddziaływania (Heckman i Vytlačil 2007). Zapewnia on, że nie będą występowały inne czynniki jednocześnie wpływające na przydzielenie oddziaływania i wynik jego działania (ang. *confounders*). Dzieje się tak, gdyż wynik działania jest niezależny od przydziału jednostki do konkretnej grupy. Zatem efekt oddziaływania może zostać oszacowany poprzez proste porównanie jednostek poddanych oddziaływaniu i jemu nie poddanych (Neyman 1923a i 1923b; Pearl 2000).

W naukach społecznych, w odróżnieniu od nauk eksperymentalnych, przeprowadzenie kontrolowanego eksperymentu i wykorzystanie randomizacji w większości przypadków nie jest możliwe. W pewnych sytuacjach nawet w naukach eksperymentalnych przeprowadzenie eksperymentu może nie być możliwe ze względu na ograniczenia prawne lub powody etyczne. Dobrym przykładem są ograniczenia w testowaniu klinicznym leków o znacznych efektach ubocznych. Z drugiej strony ocena efektu działania czynnika, np. programu czy prowadzonej polityki wymaga informacji, w jaki sposób jednostka zachowałaby się, gdyby nie została poddana oddziaływaniu tego czynnika.

Standardowe ramy analizy kontrfaktycznej i model potencjalnych wyników są przez różnych autorów różnie przypisywane: Neymanowi (1923a i 1923b), Fisherowi (1925), Royowi (1951), Quandtowi (1972) lub Rubinowi (1973a). Neyman (1923a i 1923b) w pracy dotyczącej randomizowanych eksperymentów sformułował problem w języku statystycznym. Następnie Roy (1951) zauważył, że w badaniach ankietowych (obserwacyjnych) obserwowane są wyłącznie najlepsze dla rozpatrywanych jednostek wybory. Uzależnił on podział jednostek na poddane i niepoddane oddziaływaniu od potencjalnych wyników działania. Matematyczne powiązanie obserwowanego wyniku oddziaływania ze średnią ważoną wyniku dla grupy eksperymentalnej i kontrolnej, gdzie wagami są udziały jednostek z grupy eksperymentalnej i kontrolnej w próbie, jest efektem pracy Quandta (1972). Z kolei Rubin (1973a) stworzył regułę podziału jednostek, która jest niezależna od potencjalnego wyniku. Innymi słowy, przydzielenie do jednostek poddanych albo niepoddanych oddziaływaniu nie zależy od potencjalnego wyniku działania bądź braku

działania. W celu oszacowania efektu oddziaływania zaproponował on łączenia podobnych obserwacji i porównanie ich wyników. Ten sposób analizy jest często w literaturze anglojęzycznej przypisywany Royowi (1951) oraz Rubinowi (1973b) i opisywany jako model Roya–Rubina.

Randomizacja i oddziaływanie ściśle wiążą się z pojęciem przyczynowości (Strawiński 2007). Relacja przyczynowo-skutkowa we współczesnej statystyce i ekonometrii jest przedstawiana z zastosowaniem notacji wykorzystującej wyniki kontrfaktyczne (ang. *counterfactual*). W przypadku badań nieeksperymentalnych są one wartościami hipotetycznymi, a zatem nie można tych wyników obserwować w rzeczywistości. Podstawowymi elementami analizy są jednostki, oddziaływanie i potencjalne efekty oddziaływania, czyli potencjalne wyniki. Najprostszym typem oddziaływania, jakiemu mogą być poddane jednostki, jest tzw. oddziaływanie binarne, w którym czynnik oddziałuje na jednostkę albo na nią nie oddziałuje.

Notacja odnosząca się do potencjalnych wyników została zaproponowana w pracach Neymana (1923a i 1923b) i Fishera (1925), dotyczących randomizowanych eksperymentów. Niech zmienna losowa  $P$  opisuje dwa wyróżnione stany: poddania i niepoddania oddziaływaniami. Stan poddania oddziaływaniami oznaczamy  $P_i = 1$  i mówimy wówczas, że jednostka należy do grupy eksperymentalnej (ang. *treated group*). Stan niepoddania oddziaływaniami  $P_i = 0$  nazywamy przynależnością jednostki do grupy niepoddanej oddziaływaniami (ang. *non-treated*) lub grupy kontrolnej (ang. *control treatment*). Oczywiście role stanu  $P_i = 0$  i  $P_i = 1$  można odwrócić (Heckman, LaLonde i Smith 1999). Zatem wybór stanu  $P_i = 1$  często jest arbitralny. W przypadku analiz prowadzonych w naukach społecznych stan  $P_i = 0$  jest w naturalny sposób przypisywany brakowi interwencji. Z tego powodu w literaturze ekonomicznej często zamiast o pojęciu grupy eksperymentalnej mówi się o grupie jednostek poddanych działaniu programu bądź polityki. Dodatkowo przyjmijmy następujące oznaczenia. Niech  $Y_{1i}$  będzie wynikiem działania w sytuacji, gdy jednostka  $i$  jest poddana oddziaływaniami lub należy do grupy eksperymentalnej, a  $Y_{0i}$  to wynik, gdy pozostaje poza wpływem działania lub należy do grupy pozostającej poza programem. W tym miejscu warto dokonać rozgraniczenia dwóch pojęć: grupa

niepoddana oddziaływaniu i grupa kontrolna. Do pierwszej grupy należą wszystkie jednostki niepoddane oddziaływaniom, podczas gdy grupę kontrolną tworzą wyłącznie te jednostki niepoddane oddziaływaniom, które są podobne do jednostek z grupy poddanej działaniu programu (nazywanej również grupą eksperymentalną).

W każdym badaniu analizującym wpływ oddziaływania binarnego niezależnie od tego, czy jest ono przeprowadzone metodą eksperymentalną, czy nieeksperymentalną, każda jednostka znajduje się pod wpływem czynnika, który na nią oddziałuje, albo jest tego wpływu pozbawiona. Problem identyfikacji efektu oddziaływania jest związany z faktem, że dla każdej jednostki  $i$  badacz obserwuje tylko jedną wartość wyniku  $Y_i$ . Ponadto w przypadku danych pochodzących z badania ankietowego badacz zazwyczaj nie posiada informacji, czy jednostka  $i$  była, czy nie była poddana oddziaływaniu czynnika. Formalnie zależność między obserwowanym wynikiem a rzeczywistymi wynikami działania można zapisać, wykorzystując model przełącznikowy zaproponowany przez Quandt (1972).

$$Y_i = P_i Y_{1i} + (1 - P_i) Y_{0i}. \quad (1.1)$$

W rezultacie jest obserwowany uśredniony wynik  $Y_i$ . Jest to średnia ważona wyniku poddania działaniu i wyniku pozostawiania poza jego wpływem, w której wagi są proporcjonalne do proporcji liczby jednostek poddanych oddziaływaniu względem liczby jednostek jemu niepoddanych. Jest to uśredniony efekt dla grupy.

Wielkością interesującą badaczy jest rezultat wpływu oddziaływania wobec konkretnej jednostki. Na przykład, jeżeli rozpatrujemy program szkoleń zwiększających szanse osób na rynku pracy, jest interesujące, w jakim stopniu szkolenie pomogło konkretnym uczestnikom. Efekt oddziaływania wobec jednostki  $i$  można zapisać jako:

$$\tau_i = Y_{1i} - Y_{0i}. \quad (1.2)$$

W tym miejscu pojawia się problem, ponieważ dla każdej jednostki  $i$  jest obserwowany tylko jeden potencjalny wynik  $Y_{1i}$  albo  $Y_{0i}$ . Osoba bezrobotna uczestniczy w programie szkoleń albo w nim nie uczestniczy. Nieobserwowany potencjalny wynik jest w literaturze nazywany wyni-