

Dylematy kwantyfikacji geografii

Kiedy przed blisko dwudziestu laty przygotowywałem swój pierwszy artykuł na temat metod statystycznych w geografii, trudno było znaleźć przykłady takich zastosowań. Literatura na ten temat nie zawierała zagadnień dotyczących specyfiki zastosowań w geografii. W podobnej sytuacji był również W. Garrison (1956) późniejszy czołowy propagator i twórca amerykańskiej szkoły matematycznej w geografii. Jedynie w geografii szwedzkiej sytuacja przedstawiała się lepiej dzięki pracom T. Hägerstranda.

Dzisiaj sytuacja przedstawia się radykalnie odmiennie. Już w latach sześćdziesiątych nastąpił bardzo silny rozwój zastosowań metod ilościowych w badaniach geograficznych. Zjawisko to zostało nawet nazwane przez I. Burtona (1963) „rewolucją ilościową”. Pojęcie kwantyfikacji jest traktowane jako synonim zastosowania statystyki i matematyki do badań geograficznych. Nie jest to określenie zbyt ścisłe, gdyż nie chodzi tu tylko o pojęcia ilościowe, lecz również o te właściwości i relacje jakościowe, które znajdują izomorficzną reprezentację w matematycznych właściwościach i relacjach charakteryzowanych w terminach teorii mnogości oraz topologii.

Współcześnie najsilniejszy zespół badaczy zajmujących się zastosowaniami matematyki i statystyki w analizie przestrzennej zjawisk społeczno-ekonomicznych ukształtował się w Stanach Zjednoczonych i Kanadzie. Obok geografów ekonomicznych (W. Garrison, B. J. L. Berry, M. F. Dacey, L. Curry i in.), zajmujących się przede wszystkim analizą przestrzenną, problematykę tę rozwinęli ekonomiści regionalni, czyli przedstawiciele tzw. *regional science*, której twórcą jest W. Isard. Uwaga tych ostatnich koncentrowała się jednak bardziej na budowie modeli matematycznych o charakterze optymalizacyjnym, związanych z planowaniem przestrzennym i regionalnym.

W rozwoju matematyzacji geografii najbardziej twórczy intelektualnie wkład wnieśli jednak geografowie szwedzcy, a zwłaszcza T. Hägerstrand. W ostatnich latach kierunek matematyczny rozwinął się także bardzo silnie w Związku Radzieckim i Wielkiej Brytanii. Geografowie radzieccy (J. Sauszkin, B. Gurewicz, J. Miedwiedkow) oraz ekonomiści, zajmujący się problematyką regionalną (ośrodek w Nowosybirsku), podjęli wiele badań w tym kierunku i sformułowali nowe koncepcje metodologiczne, nawiązujące do problematyki ekologicznej oraz planowania regionalnego.

Ujęcie ilościowe znamionuje też w ostatnich latach geografę brytyjską, której czołowi przedstawiciele, R. Chorley i P. Haggett, rozwinęli koncepcję systemową w geografii fizycznej i ekonomicznej.

We współczesnej światowej literaturze geograficznej kierunek matematyczny ma więc już dzisiaj utrwaloną pozycję, czego dowodem jest między innymi to, że większość czasopism publikuje prace o charakterze ilościowym, a niektóre z nich, jak na przykład *Geographical Analysis*, są wyłącznie temu poświęcone. Również na terenie Międzynarodowej Unii Geograficznej znalazło to swój wyraz w powołaniu Komisji Metod Ilościowych oraz w znacznym udziale tej problematyki w pracach innych komisji, a także różnych zjazdów i sympozjów.

Chociaż współczesna faza zastosowania metod matematycznych rozpoczęła się w Polsce w końcu lat pięćdziesiątych, to jednak jej tradycje badawcze są starsze. Chodzi mi tu o opublikowaną w 1908 r. pracę B. Janowskiego *O odległościach jako czynniku rozwoju kultury* będącą pierwszą w literaturze światowej próbą sformułowania modelu grawitacji oraz o metodę taksonomiczną J. Czekanowskiego. Nie będę szerzej omawiał tego zagadnienia, gdyż jest ono dość znane, a podsumowania dorobku okresu powojennego w tym zakresie dokonała T. Czyż (1973) na łamach *Przeglądu Geograficznego*.

Sądzę natomiast, że istotnym momentem w rozwoju zastosowań metod matematycznych jest przejście od prac programowych, kładących nacisk na rolę metod, do szerokiego ich stosowania przy rozwiązaniu różnych problemów badawczych o charakterze współuczestniczącym i to w takim zakresie w jakim jest to potrzebne do uzyskania lepszych efektów poznawczych. W takim ujęciu metody matematyczne stają się jednym z instrumentów badawczych, którego zastosowanie ma na celu rozwiązanie określonego problemu badawczego. Obecnie jesteśmy właśnie w toku przechodzenia do tego etapu.

Przejście od eksperymentowania do problemowego stosowania metod

ilościowych zmienia nasz stosunek do tego zagadnienia. Nie warto chyba już szerzej mówić o zaletach tych metod i podnosić ich wyższości, należy natomiast zwrócić uwagę na problemy związane ze swoistymi trudnościami i warunkami ich zastosowań, a więc tym co nazwałbym metodologicznymi dylematami kwantyfikacji geografii.

W związku z tym chciałbym zwrócić uwagę na niektóre z takich problemów.

Pierwszy problem dotyczy miejsca i roli tzw. modelu matematycznego w postępowaniu badawczym geografa. Nie wdając się w szersze rozważania na ten temat przypomnę, że modelem matematycznym nazywa się zwykle zbiór sformułowanych w języku matematyki postulatów (twierdzeń), charakteryzujących pod względem swojej istotności pewne procesy lub zjawiska. Korzyści wynikające z operowania językiem matematyki są na ogół znane, więc nie będę się nimi zajmował. Warto natomiast zwrócić uwagę na rolę jaką pełni model matematyczny.

W naukach społeczno-ekonomicznych przyjmuje się, że służy on aproksymacji i konkretyzacji teorii. Brak jednak dobrze rozwiniętych teorii w geografii sprawia, że większość modeli konstruowanych jest *ad hoc* i służy rekonstrukcji pewnych sytuacji problemowych, na przykład model grawitacji opisuje strukturę oddziaływania w przestrzeni społeczno-ekonomicznej. Powstaje w związku z tym zagadnienie uzasadnienia takiego postępowania. Sądzę, że prawidłowa konstrukcja modeli matematycznych wymaga realizacji dwóch procedur: idealizacji i konkretyzacji. Poprawność procedury idealizacyjnej opiera się na próbie uchwycenia istotnych czynników dla ustalanych zależności. Zasadnicza trudność tkwi w tym, że przy braku teorii nie ma sposobu ustalenia czy przyjęte w modelu czynniki są rzeczywiście istotne. Należy przy tym zauważyć, że sama procedura konkretyzacji modelu nie prowadzi do testowania teoretycznego, gdyż pomija możliwość oddziaływania czynników w modelu nie uwzględnionych.

Drugi problem dotyczy sensu teoretycznego używanych w badaniach wskaźników.

Badania ilościowe wykazują rosnący dystans między precyzją analitycznego aparatu matematycznego a jakością materiału empirycznego na jakim opiera się obliczenia. Stosowanie na wielką skalę metod numerycznych z użyciem komputerów prowadzi do obliczeń wielkiej ilości para-

metrów statystycznych, głównie współczynników korelacji o ustalonych stopniach istotności. Nie chodzi jednak tutaj o przestrzeganie przed niskim stopniem wiarygodności danych wyjściowych, gdyż jest to zagadnienie zbyt dobrze znane, aby trzeba się było nad nim tutaj rozwodzić.

Jak wiadomo, wśród przedmiotów badania wyróżnia się właściwości dostępne bezpośredniej obserwacji oraz takie, które jej nie podlegają. Chodzi mi właśnie o tę drugą grupę przedmiotów obserwacji, o dane, które stanowią wskaźniki empiryczne dla właściwości „ukrytych”, niedostępnych bezpośredniej obserwacji, takich jak postawy, dyspozycje lub cechy osobowości, właściwości grup społecznych oraz układów przestrzennych i regionalnych.

Wskaźnik empiryczny jest zawsze pewną właściwością, która połączona jest jakimś stałym związkiem z pewną inną właściwością, co pozwala stwierdzić obecność jednej na podstawie występowania drugiej. Istota tego związku jest jednak rozmaicie pojmowana, zwykle wymaga się by miał on charakter bezwyjątkowej lub statystycznej regularności.

Prawidłowa konstrukcja wskaźników empirycznych przesądza o wartości uzyskanych danych i wymaga rozwiązania oraz rozstrzygnięcia wielu problemów natury metodologicznej. Na terenie badań geograficznych występują przede wszystkim probabilistyczne wskaźniki empiryczne, a właśnie sytuacja metodologiczna tych wskaźników jest daleka od wyjaśnienia.

Jak to stwierdza T. Pawłowski (1969, s. 199) zdania wprowadzające wskaźniki nie mogą być formułowane w oderwaniu, lecz powinny być częścią składową jakiegoś zbioru twierdzeń stanowiącego teorię naukową. Zależność wskaźnika od teorii występuje szczególnie wyraźnie w przypadku wskaźników probabilistycznych. Struktura tych wskaźników oraz postać funkcji rozkładu prawdopodobieństwa muszą być wyznaczone przez założenia teorii. Tymczasem praktyka badawcza geografii z reguły rozwiązuje to zagadnienie *ad hoc* bez uwzględnienia przesłanek teoretycznych. Stąd też, gdy dane zebrane na podstawie tak ustalonych wskaźników wykorzystywane są następnie do weryfikacji hipotez, trudno uzasadnić sens takich operacji.

Ściśle z tym związane jest zagadnienie teoretycznej standaryzacji pojęć stosowanych w badaniach geograficznych. Problemy te powstają przede wszystkim w związku z porównywaniem wyników badań przeprowadzanych w ujęciu regionalnym.

Standaryzacja zbierania i opracowywania materiałów wymaga też zdania sobie sprawy z tego, że większość wskaźników jest ograniczona do pew-

nych systemów społeczno-ekonomicznych. Stąd też konieczność systematyzacji aparatury pojęciowej i określenia jakie zestawy alternatywnych wskaźników winny być stosowane w różnych układach regionalnych. Przykładem może tu być choćby pojęcie wsi i miasta.

Jednocześnie określone zespoły zmiennych posiadają wyraźne zasięgi przestrzenne (regionalne) ze względu na geograficzny zakres lokalizacji ich desygnatów. Jeśli dane pojęcie odpowiadać będzie zakresowi tych desygnatów, to teoretyczna ogólność takiego pojęcia w hipotezie będzie prowadzić do ograniczenia zakresu sensowności hipotezy i ograniczy jej możliwości porównawcze. Tak na przykład określenie związku między decyzjami producentów kólek rolniczych a typem użytkowania ziemi ogranicza zasięg empirycznej sensowności takiej hipotezy do obszaru Polski.

Trzeci problem dotyczy estymacji statystycznej i testowania modeli w geografii. Chodzi tu o dwa zagadnienia.

Pierwsze jest związane z tzw. krytyką Galtona i odnosi się do postulatu metodologicznego głoszącego, że warunkiem przydatności różnych przypadków dla potwierdzenia hipotezy statystycznej jest ich niezależność. Realizacja tego postulatu napotyka w geografii zarówno w stosunku do zjawisk społeczno-ekonomicznych, jak i fizycznych na trudności powstające wskutek tego, że pewne zjawiska lub związki zachodzące między nimi mogą wynikać z ukształtowania ich przez te same czynniki lub systemy kulturowe, a więc stanowić zależności funkcyjne pozorne. Przykładem może tu być badanie zależności jaką jest zbieżność dwóch fotografii w tysiącach egzemplarzy tej samej gazety pochodzącej od wspólnej matrycy drukarskiej.

Na problem ten zwrócił po raz pierwszy uwagę F. Galton w dyskusji nad zastosowaniem analizy korelacyjnej do uzasadniania hipotez w antropologii społecznej. W dyskusji tej Galton stwierdził: „byłoby niezmiernie pożądane dla wszystkich, którzy by chcieli posiadać dowody słuszności wniosków Taylora, aby można było stwierdzić w jakim stopniu obyczaje plemion i ras były niezależne. Mogłoby się zdarzyć, że pewne ze szczepów przyjęły je ze wspólnego źródła tak, iż były one duplikatami tego samego oryginału” (R. Narrol, 1966).

Na zagadnienie to zwrócił też uwagę S. Ossowski (1963) stwierdzając, że zależności odkrywane w ramach pewnych ciągów genetycznych, tj. serii zjawisk pochodzących ze wspólnego źródła, mogą być zależnościami pozornymi.

Sądzę, że trudności te wymagają w geografii wyjaśnienia roli genetycznej determinacji zjawisk w analizie modelowo-statystycznej. Należy ustalić jakie związki funkcjonalne są wynikiem układów ukształtowanych w przeszłości, aby uniknąć pozornych wyjaśnień funkcjonalnych.

Drugie zagadnienie dotyczy prób uzasadnienia generalizacji wykrytych na podstawie zależności statystycznych. W większości badań geograficznych mamy do czynienia z sytuacją, w której wyniki szacowanych parametrów odnoszą się do materiału obserwowanego. Wyniki obserwacji nie są wówczas losowane z żadnej populacji i odnoszą się do konkretnego obszaru. W sytuacji takiej, podobnie jak przy wyborze losowym próby, dokonujemy szacowania parametrów i oceniamy błędy średnie szacunku. Trzeba zwrócić uwagę, że przy braku założenia o losowym poborze próby traci sens operowanie pewnymi pojęciami probabilistycznymi, takimi jak wartość oczekiwana, błąd średni szacunku itp.

Powstaje pytanie czy w rozważaniu tym chodzi tylko o kwestie terminologiczne, czy też o jakąś zasadniczo różną interpretację dwóch sytuacji odmiennych: 1) w której występuje możliwość wielokrotnego pobierania próby usprawiedliwiającą częstościową interpretację średniego błędu szacunku; 2) w której szacunki i oceny ich błędów trzeba uznać tylko za charakterystykę zebranego materiału empirycznego. Rozróżnienie to prowadzi do oczywistego wniosku, że w sytuacji drugiej, tak znamiennej dla statystycznego szacowania modeli w badaniach geograficznych, nie wychodzimy poza wąską generalizację historyczną, a więc wyniki konkretyzacji takich modeli nie mają charakteru teoriiotwórczego.

Stosowanie losowego pobierania próby jest niestety w badaniach geograficznych dość rzadkie. Jedną z zasadniczych trudności jaka powstaje jest ta, że przeprowadzenie badania na podstawie próby losowej, a następnie zinterpretowanie tego wyniku jako odnoszącego się do całej zbiorowości, wynika nie tylko ze zbadania zgodności wyników z danymi empirycznymi, lecz także przeprowadzenia testowania na podstawie pomiaru jednostek z innego wielokrotnego wyboru losowego, co pozwoliłoby sprawdzić wartość przyjętej metody reprezentacyjnej. W tym drugim przypadku oczywiście możliwości interpretacyjne są dużo szersze i mogą prowadzić do ustalania zależności nieograniczonych przestrzennie.

Czwarty problem, który chciałbym poruszyć ma nieco inny charakter i odnosi się do stosunku, jaki zachodzi między rozwiązywanym problemem a stosowaną metodą.

W geografii, podobnie jak i w innych dyscyplinach, matematyka jest neutralna w stosunku do problemów o charakterze rzeczowym. Sam fakt konstruowania modeli matematycznych nie przesądza o stanowisku jakie zajmuje badacz wobec podstawowych zagadnień, które chce rozwiązać. Pomimo tego neutralnego charakteru metod matematycznych powstaje jednak zjawisko nazywane niekiedy „technikocentryzmem”, a polegające na przystosowaniu problemów do metod, zamiast postępowania odwrotnego. Jak już to stwierdziłem na początku sytuacja taka znamionuje pierwszy etap stosowania metod matematycznych. Nadmierny „technikocentryzm” mógłby prowadzić do sytuacji, w której rozwiązywano by różne banalne tematy tylko dlatego, że można je opracowywać za pomocą eleganckich i wyrafinowanych metod statystycznych. Można oczywiście zrozumieć, że w pierwszym etapie sytuacja taka jest usprawiedliwiona zafascynowana niem ścisłością matematyczną. Dalszy rozwój geografii wymaga jednak takiego podejścia, w którym o wyborze problemu decydują z jednej strony sens problemowy, tzn. potrzeba rozwiązania istotnego pytania poznawczego i wypełnienia luk w istniejącej wiedzy, a z drugiej sens metodologiczny, polegający na zastosowaniu takiego zabiegu poznawczego, który da w wyniku możliwie ściśle i prawomocne rezultaty. Stosowanie metod matematycznych na równi z innymi zabiegami poznawczymi musi być całkowicie podporządkowane takim kryteriom.

Literatura

- Burton I. (1963) *The quantitative revolution and theoretical geography*. „The Canadian Geographer” 7, 1963, s. 151 - 162.
- Czyż T. (1973) *Zastosowanie metod i modeli matematycznych w geografii polskiej*. „Przegląd Geograficzny” 45, 1, s. 29 - 49.
- Garrison W. (1956) *Applicability of statistical inference to geographical research*. „Geographical Review”, 46, 1956, s. 427 - 429.
- Narrol R. (1966) *Two solutions of Galton's problem*. W: F. Moore (ed.) *Readings in cross-cultural methodology*, New Haven.
- Ossowski S. (1963) *Dwie koncepcje historycznych uogólnień*. „Studia Socjologiczne” 2 (9), s. 51 - 62.
- Pawłowski T. (1969) *Metodologiczne zagadnienia humanistyki*. Warszawa.