

Walenty Ostasiewicz

Akademia Ekonomiczna we Wrocławiu

KWANTYFIKACJA CECH SPOŁECZNYCH

1. Wstęp

Wśród nauk społecznych, które dość wcześnie zaczęły stosować metody kwantyfikacji zjawisk stanowiących ich główny przedmiot badań, wyróżniają się psychologia i ekonomia. W wyniku systematycznego stosowania pomiaru, a także metod statystycznych do analizy uzyskanych wyników wyodrębniły się dwie odrębne dziedziny badawcze. Jedna z nich, o nazwie psychometria, już na początku lat trzydziestych zajęła się pomiarem cech psychologicznych, stosując przy tym dobrze rozwiniętą teorię pomiaru, nazywaną często teorią testów. Druga dziedzina – ekonometria – traktowana jako połączenie ekonomii matematycznej, którą stosuje się do formułowania hipotez, z analizą statystyczną, służącą do weryfikacji tych hipotez (por. [11]). Mimo swej nazwy z pomiarem nie miała ona wiele wspólnego. Powszechnie stosowane różnego rodzaju wskaźniki jako forma ilościowego charakteryzowania zjawisk ekonomicznych określona została jako pomiar bez teorii (por. [10]).

2. Wskaźnikowy opis rzeczywistości społecznej

Do opisu sytuacji społecznej, nazywanego też ewidencją społeczną, stosowane są różnego rodzaju wielkości liczbowe nazywane miernikami cząstkowymi i syntetycznymi, wskaźnikami lub indeksami. Większość z nich nie ma należytego uzasadnienia teoretycznego z punktu widzenia jakiegokolwiek teorii pomiaru. Dlatego też trudno jest podać porządną klasyfikację wszystkich proponowanych mierników.

Niektóre z mierników dotyczą relacji jednostki z resztą społeczeństwa, inne zaś charakteryzują struktury organizacyjne. Największą grupę stanowią mierniki zwane wskaźnikami społecznymi stanowiącymi tzw. rachunkowość społeczną.

Wskaźniki te wraz ze wskaźnikami ekonomicznymi „najważniejsze aspekty życia społecznego i działalności każdego kraju” opisywane są za pomocą wskaźników należących do dwóch wielkich grup:

- wskaźniki społeczne,
- wskaźniki ekonomiczne.

Wskaźniki społeczne, stanowiące ewidencję społeczną lub raport społeczny, to liczbowe charakterystyki dotyczące stanu całego społeczeństwa. Stosowanych jest kilkaset takich wskaźników. Wymienić można np. takie jak: umieralność niemowląt, liczba lekarzy na 100 000 mieszkańców, udział w dochodzie narodowym wydatków na obronność, edukację, opiekę zdrowotną, odsetek ludzi ubogich, stopa bezrobocia itp. Do wskaźników tego typu zalicza się też wskaźniki stosowane przez ONZ; najbardziej znanym wskaźnikiem jest wskaźnik znany pod kryptonimem HDI.

Oprócz tych dwóch wielkich grup wskaźników stosowanych jest wiele mierników dotyczących ściśle sprecyzowanych aspektów życia społecznego. Jednym z ważniejszych takich szczególnych aspektów jest status społeczny poszczególnych członków społeczeństwa. Do oceny takiego statusu wykorzystywane są cztery następujące cechy: zawód, wykształcenie, dochód, zamieszkanie. Ponieważ trudno o wiarygodne dane o poziomie dochodów, więc większość konstruowanych wskaźników charakteryzujących status społeczny jednostki uwzględnia tylko dwie pierwsze cechy.

Na przykład popularny indeks Duncana (*socio economic index*) stanowi odpowiednio ważoną sumę wartości liczbowych przypisanych poszczególnym zawodom i wykształceniu. Ponad 500 zawodów podzielono na 7 kategorii, takich jak kadra kierownicza wielkich koncernów, menedżerowie, personel administracyjny, urzędnicy, robotnicy wykwalifikowani, operatorzy maszyn, robotnicy niewykwalifikowani.

Również wykształcenie klasyfikowane jest według 7 kategorii: wyższe, studia licencjackie, nieukończone studia licencjackie, wykształcenie średnie, nieukończone średnie, podstawowe, niepełne podstawowe (por. [12]).

Niektóre inne mierniki, nazywane często skalami, wymieniono niżej (por. [12]):

1. Uczestnictwo społeczne
2. Morale i odczucia społeczne.
3. Postawy społeczne.
4. Struktury organizacyjne społeczeństwa.
5. Prestiż społeczny.
6. Skala osobowości.
7. Zaufanie społeczne.

3. Teoria pomiaru i jej niedostatki

Przez pojęcie pomiaru rozumie się przyporządkowywanie liczb cechom rzeczywistych obiektów lub rzeczywistych zjawisk.

Przyjmijmy, że rozpatrywany jest jakiś fragment rzeczywistości, który symbolicznie oznaczony tu będzie symbolem X .

Przyporządkowywanie liczb obiektom (przedmiotom), które są reprezentowane elementami zbioru X , musi zachowywać wszystkie sensowne związki między elementami tego zbioru. Aby ten podstawowy warunek pomiaru można było sprawdzić, potrzebny jest dokładny, precyzyjny opis tych związków w strukturze empirycznej, którą w odróżnieniu od struktury numerycznej nazywa się strukturą jakościową (por. [14]).

Opis taki, podawany w postaci aksjomatycznej, traktowany jest jako jakościowy opis praw empirycznych. Same aksjomaty zaś, jeśli umożliwiają sensowną reprezentację numeryczną, nazywane są aksjomatami pomiaru. Podstawę takiej reprezentacji stanowi tzw. twierdzenie o reprezentacji. Twierdzenie to orzeka, że jeśli dana struktura relacyjna *spełnia* określone aksjomaty, to istnieje *homomorfizm* relacyjnej struktury empirycznej $(X, \succ, \approx)$ w numeryczną (teoretyczną) strukturę relacyjną $(R, >, =)$. Homomorfizm ten nazywa się *skalą pomiarową* lub krótko *skalą*.

Aby można było dokonać pomiaru, trzeba przede wszystkim wiedzieć, czy jest on możliwy. Otóż odpowiedź na takie pytanie stanowi pierwszy podstawowy problem teorii pomiaru. Oznacza on określenie warunków przynajmniej wystarczających, tzn. gwarantujących istnienie homomorfizmu. „Idealne” rozwiązanie problemu istnienia homomorfizmu oznaczałoby również sformułowanie warunków zarówno koniecznych, jak i dostatecznych.

Drugim podstawowym problemem teorii pomiaru jest problem *jednoznaczności*. Polega on na tym, aby określić wszystkie możliwe homomorfizmy.

S.S. Stevens w 1946 r. w czasopiśmie „Science” opublikował artykuł pt. *On the Theory of Scales of Measurement*, w którym wszystkie skale pomiarowe podzielił na cztery grupy o następujących nazwach (por. [15]):

- 1) skale nominalne,
- 2) skale porządkowe,
- 3) skale interwałowe,
- 4) skale ilorazowe.

Typ skali został zdefiniowany przez zbiór wszystkich dopuszczalnych przekształceń.

Przekształcenie φ jest dopuszczalne dla skal porządkowych wtedy i tylko wtedy, gdy

$$m(x) > m(y) \Rightarrow \varphi(m(x)) > \varphi(m(y)).$$

Czyli dowolne przekształcenie monotoniczne wartości $m(x)$ i $m(y)$ jest dopuszczalnym przekształceniem dla skali porządkowej.

Przekształcenie φ jest dopuszczalne dla skali interwałowej wtedy i tylko wtedy, gdy istnieje stała c taka, że

$$m(x) - m(y) = c[\varphi(m(x)) - \varphi(m(y))].$$

W przypadku skali ilorazowej warunek jest podobny, a mianowicie zamiast odejmowania występuje dzielenie, czyli warunek jest następujący:

$$\frac{m(x)}{m(y)} = c \cdot \frac{\varphi(m(x))}{\varphi(m(y))}.$$

Skala nominalna określona jest przez dowolne wzajemnie jednoznaczne odwzorowanie zbioru liczb rzeczywistych.

Klasyfikacja skal pomiarowych wprowadzona została przez psychologa i przez psychologów została spopularyzowana. Zadziwiające jest więc to, że te cztery typy skal pomiarowych stały się wyjątkowo popularne nawet wśród statystyków. Omawiane są w wielu podręcznikach, nawet w takich przypadkach, gdy nie ma ku temu żadnej przesłanki. Podawane przykłady dotyczą zawsze wielkości fizycznych. Brak jest natomiast przykładów z nauk społecznych. W naukach takich bowiem skale te mają wyjątkowo małe zastosowanie. Do klasyfikacji pomiarów charakteryzujących zjawiska ekonomiczne lepsza jest typologia powszechnie stosowana przez statystyków.

- 1) stopnie (*grades*): dobry, dostateczny itp.,
- 2) rangi (*ranks*): liczby naturalne większe,
- 3) częstości, frakcje (*fraction*): ułamki właściwe lub procenty,
- 4) liczebności, liczności (*counts*): liczby nieujemne całkowite,
- 5) ilości (*amounts*): nieujemne liczby rzeczywiste.

Mimo iż istnieje wiele innych klasyfikacji, to w podręcznikach statystycznych wciąż „króluje” mało przydatna klasyfikacja Stevensa.

Zapytajmy na przykład, w jakiej skali mierzona jest zależność liniowa dwóch cech ekonomicznych. Odpowiedź jest prosta: w żadnej z czterech podanych przez Stevensa. Stosowanie procedur pomiarowych w psychologii i innych naukach społecznych stało się powszechną praktyką już na początku lat trzydziestych. Procedury te różniły się jednak bardzo istotnie od procedur proponowanych w ramach dość dobrze rozwiniętej w tym czasie teorii pomiaru. Powstał więc bardzo ważny problem: czy możliwy jest pomiar (w sensie istniejącej teorii pomiaru) doznań ludzkich, ich postaw, odczuć itp.

W celu rozwiązania problemu S.S. Stevens zaproponował zmianę definicji pomiaru tak, aby obejmowała ona procedury stosowane przez psychologów (on sam był psychologiem).

Pomiarem nazwał on przyporządkowywanie liczb obiektom lub zdarzeniom według określonych reguł. Definicją taką nie byli jednak usatysfakcjonowani przedstawiciele nauk społecznych, takich jak socjologia czy ekonomia. Przede wszystkim dlatego, że podstawowym pojęciem tych nauk jest zjawisko, które nie jest ani obiektem, ani jednorazowym zdarzeniem. Jako przykłady można tu wymienić takie zjawiska, jak alienacja, wyeliminowanie, anomia, korupcja itp. Przedstawicielom nauk społecznych nie odpowiadało też zbyt „prymitywne” rozumienie pomiaru jako procedury niemal mechanicznej, polegającej na przyporządkowywaniu liczb według ściśle ustalonych reguł. Pomiar przecież od dawna, w sensie filozoficznym, traktowany jest jako pomost łączący teorię z praktyką.

4. Statystyczne podejście do skalowania

Pomiarem w naukach społecznych proponuje się nazwać nie przypisanie liczb obiektom ze względu na określone cechy, lecz *proces* łączący teoretyczne pojęcia, zwykle abstrakcyjne, z ich obserwowalnymi indykatorami (wskaźnikami) empirycznymi (por. [4; 6]). Pomiar w naukach społecznych nazywa się często *skalowaniem*. Przy czym odróżnia się skalowanie jednowymiarowe od wielowymiarowego. Metody skalowania jednowymiarowego klasyfikowane są według wielu różnych kryteriów. Jedno z kryteriów wykorzystuje pojęcie funkcji charakterystycznej, funkcji śladowej, funkcji odpowiedzi, funkcji reakcji itp.

W celu precyzyjnego określenia tej funkcji psychometrycznej przyjmijmy, że mierzona cecha oznaczona będzie symbolem Y , jej wartości zaś małą literą y . Załóżmy ponadto, że cecha ta nie jest bezpośrednio obserwowalna. Obserwowalne zaś są cechy $X_1, X_2, \dots, X_n$, które traktujemy tu jako zmienne losowe zero-jedynkowe. Cechę Y można więc potraktować jako pojęcie teoretyczne, cechy $X_1, X_2, \dots, X_n$ zaś jako jej indykatory empiryczne.

Oznaczmy rozkład cech $X_1, X_2, \dots, X_n$ symbolem $f(x) = f(x_1, x_2, \dots, x_n)$, a rozkład cechy Y w populacji symbolem $h(y)$. Związek między cechą nieobserwowalną a jej indykatorami określany jest w postaci rozkładu warunkowego $g(x | y)$. Łączny rozkład cech obserwowalnych można wówczas przedstawić w postaci (por. [2]):

$$f(x) = \int_{DY} h(y)g(x | y)dy,$$

gdzie DY oznacza dziedzinę wartości zmiennej Y .

Zadanie pomiaru cechy Y polega, przy tak sformułowanym problemie, na określeniu jakiegoś parametru pozycyjnego rozkładu warunkowego $h(y | x)$, np. wartości oczekiwanej tego rozkładu.

Rozkład $h(y | x)$ określa się z twierdzenia Bayesa następująco:

$$h(y | x) = \frac{h(y)g(x | y)}{f(x)}.$$

Praktyczne korzystanie z tego wzoru jest możliwe tylko w przypadku przyjęcia dodatkowych ograniczeń. Jednym z nich jest tzw. niezależność lokalna zwana też niezależnością warunkową (por. [2]):

$$g(x | y) = \prod_{i=1}^n g_i(x_i | y).$$

Funkcje g_i nazywa się funkcjami charakterystycznymi indykatorów X_i , $i = 1, 2, \dots, n$. Funkcje te definiowane są w postaci parametrycznej, parametry zaś są estymowane na podstawie obserwacji $x_1, x_2, \dots, x_n$. Jeżeli funkcja charakterystyczna jest monotoniczna, to mówi się o skalowaniu sumacyjnym lub skalowaniu Likerta. Jeżeli zaś funkcje g_i mają kształty funkcji gęstości rozkładu normalnego, to mówi

się wówczas o skalowaniu Thurstona, które określa się też mianem skalowania komparacyjnego. Gdy funkcja charakterystyczna jest funkcją schodkową, to mówi się o skalowaniu Guttmana, które określa się też mianem skalowania kumulacyjnego (por. [16; 21]).

5. Uwagi końcowe

Pojęcie pomiaru jest podstawowym pojęciem tzw. reprezentacyjnej teorii pomiaru, która w pełni rozwinięta została w latach siedemdziesiątych. Pomiarom w tej teorii nazywa się homomorfizm między tzw. relacyjną strukturą empiryczną i teoretyczną strukturą liczb rzeczywistych. Głównym problemem tej teorii jest problem istnienia homomorfizmu i jego jednoznaczności. Warunki dostateczne na jego istnienie nazywa się aksjomatami pomiaru, które mogą być traktowane jako prawa empiryczne opisujące w sposób jakościowy strukturę empiryczną. Jeśli te prawa są spełnione w pewnym fragmencie świata, to fragment ten można sensownie opisać liczbowo. Problem sensowności jest to trzeci podstawowy problem teorii pomiaru. Teoria ta ma zastosowanie głównie do pomiaru cech fizycznych.

Teoria ta jest jednak prawie bezużyteczna w naukach społecznych, szczególnie tych o charakterze behawioralnym. W naukach tych bowiem zachodzi potrzeba pomiaru postaw ludzi, ich zachowań, a także różnych zjawisk społecznych. W większości takich przypadków cecha, którą trzeba zmierzyć, istnieje tylko hipotetycznie w umysłach ludzkich. Pomiar polegać więc może tylko poprzez zadawanie pytań w formie bezpośredniej lub w postaci ankiet.

Odpowiedzi na zadawane pytania kodowane liczbowo określa się zwykle mianem *skalowania*. W teorii skalowania rozróżnia się trzy rodzaje skal: skale sumacyjne, skale komparacyjne i skale kumulacyjne (por. [4; 9]).

W odróżnieniu od reprezentacyjnej teorii pomiaru, podstawowym problemem teorii skalowania jest problem *trafnego* wyboru indykatorów mierzonej wielkości hipotetycznej, a także problem precyzji pomiaru, zwany inaczej problemem *rzetelności* pomiaru.

W teorii skalowania pomiar jest zawsze możliwy, problem polega tylko na tym, jak go trafnie i rzetelnie przeprowadzić i jak te dwa podstawowe kryteria zwyfikować.

Literatura

- [1] Adams E.W., Fagot R., Robinson R.E. (1965), *A Theory of Appropriate Statistics*, „Psychometrika” 30, s. 99-127
- [2] Bartholomew D.J., Knott M. (1999), *Latent Variable Models and Factor Analysis*, Arnold, London.
- [3] Biehler R. (1982), *Explorative Datenanalyse – eine Untersuchung aus der Perspektive einer deskriptiv-empirischen Wissenschaftstheorie*, Universität Bielefeld, Bielefeld.

- [4] Blalock H.M. (red.) (1974), *Measurement in the Social Sciences*, Aldine-Atherton, Chicago.
- [5] Burke C.J. (1953), *Additive Scales and Statistics*, „Psychological Review”, vol. 60, nr 1, 73-75.
- [6] Carmines E.G., Zeller R.A. (1980), *Reliability and Validity Assessment*, Sage Publications, Beverly Hills.
- [7] Coombs C.H. (1964), *Theory of Data*, J. Wiley, New York.
- [8] Duncan O.D. (1984), *Notes on Social Measurement*, Russel Sage Foundation, New York.
- [9] Goldberger A.S. (1971), *Econometrics and Psychometrics: a Survey of Communalities*, „Psychometrika” 36(2), 83-107.
- [10] Koopmans T.C. (1974), *Measurement without Theory*, „The Review of Economics and Statistics” 29, 161-172.
- [11] Landreth H., Colander D.C. (1998), *Historia myśli ekonomicznej*, PWN, Warszawa.
- [12] Miller D.C., Salkind N.J. (2002), *Handbook of Research Design*, Sage Publications, London.
- [13] Mosteller F., Tukey J.W. (1977), *Data Analysis and Regression*, Addison-Wesley, Boston.
- [14] Narens L. (1985), *Abstract Measurement Theory*, MIT Press, Massachusetts.
- [15] Narens L. (1981), *On the Scales of Measurement*, „J. Math. Psychology”, 241, 249-275.
- [16] Ostasiewicz W. (2003), *Istota pomiaru statystycznego*, [w:] *Pomiar statystyczny*, red. W. Ostasiewicz, AE, Wrocław.
- [17] Pfanzagl J. (1971), *Theory of Measurement*, Physica Verlag, Wien.
- [18] Roberts F.S., Rosenbaum Z. (1986), *Scale Type, Meaningfulness and Possible Psychophysical Laws*, „Math. Social Sciences” 12, 77-95.
- [19] Robinson E.E. (1965), *Measurement and Statistics: Towards a Classification of the Theory of „Permissible Statistics”*, „Phil. of Science” 32, 229-243.
- [20] Savage I.R. (1957), *Nonparametric Statistics*, JASA, 52, 331-344.
- [21] Torgerson W.S. (1958), *Theory and Methods of Scaling*, J. Wiley, New York.
- [22] Zimbardo P.G. (1999), *Psychologia i życie*, PWN, Warszawa.

QUANTIFICATION OF SOCIAL FEATURES

Summary

The modern measurement theory which began in 1901 with Hölder, and culminated with monograph by Krantz D.H., Luce R.D., Suppes P. and Tversky A. is of a little use in social and behavioral sciences. In the paper there are discussed some shortcomings of this theory, and indicated some new developments appropriate for scaling in social sciences.

Particularly, the theory proposed by D. Bartholomew seems to have a good perspective for wide applications. This theory is shortly highlighted.