

Tadeusz Sozański

ZMIENNE

Notatka do wykładu (10/1999, aktualizacja: 10/2004)

Definicja zmiennej

Zmienna (w statystyce używa się też terminu "cecha") jest to przyporządkowanie wartości liczbowych elementom zbioru przedmiotów wyodrębnionego na użytek badań; przedmioty te nazywamy *jednostkami analizy*. W socjologii w tej roli występują ludzie lub zbiorowości społeczne. Dla zmiennej X : $O \rightarrow V$, gdzie O oznacza zbiór jednostek analizy, a V zbiór wartości, zapis " $X(o)=v$ " czytamy: " v jest wartością zmiennej X dla obiektu o " lub "zmienna X przyjmuje wartość v dla obiektu o ".

Jeśli zmienna przyjmuje tylko jedną wartość nazywamy ją *stałą*. Często zbiór jednostek analizy wyodrębnia się z szerszego zbioru przez wskazanie szczególnej wartości jakiejś zmiennej określonej w tym szerszym zbiorze. Wówczas w węższym zbiorze zmienna ta będzie stałą. Termin "zmienna" zasadniczo rezerwuje się jednak dla przypadku, gdy zmienna przyjmuje co najmniej dwie wartości. Przykładem zmiennej o dwu wartościach jest płeć: O jest zbiorem ludzi, a $V=\{0,1\}$, gdzie 0 oznacza, że dana osoba jest mężczyzną, a 1 – że jest kobietą.

"Pojęcia" a zmienne.

Nieformalny dyskurs socjologiczny obok nazw rozmaitych *obiektów społecznych* (jednostek i całości ponadjednostkowych, np. "uczeń", "klasa szkolna") obfituje w liczne nieostre "pojęcia", mające oznaczać cechy obiektów, zjawiska, procesy, takie jak np. "dochód" i "bezrobocie". W kontekście badań obiekty społeczne stają się jednostkami analizy, natomiast "pojęcia", aby zyskać sens naukowy muszą być "przełożone" na zmienne. Przekład "pojęcia" na zmienną wymaga, po pierwsze, wyodrębnienia zbioru jednostek analizy, po drugie, wskazania zbioru wartości liczbowych, po trzecie, określenia sposobu przyporządkowania wartości jednostkom analizy. Np. "dochód" można przekształcić w zmienną charakteryzującą zbiór "podatników polskich w roku 1998", przyjmując za wartość tej zmiennej dla danego podatnika "roczny dochód w zł wykazany w odpowiedniej rubryce PIT-u". Jeśli "bezrobocie" interesuje nas jako "zjawisko" zlokalizowane w "przestrzeni społecznej", możemy wziąć gminy jako jednostki analizy, a wartość zmiennej dla danej gminy określić jako liczbę bezrobotnych zarejestrowanych w tej gminie w ustalonym momencie czasowym. Wybrawszy jednostkę analizy, trzeba następnie nieformalny sens "pojęcia" przełożyć na konkretne określenie zbioru wartości zmiennej i sposobu ich przypisywania jednostkom analizy.

Możliwe są trzy sytuacje: (1) istnieje jedyny przekład podyktowany przez znaczenie pojęcia; w przypadku płci wiadomo, że są dwie wartości a sposób ich zakodowania nie jest istotny; (2) sens pojęcia jest jasny, ale zbiór wartości można określić na różne sposoby; (3) pojęcie dopuszcza dużą swobodę interpretacyjną; tu zaliczamy pojęcia wielowymiarowe (np. "aktywność kulturalna"), które reprezentuje się przez układ zmiennych.

Podziały zmiennych

Ogólne podziały formalne

1. Według typu pomiaru (zmienna jest traktowana jako funkcja pomiarowa w ramach pewnego systemu pomiarowego) rozróżniamy zmienne: *nominalne*, *porządkowe*, *interwałowe*, *stosunkowe* i

absolutne. Zmienne nominalne i porządkowe nazywamy *jakościowymi*, mierzalne w mocniejszym sensie *ilościowymi*.

2. Według mocy zbioru wartości zmienne dzielą się na *skokowe* (inaczej *dyskretne*) i *ciągłe* (jest to podział popularny w statystyce):. W pierwszym przypadku zbiór wartości jest skończony lub przeliczalny (daje się ponumerować) w drugim przypadku ma postać przedziału w zbiorze liczb rzeczywistych. Zmienne skokowe mogą być ilościowe lub jakościowe, zmienne ciągłe muszą być ilościowe. Zmienne skokowe o skończonej liczbie wartości dzielą się na *dwuwartościowe* (inaczej *dychotomiczne*) i *wielowartościowe*.

3. Według sposobu określenia rozróżnić można zmienne *pierwotne* (inaczej *proste*, *źródłowe*) i *wtórne* (inaczej *skonstruowane*). Wartość zmiennej pierwotnej przypisuje się obiektowi wykonując elementarne czynności badawcze (obserwacja, pomiar za pomocą przyrządu, zakodowanie informacji odczytanej ze źródła pisanego). Ze zmiennych pierwotnych konstruuje się zmienne wtórne za pomocą operacji matematycznych (np. $Y=X_1+X_2$) i funkcji (np. $Y=\ln X$), a także poprzez redukcję zbioru wartości.

4. Zmienne *teoretyczne* charakteryzujące abstrakcyjne obiekty teoretyczne odróżnia się od zmiennych *empirycznych* określonych dla obiektów empirycznych. Zmienne empiryczne mogą być obserwowalne, czyli *jawne*, albo nieobserwowalne, czyli *ukryte*. Jawne zmienne empiryczne nazywa się *operacyjnymi*, gdy wiadomo jakie konkretne czynności należy wykonać, aby przypisać wartość zmiennej danemu obiektowi empirycznemu .

5. Zmienne *lokalne* (historyczne) określone są na zbiorze obiektów zdefiniowanym historycznie (przez wskazanie miejsca i czasu) i w sposób, który nie daje się przenieść na szerszy zbiór obiektów zdefiniowany typologicznie (bez odwoływania się do wyznaczników czasoprzestrzennych). Przykładem jest zmienna trójwartościowa określona jako postawa obywateli polskich wobec obecnego prezydenta (postawa negatywna, neutralna, pozytywna). Zmienne *uniwersalne* charakteryzują klasę obiektów zdefiniowaną typologicznie (np. płeć i wiek są zmiennymi określonymi dla wszystkich ludzi), a sposób określenia jest ten sam w całym zbiorze.

Podział zmiennych socjologicznych Lazarsfelda i Menzela

6. Ze względu na typ charakteryzowanych obiektów społecznych (jednostki i układy ponadjednostkowe) zmienne dzielą się na: *indywidualne* (*jednostkowe*) i *kolektywne* (inaczej, *zbiorowe* lub *grupowe*).

Zmienne indywidualne rozpadają się na dwie kategorie: zmienne *absolutne* (przysługujące jednostkom niezależnie od relacji zachodzących między nimi np. płeć) i *relatywne* (opisujące relacje jednostek z innymi jednostkami lub obiektami wyższego rzędu).

Zmienne relatywne dzielą się na trzy podtypy: zmienne *relacjonalne* (mówi się też *relacyjne*) opisują powiązania jednostki z innymi jednostkami w grupie; zmienne *porównawcze* powstają w ten sposób, że wartość pewnej cechy absolutnej (np. wynik egzaminu wyrażony jako ocena na skali szkolnej) zastępuje się przez miarę opisującą miejsce jednostki w zbiorze w porównaniu z innymi jednostkami (np. numer 1 przypisujemy osobie, która w danej grupie uzyskała najlepszy wynik); zmienne *kontekstowe* określa się przenosząc na jednostkę wartość zmiennej kolektywnej dla układu, do którego ta jednostka należy (np. uczniowi można przypisać średnią ocen w jego klasie).

Zmienne zbiorowe podzielone zostały przez Lazarsfelda i Menzela na trzy typy. Zmienne *analizyczne* powstają poprzez wykonanie operacji (np. sumowanie, uśrednienie) na wartościach jakiejś zmiennej indywidualnej zaobserwowanych dla członków grupy (np. średnia ocen w klasie). Zmienne *strukturalne* to zmienne określone na podstawie informacji dotyczących relacji między członkami

grupy (przykładem może być liczba odwzajemnionych wyborów w małej grupie). Zmienne *globalne* (emergentne) to pozostałe zmienne kolektywne. W tym przypadku do przypisania grupie wartości zmiennej wykorzystuje się inne dane niż te, które odnoszą się do jednostek lub powiązań między jednostkami (przykładem zmiennej globalnej jest dochód grupy, o ile nie został określony jako suma dochodów jej członków, bo wówczas byłby zmienną analityczną).

Miejsce zmiennych w badaniach

7. W kontekście procesu badawczego można wyróżnić cztery kategorie zmiennych: A. Zmienne *badane* (mierzone), w tym *stałe* (użyte do wyodrębnienia zbioru jednostek analizy) i *właściwe*, czyli te, ze względu wartości których badane obiekty różnią się między sobą; B. Zmienne celowo *pominięte* przez badacza, lecz możliwe do badania; C. Zmienne, *znane* badaczowi, lecz niemożliwe do badania ze względu na brak narzędzi pomiarowych; D. Zmienne *nieznane* badaczowi. W badaniach eksperymentalnych wśród zmiennych badanych wyróżnia się zmienne *manipulowane*. Są to zmienne, które badacz sam wytwarza za pomocą stosownych procedur i może zdecydować jaką wartość nadać dowolnemu obiektowi.

8. Badanie może mieć na celu jedynie poznanie z jaką częstością występują w badanym zbiorze wartości zmiennych interesujących badacza, np. jaka jest liczba zwolenników i przeciwników kary śmierci wśród dorosłych obywateli danego kraju. Zwykle jednak celem badania jest poznanie związku między zmiennymi, np. jaki jest związek między poziomem wykształcenia a stosunkiem do kary śmierci. Ze względu na rolę w hipotezach opisujących związki między zmiennymi, zmienne dzieli się na *niezależne* (wyjaśniające) i *zależne* (wyjaśniane). Między zmienną niezależną a zależną można wstawić zmienną *interweniującą* (pośredniczącą). W modelach wielozmiennowych wprowadza się także zmienne *zakłócające*, których działanie tłumaczy odchylenia od założonej w teorii idealnej formy związku między zmiennymi.



<http://www.cyf-kr.edu.pl/~usozans/>

04/10/2004