

Tadeusz Miller

Uwagi na temat metody ustalania częstości korzystania z wybranych dóbr i usług w badaniach socjologicznych

Przegląd Socjologiczny Sociological Review 30, 123-130

1978

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

TADEUSZ MILLER
Uniwersytet Łódzki

UWAGI NA TEMAT METODY USTALANIA CZĘSTOŚCI KORZYSTANIA Z WYBRANYCH DÓBR I USŁUG W BADANIACH SOCJOLOGICZNYCH

Częstość korzystania z rozmaitych dóbr i usług jest przedmiotem zainteresowania socjologów empiryków. Niekiedy jest to centralny problem badań lub jego ważna część, niekiedy różnice w częstości korzystania stanowią główne lub pomocnicze kryterium stratyfikacji badanej zbiorowości, innym razem wreszcie szuka się związków między częstością korzystania a innymi cechami demograficznymi, ekonomicznymi, społecznymi, psychicznymi, przy czym częstość korzystania jest czasem zmienną niezależną, a czasem zależną.

Wyniki niektórych badań prowadzonych przez instytucje centralnej administracji, a dotyczące częstości korzystania z wybranych dóbr i usług, publikowane są co roku w „Roczniku Statystycznym” wydawanym przez Główny Urząd Statystyczny. Dla przykładu można tu wymienić kilka.

1. W 1975 r. w miastach, w których istniały czynne trasy komunikacji miejskiej, przeciętnie 1 mieszkaniec korzystał z jazdy autobusem, tramwajem lub trolejbusem 1,29 razy w ciągu doby¹. Rozpatrując częstość korzystania wg województw spostrzec można poważne różnice. Największe przeciętne są: w woj. stołecznym warszawskim — 2,66; w woj. miejskim łódzkim — 2,07; w woj. miejskim krakowskim — 1,97; a najmniejsze: w woj. tarnobrzesckim — 0,32; sieradzki — 0,07 i skierniewickim — 0,01.

2. W 1975 r. mieszkaniec Polski kupił przeciętnie 76 egzemplarzy krajowych gazet i 23 egzemplarze krajowych czasopism². W 1970 r.

¹ „Rocznik Statystyczny”, 1976, tablica 18 (612) s. 416.

² „Rocznik Statystyczny”, 1976, tablica 6 (699), s. 462, oraz tablica 25 (718), s. 471.

przeciętne te wynosiły 70 egzemplarzy gazet i 20 egzemplarzy czasopism. Rozpatrując częstość kupowania wg województw wykrywa się ogromne różnice.

Dla krajowych gazet największe przeciętne występują: w woj. miejskim łódzkim — 122 egz., w woj. bydgoskim — 120 i wrocławskim — 115; a najmniejsze: w woj. ciechanowskim i łomżyńskim — po 35 egz., w woj. ostrołęckim — 28 i w woj. skierniewickim — 25.

Dla krajowych czasopism największe przeciętne występują w woj. szczecińskim — 34 egz., w woj. olsztyńskim i stołecznym warszawskim — po 32 egz., a najmniejsze: w woj. konińskim, ostrołęckim i sieradzkim po 10 egz., w woj. radomskim i skierniewickim — po 9 egz.

3. W 1975 r. wśród czytelników korzystających z bibliotek publicznych, związków zawodowych i rozmaitych instytucji przeciętna liczba wypożyczeń wynosiła 16,1³, a mieszkaniec Polski korzystał z wypożyczeń przeciętnie 4,124 razy w ciągu tego roku⁴.

4. W 1975 r. mieszkaniec Polski był przeciętnie 4,14 razy w kinie, 0,43 razy na przedstawieniach estradowych, 0,13 razy w cyrku⁵.

5. Pracownik ubezpieczony w Zakładzie Ubezpieczeń Społecznych korzystał przeciętnie w 1975 r. przez 14,8 dni ze zwolnienia lekarskiego z pracy z powodu choroby lub wypadku, podczas gdy w 1960 r. przeciętna ta wynosiła 10,6 dni, w 1965 — 11,9 dni, w 1970 — 13,0 dni⁶.

W podanych przykładach ustalenie częstości było jednym z głównych celów badania. Wyniki badania pozwoliły na zanalizowanie dynamiki zjawiska i jego terytorialnego zróżnicowania. Metody uzyskania informacji były we wszystkich tych przykładach podobne: na podstawie obowiązującej przedsiębiorstwa i instytucje państwowe sprawozdawczości otrzymywano dane o liczbie widzów, pasażerów, sprzedanych gazet itd. i dzieląc te liczby przez liczbę ludności Polski, liczbę ubezpieczonych w ZUS itd. otrzymywano przeciętną.

Można chyba doświadczenia płynące z podanych przykładów uogólnić. Wydaje się, że jeżeli badania dotyczą wszystkich mieszkańców danego terytorium, bez dzielenia ich na warstwy, i jeżeli informacje o liczbie tych mieszkańców i łącznej liczbie (krotności) udzielonych usług lub sprzedanych dóbr zawarte są w obowiązującej odpowiedniej instytucji sprawozdawczości, to problem ustalenia przeciętnej nie naręcza żadnych trudności.

³ „Rocznik Statystyczny”, 1976, tablica 8 (701), s. 463.

⁴ „Rocznik Statystyczny”, 1976, tablica 9 (702), s. 464.

⁵ „Rocznik Statystyczny”, 1976, tablica 15 (708), s. 466, i tablica 12 (705), s. 465.

⁶ „Rocznik Statystyczny”, 1976, tablica 22 (744), s. 484.

Zupełnie inaczej przedstawia się sytuacja, gdy dla realizacji celu badania konieczne jest uzyskiwanie informacji o częstości korzystania z określonych dóbr i usług przez poszczególne jednostki badanej zbiorowości bądź ich grupy lub gdy chodzi o korzystanie z takich usług, których instytucje nie mają obowiązku lub nie mogą rejestrować bez specjalnie zaprogramowanych i przeprowadzonych badań. Można tu przykładowo przytoczyć słuchanie audycji radiowych, oglądanie telewizji czy czytanie książek.

Podstawową techniką badawczą, jakiej używa się w sytuacjach, o których mowa wyżej, to zadawanie jednostkom badanej zbiorowości pytań na interesujący temat. Zwykle pytania takie przybierają jedną z dwóch postaci:

- jak często P. na ogół (przeciętnie) korzysta...
- ile razy P. w ciągu... korzystał...

Przy pierwszym sformułowaniu badacz przerzuca na respondenta ogromny trud ustalenia tej „przeciętnej krotności”. Można przyjąć za pewnik, że ogromna większość respondentów nigdy przedtem nie zastanawiała się nad taką wielkością i faktycznie nie jest w stanie bez pomocy badacza na takie pytanie odpowiedzieć. Zdając sobie z tego sprawę, badacz do tak sformułowanego pytania dodaje kafeterię odpowiedzi typu „bardzo często”, „często”, „raczej rzadko”, „bardzo rzadko” itp., albo typu „nigdy”, „raz na rok”, „raz na miesiąc”, „raz na tydzień” itp. W pierwszym przypadku badacz zadowala się subiektywną, zależną od wielu czynników oceną respondenta, co w licznych sytuacjach uniemożliwia porównania i stawia pod znakiem zapytania wartość poznawczą badania. W drugim przypadku próba obiektywizacji też nie przynosi dobrych (w sensie bliskich rzeczywistości) rezultatów. Trud obliczenia takiej przeciętnej wcale nie zostaje z respondenta zdjęty, a ponieważ, jak już wspomniano wyżej, najczęściej jest on pierwszy raz o taką sprawę pytany i pierwszy raz się nad nią zastanawia, wybiera wariant odpowiedzi, który mu się w tej chwili z różnych powodów wydaje najwłaściwszy. Wszelkie podsumowania, obliczenia miar centralnej tendencji, zróżnicownia, a nawet zwykłych odsetków są więc zwykle bardzo mało diagnostyczne.

Przy drugim sformułowaniu zasadniczym problemem, który musi rozstrzygnąć badacz, jest wybór długości czasu, użyty w pytaniu. Oczywiście wybór okresu zależny jest w pierwszym rzędzie od badanej problematyki. Inaczej dobierać będzie badacz długość okresu pytając o częstość chodzenia do kina, inaczej pytając o częstość zmiany samochodu. Wydaje się, że można tu wprowadzić zasadnicze rozróżnienie między pytaniami mającymi na celu ustalenie częstości korzystania z takich dóbr lub usług, przy których można się spodziewać przeciętnej

korzystania rządu co najmniej kilku razy w roku, a jednocześnie do rzadkości należeć będą jednostki korzystające mniej niż jeden raz w roku, a pytaniami, dla których spodziewana przeciętna będzie mniejsza niż kilka razy w roku, a liczba jednostek korzystających mniej niż raz w roku będzie liczna. Ta ostatnia kategoria badań, sporadycznie zresztą spotykana w Polsce, nie nasuwa poważniejszych problemów metodologicznych w zakresie formułowania pytań. Badania częstości korzystania z usług typu pierwszego są u nas natomiast na porządku dziennym, a okres, o który badacz pyta respondenta, wymaga specjalnego omówienia. W tej sytuacji dalsza dyskusja dotyczyć będzie sformułowania pytań i dalszej analizy tylko dla tych „dużych” częstości.

W ustaleniu częstości korzystania z usług i dóbr, z których na ogół korzysta się dość często w ciągu roku, w grę wchodzi może sezonowość. Sądzić można, że w okresie wakacyjno-urlopowym więcej się czyta beletrystyki, częściej się chodzi do kina, natomiast rzadziej ogląda się telewizję. W zimie częściej się chodzi do lekarza (przeziębienie, grypa) niż w lecie itd. Ale taka sezonowość nie zawsze musi być powszechna. U osób przedkładających urlop w zimie nad urlop w lecie sezonowość chodzenia do kina, czytania beletrystyki itp. może i pewnie jest inna niż u pozostałych. Niektóre schorzenia dają mniejsze dolegliwości w zimie niż na wiosnę lub na jesieni, co na pewno odbija się na częstości korzystania z porad lekarskich. Dlatego najbardziej właściwym okresem dla ustalania częstości korzystania z takich dóbr i usług, w których może odgrywać rolę sezonowość (a takich jest chyba przytłaczająca większość), jest pełny rok.

Optymalność rocznego okresu dla określenia częstości korzystania nie oznacza jednak konieczności formułowania pytania typu „Ile razy w ciągu roku korzystał P...”. Takie formułowanie nie tylko nie jest konieczne, ale jest chyba bardzo niepożądane. Stwarza ono podobnie jak sformułowanie: „jak często na ogół korzystał P...”, bardzo wielkie trudności dla respondenta. Chyba do wyjątków należą respondenci, którzy potrafiliby zgodnie z rzeczywistością odpowiedzieć na pytanie typu: „ile razy był P. w 1976 roku w kinie” albo „ile razy był P. w ostatnim roku w kinie”. Podobne trudności wystąpiłyby w pytaniach o ilość porad lekarskich, ilość przeczytanych książek itd. Szczególnie wielkie trudności mogłyby wystąpić u ludzi starszych ze względu na osłabienie pamięci, ale i u ludzi młodych bardziej dokładnych odpowiedzi spodziewać się nie można. Po prostu rok jest okresem za długim, aby można było (oczywiście bez prowadzenia specjalnych zapisów) dokładnie zapamiętać krotność korzystania z takich dóbr i usług, z których na ogół korzysta się dość często.

Przytoczone wyżej i bezsporne chyba argumenty prowadzą do wniosku, że należy pytać o okresy krótsze niż pełny rok. Jak pogodzić to jednak z niebezpieczeństwem, stwarzanym przez sezonowość?

Zanim sformułowana zostanie odpowiedź na to pytanie, należy zwrócić uwagę na dwojaki sposób analizy możliwy do przeprowadzenia po zdobyciu odpowiednich informacji od wszystkich respondentów. Jeden z nich to zaliczenie każdego respondenta do odpowiedniej klasy częstościowej i ustalenie na tej podstawie struktury zbiorowości wg częstości korzystania. Przykładowo w rezultacie takiego typu analizy można by otrzymać szereg rozdzielnicy, w którym zawarte byłyby informacje, że 0 razy korzystało w ciągu roku 10% badanej zbiorowości, 1 raz korzystało 20% itd. Zaliczenie każdego respondenta do jakiejś klasy częstości umożliwiłoby wyznaczanie równań regresji i przewidywanie na ich podstawie wartości innych cech skorelowanych z częstością korzystania w ciągu roku. Możliwe byłoby wyciąganie i innych daleko idących wniosków.

Drugi sposób analizy to ustalenie średnich częstości korzystania w ciągu roku bądź dla całej badanej zbiorowości, bądź dla poszczególnych jej warstw, tworzonych według innych cech niż badana częstość korzystania. Przykładowo można by otrzymać w rezultacie takiego typu analizy tablicę statystyczną, w której zawarte byłyby informacje o średniej częstości korzystania w ciągu roku wśród kobiet w wieku 35—44 lat z wykształceniem średnim itp.

Aby możliwy był pierwszy typ analizy, konieczne jest zadawanie pytania o częstość korzystania w ciągu całego roku. Ale, jak starano się uzasadnić wcześniej, takie sformułowania nie mogą doprowadzić do mniej więcej prawdziwej i dokładnej odpowiedzi. Inaczej ma się sprawa, gdy celem badania nie jest ustalenie struktury zbiorowości według badanej częstości korzystania, natomiast częstość korzystania traktuje się jako cechę zależną, której średnie wartości mogą się zmieniać od warstwy do warstwy. Innymi słowy inaczej jest, gdy badacz może zadowolić się obliczeniem średniej ogólnej częstości korzystania w ciągu roku i średnich warunkowych (warstwowych, cząstkowych).

Niech n oznacza, ile razy korzystał respondent z danej usługi lub dobra w ciągu roku ($n = 0, 1, 2, \dots, 365$).

Niech l oznacza długość okresu w dniach, o który badacz pyta respondenta ($l = 1, 2, \dots, 365$).

Niech k oznacza, ile razy korzystał respondent z danej usługi lub dobra w ciągu l dni ($k = 0, 1, 2, \dots, l$).

Jeżeli przyjąć założenie, że przy n -krotnym korzystaniu z danej usługi lub dobra ma się do czynienia ze zdarzeniami niezależnymi, to

prawdopodobieństwo, że respondent, który korzystał n razy w ciągu roku z danej usługi, skorzystał z niej k razy w ciągu l dni wynosi

$$P_l(k|n) = \frac{C_l^k C_{365-l}^{n-k}}{C_{365}^n},$$

gdzie $C_l^k = \frac{l!}{k!(l-k)!}$, $C_{365-l}^{n-k} = \frac{(365-l)!}{(n-k)!(365-l-n+k)!}$,

$$C_{365}^n = \frac{365!}{n!(365-n)!}$$

gdzie $n! = 1 \times 2 \times 3 \dots (n-2)(n-1) \times n$.

Przyjmując przykładowo, że badacz pyta respondenta o krotkość korzystania w ciągu ubiegłych 14 dni (a więc $l = 14$), to prawdopodobieństwo, że respondent (prawdomówny), który zwykle korzysta 7 razy w ciągu roku (a więc $n = 7$), odpowie, że korzystał w ciągu tych dwóch tygodni 0 razy, 1 raz, 2 razy, 3 razy, 4 razy, 5 razy, 6 razy, 7 razy, wynosi odpowiednio:

$$P_{14}(0|7) = \frac{C_{14}^0 C_{351}^7}{C_{365}^7} = \frac{14!}{0!14!} \cdot \frac{351!}{7!344!} = 0,7587,$$

$$P_{14}(1|7) = \frac{C_{14}^1 C_{351}^6}{C_{365}^7} = 0,2155,$$

$$\left. \begin{aligned} P_{14}(2|7) &= 0,0243, \\ P_{14}(3|7) &= 0,0014, \\ P_{14}(4|7) &= 0,0000 \\ P_{14}(5|7) &= 0,0000 \\ P_{14}(6|7) &= 0,0000 \\ P_{14}(7|7) &= 0,0000 \end{aligned} \right\} \text{łącznie } 0,0001.$$

W tabeli podane zostały przykładowo prawdopodobieństwa obliczone dla sytuacji, gdy badacz ustala częstość korzystania w ciągu ostatnich 14 dni dla $n = 0, 1, 2, \dots, 20$ i $k = 0, 1, 2, 3, 4, 5$.

Interpretacja obliczonych prawdopodobieństw prowadzi do ważnych wniosków. Jeżeli badacz bada dużą zbiorowość, w której są także duże liczebności osób korzystających z pewnej usługi lub dobra w ciągu ubiegłego roku 0, 1, 2 i 3 razy, i zadaje im pytanie o częstość

$$\text{Wybrane prawdopodobieństwa } P_{14}(k/n) = \frac{C_{14}^k C_{351}^{n-k}}{C_{365}^n}$$

n/k	0	1	2	3	4	5
0	1,0000					
1	0,9616	0,0384				
2	0,9247	0,0740	0,0014			
3	0,8890	0,1070	0,0040	0,0000		
4	0,8546	0,1375	0,0077	0,0002	0,0000	
5	0,8215	0,1657	0,0124	0,0004	0,0000	0,0000
6	0,7895	0,1917	0,0180	0,0008	0,0000	0,0000
7	0,7587	0,2155	0,0243	0,0014	0,0000	0,0000
8	0,7291	0,2374	0,0313	0,0022	0,0001	0,0000
9	0,7005	0,2573	0,0389	0,0032	0,0002	0,0000
10	0,6729	0,2755	0,0470	0,0044	0,0002	0,0000
11	0,6464	0,2919	0,0555	0,0058	0,0004	0,0000
12	0,6208	0,3068	0,0643	0,0075	0,0005	0,0000
13	0,5962	0,3201	0,0734	0,0095	0,0008	0,0000
14	0,5725	0,3320	0,0827	0,0117	0,0010	0,0001
15	0,5497	0,3425	0,0922	0,0141	0,0014	0,0001
16	0,5277	0,3518	0,1018	0,0169	0,0018	0,0001
17	0,5065	0,3598	0,1114	0,0198	0,0023	0,0002
18	0,4861	0,3668	0,1210	0,0230	0,0028	0,0002
19	0,4665	0,3727	0,1305	0,0265	0,0035	0,0003
20	0,4476	0,3775	0,1400	0,0302	0,0042	0,0004

korzystania z tej usługi w ciągu ubiegłych 14 dni, to można się spodziewać, że

— wśród tych, którzy w ciągu ubiegłego roku nie korzystali ani razu, 100% odpowie 0 razy;

— wśród tych, którzy w ciągu ubiegłego roku korzystali 1 raz, ok. 96% odpowie 0 razy i 4% odpowie 1 raz;

— wśród tych, którzy w ciągu ubiegłego roku korzystali 2 razy, ok. 92% odpowie 0 razy i ok. 7% odpowie 1 raz i ok. 1% odpowie 2 razy;

— wśród tych, którzy w ciągu ubiegłego roku korzystali 3 razy, ok. 89% odpowie 0 razy, ok. 10% odpowie 1 raz i ok. 1% odpowie 2 razy.

Rozumowanie takie można oczywiście rozciągnąć na inne n i k . Znając więc w badanej zbiorowości rozkład korzystania w ciągu całego roku, bardzo łatwo stosując podobne rozumowanie obliczyć średnią arytmetyczną korzystania w ciągu dowolnie wybranych 14 dni. A odwrotnie? Co prawda znając średnią 14-dniową nie można odtworzyć rozkładu rocznego, można jednak w przybliżeniu ustalić średnią roczną. Średnia 14-dniowa nie jest przecież wartością przypadkową. Wręcz

przeciwnie, jest ona, jak widać, bardzo silnie uzależniona od rozkładu rocznego korzystania. A więc?

Obliczona średnia arytmetyczna częstości korzystania w ciągu ostatnich 14 dni (ogólna bądź dla poszczególnych warstw, o ile tylko warstwy były dostatecznie liczne) pomnożona przez 26 a ogólnie przez 365/l, będzie w przybliżeniu odpowiadać na pytanie, jaka była średnia częstość korzystania w ciągu ubiegłego roku. Przybliżenie będzie tym dokładniejsze, im większa liczba jednostek była badana. Cel badania jest więc osiągnięty.

Należy teraz rozpatrzyć słuszność założenia o niezależności zdarzeń, jakimi są jednorazowe korzystania z pewnej usługi lub dobra i ustosunkować się do zagadnienia niebezpieczeństwa sezonowości przy pytaniu nie o cały rok, a tylko pewną jego część. Sprawy te zresztą dość silnie się łączą.

Wylimitowanie niebezpieczeństwa sezonowości jest tylko możliwe pod warunkiem, że zbieranie informacji (wywiady, ankieta pocztowa itp.) prowadzone będzie przez cały rok. Badaną zbiorowość należy podzielić na części i każdą z nich badać w innym terminie tak, aby żadna pora roku nie była pominięta. W ten sposób pytanie brzmiące: „ile razy w ciągu ostatniego miesiąca (tygodnia, dwóch tygodni) skorzystał P. z...”, będzie dla każdego respondenta (ewentualnie niewielkiej liczby respondentów łącznie) dotyczyć innego okresu, obejmie cały rok i doprowadzi do eliminacji wpływu sezonowości przy obliczaniu wartości średnich. Ogromną wagę należy jednak przyłożyć do właściwego podziału zbiorowości. Podziału tego nie wolno dokonywać według jakichkolwiek wyróżnionych cech (np. płeć, wiek, miejsce zamieszkania, miejsce pracy) — bezwzględnie konieczny jest tu dobór losowy.

Wydaje się, że przy takim trybie zbierania informacji problem traktowania jednorazowego skorzystania z pewnego dobra lub usługi jako zdarzenia niezależnego traci bardzo na ważności. Jeżeli nawet zdarzenia te nie są niezależne (np. zadowolenie z pójścia do kina przyspiesza ponowne pójście, a niezadowolenie je opóźnia), to proponowany sposób sformułowania pytania i zbierania informacji od różnych respondentów w różnym czasie pozwala na szacowanie rocznych średnich bez systematycznego błędu.

Kończąc rozważania na temat sposobu ustalania częstości korzystania z wybranych usług i dóbr warto dodać, że w analogiczny sposób można badać i inne zachowania ludzkie. Należy jednak pamiętać, że jedynym dopuszczalnym sposobem analizy jest wtedy obliczanie średnich ogólnych i grupowych, natomiast niemożliwe jest ustalenie struktury zbiorowości według badanego zachowania.