



Tadeusz Z. Leszczyński

Metoda matchingu w naukach o bezpieczeństwie*

Badacz zamierzający prowadzić badania w obszarze bezpieczeństwa państwa, który pragnie przetestować własną hipotezę badawczą, staje przed potrzebą wcześniejszego rozwiązania szeregu problemów. Ułatwia mu to przyjęcie prawidłowo skonstruowanego planu badawczego, rozumianego jako logiczny model wnioskowania przyczynowo-skutkowego. Najczęściej jest to plan eksperymentalny, którym w naukach społecznych obejmowane są dwie grupy osób: grupa eksperymentalna (poddawana działaniu zmiennej niezależnej) i grupa kontrolna (niepoddawana powyższemu działaniu), przy czym przyporządkowanie osób do poszczególnych grup w metodach eksperymentalnych odbywa się na drodze losowej. Oszacowanie wpływu zmiennej niezależnej polega na dwukrotnym pomiarze zmiennej zależnej w obu grupach oraz porównaniu wielkości różnicy miar pomiędzy pomiarem początkowym i pomiarem końcowym w każdej grupie.

Wiele zjawisk interesujących badaczy w naukach społecznych nie poddaje się bezpośredniemu eksperymentowaniu. Wówczas stosowane są plany badań przekrojowych, plany *quasi*-eksperymentalne i *pre*eksperymentalne, a także plany łączone, stanowiące kombinację powyższych planów. Badania z wykorzystaniem planu przekrojowego rozpoczyna się od utworzenia losowej, reprezentatywnej próby badawczej (mężczyzn i kobiet), której następnie zadawane są pytania sondażowe. Wyniki badań sondażowych wykorzystywane są zwykle do określania relacji przyczynowo-skutkowych pomiędzy właściwościami a dyspozycjami bądź do opisu zależności pomiędzy zmiennymi. Plan przekrojowy ma wiele ograniczeń metodologicznych, np. dotyczących jego trafności wewnętrznej, stąd do ich przewyciężenia wykorzystywane są metody statystyczne odpowiadające niektórym operacjom wbudowanym w plan eksperymentalny. Stosując plany *quasi*-eksperymentalne, tak jak plany przekrojowe słabsze

* Artykuł wyraża poglądy i obserwacje autora, które zostaną przedstawione w książce pod tym samym tytułem (przyp. red.).

pod względem trafności wewnętrznej od planów eksperymentalnych, jako metody kontroli należy stosować techniki analizy danych. W metodach *quasi*-eksperymentalnych można losowo pobierać próbę z populacji, lecz nie należy losowo przyporządkowywać osób badanych do porównywanych grup. Plany preeksperymentalne, często pozbawione grupy kontrolnej, są najslabszymi planami badawczymi, gdyż nie kontroluje się w nich większości źródeł trafności wewnętrznej i zewnętrznej. Ze względu na duże prawdopodobieństwo popełnienia błędu przy wyprowadzaniu wniosków przyczynowo-skutkowych plany te najczęściej stosuje się przy wstępnym testowaniu niektórych hipotez badawczych bądź w badaniach eksploracyjnych (np. analizie przypadku polegającej na dokonaniu pomiaru jednej grupy lub jednego zdarzenia w jednym punkcie czasowym).

Porównując możliwe do zastosowania w naukach społecznych plany badawcze, należy stwierdzić, że plany umożliwiające uzyskanie jednoznacznych danych dotyczących przyczynowości, tzn. mocne z punktu widzenia trafności wewnętrznej, zwykle prowadzą do obniżenia możliwości uogólniania, tj. są słabe ze względu na trafność zewnętrzną. Tendencję tę potwierdza obserwacja, że plany badawcze słabe pod względem trafności wewnętrznej są także słabe pod względem trafności zewnętrznej, gdyż niemożliwe jest wówczas dokonywanie jakiegokolwiek generalizacji. Jak twierdzą C. Frankfort-Nachmias oraz D. Nachmias¹, eksperymenty są silne ze względu na możliwość kontroli, lecz słabe ze względu na reprezentatywność, zaś badania przekrojowe i *quasi*-eksperymentalne wykazują tendencje odwrotne. Ponadto eksperymenty umożliwiają wyprowadzanie wniosków przyczynowo-skutkowych dzięki kontroli czynników wewnętrznych i zewnętrznych względem planu badawczego oraz pozwalają na kontrolę procedury wprowadzania zmiennej niezależnej, a stąd na określanie kierunku zależności przyczynowo-skutkowej, zaś pozostałe przedstawione powyżej plany badawcze są tych właściwości pozbawione.

Spółeczeństwa i ich państwa muszą przygotowywać się do przyszłych działań zmierzających do reagowania na przewidywane oraz nawet ledwie uświadamiane zagrożenia bezpieczeństwa. Stąd hipotezy badawcze w obszarze bezpieczeństwa państwa zwykle dotyczą ewentualnego naruszenia żywotnych interesów państwa i jego obywateli w stopniu, który stwarza potencjalne zagrożenia dla ich przetrwania. W naukach społecznych hipoteza badawcza to nic innego jak proponowana przez badacza odpowiedź, jakiej można udzielić na pytanie badawcze, wyrażona w postaci związku pomiędzy zmienną zależną i zmienną niezależną, którą można zweryfikować dopiero po przeprowadzeniu badań. Jedną z metod kontroli wykorzystywanych w metodach badawczych stosowanych w naukach społecznych polegającą na wyrównywaniu grupy eksperymentalnej i grupy kontrolnej ze względu na wyniki niezwiązane z planem badawczym, o których można założyć, że mają związek z hipotezą badawczą, jest metoda *matching* (z ang. dobrze dobrany, pasujący), zwana także metodą doboru wiązanego.

Matching wywodzi się z fotogrametrii, w której pierwotnie wykorzystywano metody analogowe polegające na „ręcznym” pomiarze i porównywaniu punktów na zdjęciach. Powszechnie wykorzystywany był w czasie II wojny światowej do analizy zdjęć lotniczych. Stąd już w latach pięćdziesiątych ubiegłego wieku opracowano korelatory poziomów szarości dwóch obrazów, co znacznie przyspieszyło analizy rozpoznawcze.

¹ Por. C. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Poznań 2001, s. 164.

Zalety powyższej metody mogły zostać uwypuklone dopiero wraz z rozwojem fotogrametrii cyfrowej, której początki datuje się na lata osiemdziesiąte XX wieku, kiedy rozwój komputerów umożliwił obróbkę obrazów cyfrowych, a kamery cyfrowe umożliwiły pomiary w trybie *on line* lub nawet w czasie rzeczywistym. Metoda została wykorzystana m.in. do automatycznego nadzorowania i sterowania procesami przemysłowymi, badań w medycynie, transporcie, a także w przemyśle kosmicznym, lotniczym, okrętowym, nuklearnym, motoryzacyjnym itd.

Aktualnie można mówić o dwóch głównych sposobach pozyskiwania obrazów cyfrowych, np. obszarów występowania sytuacji kryzysowej:

- sposób bezpośredni – zapis przestrzeni przedmiotowej za pomocą urządzeń pozwalających rejestrować obraz bezpośrednio w formie cyfrowej za pomocą odpowiednich sensorów,
- sposób pośredni – doprowadzanie do postaci cyfrowej istniejących materiałów analogowych, np. poprzez skanowanie zdjęć.

Wykorzystywana w fotogrametrii cyfrowej metoda matchingu polega na identyfikacji i pomiarze punktów homologicznych na dwóch lub więcej obrazach. Jej celem jest wybranie obiektu na jednym obrazie i odszukanie odpowiednika na drugim, stąd proces ten nazywany jest z ang. *image matchingiem*. Natomiast w aerotriangulacji wykorzystywany jest *matching obrazów sekwencyjnych*, tj. obrazów pozyskanych kamerami wideo.

Metody analizy obrazów można podzielić na:

- *Area Based Matching* – porównywanie skali szarości grupy pikseli, a w przypadku obrazu kolorowego wykorzystywanie jednego z kolorów do korelacji,
- *Feature-Based Matching* – porównywanie krawędzi lub innych obiektów z obrazów oryginalnych z odpowiednimi homologicznymi obiektami na drugim lub kilku innych obrazach,
- *Symbolic Matching* – porównywanie własności topologicznych obiektów zaimplementowanych do systemu jako np. grafy, drzewa lub sieci semantyczne.

Wykorzystanie matchingu sprowadza się w praktyce do czterech etapów:

- wyboru elementów dopasowania,
- znalezienia ich odpowiedników w innym obiekcie lub procesie,
- określenia współczynnika korelacji, np. przestrzennego dopasowania elementów,
- oszacowania dokładności dopasowania.

Należy zaznaczyć, że metoda matchingu wykorzystywana jest w fotogrametrii do różnych celów, w tym:

- kalibracji,
 - orientacji bezwzględnej,
 - orientacji wewnętrznej,
 - orientacji wzajemnej,
 - aerotriangulacji,
 - generowania numerycznego modelu terenu.
- Współczynnik korelacji (ρ) można zdefiniować jako²:

² D.B. Rubin, *Assignment to treatment group on the basis of a covariate*, „Journal of Educational Statistics” 1977, nr 2, s. 1-26.

$$\rho = \frac{\sigma_{LR}}{\sigma_L \sigma_R}$$

gdzie:

- σ_L - odchylenie standardowe obrazu (procesu) L - wzorca,
- σ_R - odchylenie standardowe obrazu (procesu) R - porównywanego ze wzorcem,
- σ_{LR} - kowariancja fragmentów obrazów (elementów procesów) L i R.

Znormalizowany współczynnik korelacji może przyjmować wartości z przedziału zamkniętego $\langle -1; 1 \rangle$, w tym w szczególności:

- -1 w przypadku porównywania diapozytywu i negatywu oraz procesu odtwarzanego w odwrotnej kolejności względem procesu wzorcowego,
- 0 przy braku korelacji,
- 1 w przypadku, gdy macierz wzorca pokrywa się z macierzą porównywanego obrazu (procesu).

Metoda matchingu znajduje praktyczne zastosowanie m.in. w zapewnianiu bezpieczeństwa personalnego i ekonomicznego organizacji, np. przy wartościowaniu stanowisk pracy w zarządzaniu personelem, zabezpieczaniu przedsiębiorstw i instytucji publicznych przed ryzykiem walutowym oraz zarządzaniu kryzysowym.

Wartościowanie stanowisk pracy, stanowiące podstawę każdego systemu wynagrodzeń, polega na określeniu relatywnego miejsca poszczególnych stanowisk w hierarchii płacowej ogółu stanowisk w instytucji lub przedsiębiorstwie i jest środkiem do utworzenia *tabel zaszerogowania stanowisk*, czyli również zbudowania hierarchii płacowej grup stanowisk o podobnej złożoności. Powyższy proces rozpoczyna się zawsze od dokładnego poznania organizacji instytucji (firmy), w tym zakresu zadań na poszczególnych stanowiskach służbowych oraz ich charakteru i wpływu na wynik całej organizacji. W efekcie zidentyfikowane zostają grupy stanowisk o podobnym charakterze pracy, ale o zróżnicowanym poziomie wymagań kwalifikacyjnych. Następnie ogół stanowisk dzielony jest na grupy o podobnym charakterze pracy i wymogach, po czym z pomocą kierownictwa organizacji wytyczane są ścieżki rozwoju oraz określone wymogi kompetencyjne dla każdego stanowiska w ścieżce. Tak stworzone mapy kariery zwykle zostają wyposażone dodatkowo w profile typowych stanowisk w poszczególnych rodzinach stanowisk. *Job matching* zwykle stosowany jest w organizacjach, które w znacznym stopniu są zorientowane na konkurencję na rynku pracy, dążą do szybkiej integracji różnych narzędzi kadrowych na bazie języka kompetencji bądź chcą prowadzić zróżnicowaną politykę płacową dla poszczególnych rodzin stanowisk. Powyższą metodę wykorzystują także firmy HRS świadczące usługi w zakresie wyszukiwania pracowników, portale kojarzące pary, a także firmy oferujące pośrednictwo w znalezieniu partnerów handlowych. Po otrzymaniu wymagań sprecyzowanych przez kierownictwo organizacji, firmy te dokonują dokładnej analizy danych zebranych dzięki testowi osobowości lub po zapoznaniu się z wygenerowanymi potrzebami biznesowymi wyszukują za pomocą metody matchingu osoby najbardziej zbliżone do wzorca (ideału), wyznające podobne wartości czy mające zbliżone zainteresowania lub partnerów strategicznych dostosowanych do wymagań biznesowych zleceńodawcy.

Matching w finansach jest jedną z naturalnych metod zabezpieczania ryzyka walutowego, a więc zwiększenia bezpieczeństwa ekonomicznego, która polega na wewnętrznej kompensacie należności i zobowiązań denominowanych w jednej walucie, powodującej, że na ryzyko walutowe narażone jest tylko saldo (wartość netto) przepływów. Metoda ta jest szczególnie przydatna w przedsiębiorstwach (rzadziej instytucjach publicznych), które dysponują stałymi dwukierunkowymi przepływami finansowymi w tej samej lub w różnych walutach obcych. Przedsiębiorstwa i instytucje tego typu zabezpieczają się przed ryzykiem walutowym, wykorzystując metodę matchingu do podporządkowania wielkości zobowiązań i terminów ich spłaty pod wielkość należności i terminy ich wpływu. W powyższym celu wykorzystywany jest matching naturalny lub matching równoległy. Matching naturalny polega na powiązaniu wpływów i płatności wyrażonych w tej samej walucie obcej, w taki sposób, by były one zbliżone pod względem wartości i czasu występowania. Matching równoległy polega na wiązaniu wpływów i płatności wyrażonych w różnych walutach obcych (dodatkowo ze sobą skorelowanych). W obu powyższych przypadkach dopasowanie płatności do wpływów może zapewnić pełne pokrycie lub zmniejszenie otwartych pozycji oraz wyeliminowanie zbędnego obrotu środków płatniczych na rynku. Metoda ta jest dostępna zwłaszcza dla większych przedsiębiorstw, które w dowolny sposób mogą generować przychody lub zaciągać zobowiązania w różnych walutach. W szczególności przedsiębiorstwa eksportujące swoje produkty stale są narażone na wahania kursu walutowego, które można częściowo ograniczyć poprzez zwiększenie kosztów indeksowanych tą samą walutą. Eksporter może np. importować surowce lub półprodukty od dostawców, którym będzie płacił w walucie, w jakiej realizuje eksport. Spadek kursu walutowego spowoduje czasowy względny spadek przychodów (w PLN), gdyż spadną wówczas także koszty operacyjne, ze względu na zmniejszenie kosztów zakupu surowców i materiałów (wzrost kursu spowoduje wzrost przychodów eksportera oraz wzrost kosztów operacyjnych).

Z zastosowaniem matchingu mamy do czynienia w zapewnianiu bezpieczeństwa na lotniskach. Metodę tę można wykorzystać również np. w metrze, w porcie morskim, na dworcu kolejowym i w innych miejscach publicznych. Używane są tam systemy do rozpoznawania twarzy osób poszukiwanych bądź niepożądanych. W tym celu wykonywane są stereogramy twarzy podejrzanych osób na podstawie dostępnych zdjęć bądź nagrań wideo, zwykle sporządzone z wykorzystaniem dwóch manualnie ogniskowanych kamer. Następnie wytworzone stereogramy w postaci modeli 3D, przedstawiające wirtualne odwzorowanie odpowiednio dobranych fotopunktów twarzy na siatce przestrzennej gromadzone są w bazach danych, z którymi porównywane są obrazy twarzy osób przebywających w miejscu objętym monitoringiem uzyskiwane z kamer rozmieszczonych w powyższych instytucjach oraz sprzężone z innymi systemami ochrony. Metoda ta zdobywa coraz większe zastosowanie w medycynie, zwłaszcza w laparoskopii, tomografii komputerowej i rezonansie magnetycznym.

Metoda matchingu wykorzystywana jest także w ewaluacji *ex-post* do oszacowania efektu przyczynowego podejmowanych interwencji. W przypadkach gdy zastosowanie planu eksperymentalnego nie jest możliwe, np. ze względu na koszty realizacji lub konieczność nieetycznego przymusu partycypacji w grupach porównaw-

czych³, wówczas użyteczna jest m.in. metoda oceny skuteczności interwencji (ang. *propensity score matching*). Jest to metoda nieeksperymentalna, należąca do grupy metod wykorzystywanych w tzw. badaniach obserwacyjnych, zbliżona do podejścia stanu kontrfaktycznego, dominującego dotychczas w analizie przyczynowej w badaniach ewaluacyjnych. Jednak w przeciwieństwie do swej poprzedniczki, metoda ta oferuje sposób korekty obciążenia szacunków efektu netto spowodowanego selekcją, tj. próbuje wyeliminować obciążenia wynikające z nieodłącznych różnic pomiędzy grupą interwencji a grupą kontrolną. Wynika to z faktu, że zwykle brak jest informacji, jakie jednostki zostały objęte działaniem danego zdarzenia, a jakie nie. Podstawą metody *propensity score matching* jest zbalansowanie X , tj. utworzenie grupy kontrolnej, która będzie miała taki sam rozkład zmiennych niezależnych jak grupa interwencji. Służy temu wykorzystanie pewnej funkcji X , która posiada tzw. właściwości balansujące. Funkcja ta spełnia warunek⁴:

$$X \perp D \mid b(X)$$

Powyższy zapis oznacza, że warunkowy rozkład X , względem wartości funkcji $b(X)$ jest taki sam dla jednostek w grupie eksperymentalnej ($D=1$), jak dla jednostek w grupie kontrolnej ($D=0$). Stąd warunkowe prawdopodobieństwo doboru obserwacji do zdarzenia ($D=1$) w metodzie *propensity score matching* szacowane jest względem wektora zmiennych X :

$$P(X) = P(D = 1|X)$$

Omawiana metoda ma na celu utworzenie grupy kontrolnej, składającej się z jednostek w jak największym stopniu podobnych do tych, które znalazły się w grupie eksperymentalnej. Wykorzystywane w niej warunkowe prawdopodobieństwo doboru obserwacji do zdarzenia (*propensity score*) stanowi jedyną podstawę, według której dokonywane jest dopasowywanie. Zmienna ta, czasem definiowana jako wskaźnik skłonności do partycypacji w warunkach interwencji, umożliwia więc redukcję cech / wymiarów, za pomocą których można opisywać obserwacje w zbiorze danych. Od strony technicznej metoda *propensity score matching* obejmuje:

- wyliczenie wartości *propensity score*, np. z wykorzystaniem modelu regresji logistycznej, w którym zmienną zależną jest fakt bycia w grupie objętej oddziaływaniem bodźca, zaś zmiennymi niezależnymi cechy, które mają wpływ na wynik oraz na uczestnictwo w danym działaniu,
- dobór jednostek do grupy kontrolnej na podstawie oszacowanego wyżej wskaźnika, np. z użyciem metody najbliższego sąsiada polegającej na dopasowaniu jednostek najbardziej podobnych, efektem łączenia będzie grupa kontrolna, która będzie miała zbalansowane wszystkie zmienne obserwowalne, wykorzystane w modelu prawdopodobieństwa, a więc w zakresie wybranego zestawu cech podobne do istniejącej grupy interwencji,
- analiza efektów w oparciu o porównanie grupy interwencji z grupą kontrolną.

³ Jak zauważają R. Konarski i M. Kotnarowski, osoby biorące udział w badaniu mogą nie wyrażać zgody na poddanie się manipulacji eksperymentalnej, a równocześnie brak takiej manipulacji może być nieetyczny, zwłaszcza gdyby była ona korzystna. Por. R. Konarski i M. Kotnarowski, *Zastosowanie metody propensity score matching w ewaluacji ex-post*, [w:] *Ewaluacja ex-post. Teoria i praktyka badawcza*, red. A. Huber, PARP, Warszawa 2007, s. 183.

⁴ R. Trzciński, *Wykorzystanie techniki propensity score matching w badaniach ewaluacyjnych*, Warszawa 2009, s. 27 i kolejne.

Metoda matchingu w naukach o bezpieczeństwie

Punktem wyjścia do zastosowania metody matchingu jest dostępność wzorca. W niektórych sytuacjach możliwe jest uzyskanie dostępu do wzorca kosztem poniesienia niezbędnych nakładów finansowych lub zużycia jednostek czasu, jednak w większości przypadków parametry wzorca dostępne są jedynie z dającym się oszacować – bądź nie – prawdopodobieństwem. By uzyskać przybliżony model wzorca, należy dysponować pewną minimalną grupą kontrolną elementów istotnych, które winny zostać zebrane w zestandaryzowany sposób oraz w określonym czasie. Istotne są tylko te elementy modelu wzorca, których zmiana będzie oddziaływała na wynik końcowy oraz na wybór relacji przekształcającej. Ponadto niezbędna jest ocena precyzji estymacji modelu wzorca.