

Lidia Kuczmierowska

**Badania ilościowe
w ewaluacji
- do czego potrzebna
jest nam statystyka?
cz. II**

Artykuł powstał w ramach projektu „Postaw na jakość – ewaluacja jako narzędzie zarządzania projektami w instytucjach publicznych zajmujących się polityką społeczną”



Projekt współfinansowany przez Unię Europejską
w ramach Programu Środki Przejściowe 2005
Wielosektorowy Projekt Wzmocnienia Zdolności Administracyjnych

Niniejszy dokument został opublikowany dzięki pomocy finansowej Unii Europejskiej. Za treść tego dokumentu odpowiada Stowarzyszenie BORIS, poglądy nim wyrażone nie odzwierciedlają w żadnym razie oficjalnego stanowiska Unii Europejskiej.

Stowarzyszenie Biuro Obsługi Ruchu Inicjatyw Społecznych BORIS
ul. Ogrodowa 50 lok.1, 00-876 Warszawa
tel./fax (22) 620 31 92, (22) 890 94 49,
www.boris.org.pl, e-mail: boris@boris.org.pl



Warszawa 2008

Badania ilościowe w ewaluacji - do czego potrzebna jest nam statystyka - część II

W pierwszym artykule dotyczącym zastosowań statystyki w ewaluacji skupiliśmy się na tym, jak przy pomocy mierników statystycznych opisywać badaną zbiorowość. Umiejętność tę możemy wykorzystywać do porządkowania i analizowania różnych, dających się wyrazić w liczbach informacji na temat działań i uczestników projektu. Omawiając podstawowe zagadnienia związane z opisem statystycznym, pominęliśmy wówczas kwestię dostępności danych. W przypadku małych projektów, gdzie zbadanie wszystkich osób i podmiotów biorących udział w naszym przedsięwzięciu jest proste i niedrogie, nie musimy kłopotać się tym, czy uzyskane informacje są pełne. Skoro nikogo nie pominęliśmy – mamy wszystkie możliwe dane.

Sprawa, niestety, komplikuje się, kiedy realizujemy projekt wielkoskalowy i nie jest możliwe ani organizacyjnie, ani finansowo gromadzenie zbyt wielu danych o wszystkich beneficjentach. Musimy wówczas ograniczyć się do zbadania tylko niektórych jednostek zbiorowości statystycznej (osób, firm, innych instytucji i grup). Mówimy, że prowadzimy badanie na próbie statystycznej. Jest ono co prawda tańsze, mniej czaso- i pracochłonne, jednak obciążone ryzykiem błędu przy uogólnianiu wyników na całą zbiorowość. Skąd będziemy wiedzieć, że to, co uzyskaliśmy badając 50 osób, będzie prawdziwe dla wszystkich 500 uczestników? Z jaką dokładnością możemy twierdzić, że tak jest? Podstawowym warunkiem wyciągania takich uogólniających, odnoszących się do całej zbiorowości wniosków jest losowy dobór próby.

Co to jest losowy dobór próby, co to jest próba losowa?

Z losowymi metodami doboru próby mamy do czynienia, kiedy na wybór jednostek do badania nie ma wpływu ani badacz, ani sam uczestnik – faktycznie są one niejako losowane i dostają się do próby przez przypadek. W ten sposób uzyskujemy próbę losową, czyli taką, której struktura ze względu na interesujące nas cechy statystyczne jest co najwyżej losowo różna od struktury całej zbiorowości. Mówimy inaczej, iż próba losowa jest reprezentatywna. Jeśli np. 1/3 wszystkich uczestników naszego projektu stanowiły osoby zamierzające założyć własne firmy, to w dobrze dobranej próbie losowej ich udział będzie taki sam. W praktyce nasze rozumowanie przebiega, oczywiście, w odwrotnym kierunku: w badanej próbie 1/3 – to osoby planujące założenie własnego biznesu, a ponieważ próba jest losowa, możemy wnioskować, iż wśród wszystkich uczestników mamy 1/3 przyszłych przedsiębiorców. Na razie pozostawiamy na boku kwestię, jak pewni możemy być takich uogólniających wniosków.

Wróćmy na chwilę do sposobów losowania próby. Jak było to już wspomniane wcześniej, aby próbę można było uznać za losową, a więc dającą podstawy do wnioskowania o całej zbiorowości, musi ona być dobrana w sposób losowy. Losowanie próby może przebiegać w różny sposób; statystycy mówią o różnych schematach losowania. Wyróżnia się dwa podstawowe schematy: losowanie niezależne i losowanie zależne. W przypadku losowania niezależnego wszystkie elementy próby losujemy z całej zbiorowości – jest to tzw. losowanie ze zwracaniem. Po wylosowaniu danego elementu z powrotem dołączamy go do całego zbioru i losujemy kolejny raz. Może więc tak się zdarzyć, iż jakaś jednostka wystąpi w próbie losowej więcej niż jeden raz. W przypadku losowania zależnego każde następne losowanie odbywa się „bez zwracania” – losujemy z coraz mniejszej zbiorowości. Każda jednostka

może znaleźć się w próbie tylko raz. Ponieważ ewaluacja nie ma eksperymentalnego charakteru, w jej przypadku stosujemy schematy losowania zależnego.

Kiedy rozstrzygniemy już problem wyboru właściwego schematu losowania, musimy zastanowić się, co będzie stanowić tzw. operat losowania. Operat – to inaczej baza danych służąca jako podstawa do losowania próby; może nią być lista jednostek danej populacji, rejestr, zestawienie - zbiór, z którego można wyselekcjonować próbę badawczą. Dobry operat musi być kompletny i aktualny, ponadto powinien zawierać dodatkowe informacje o elementach populacji, które umożliwią wydzielenie różnych grup (według interesujących nas kryteriów takich jak wiek, płeć, wykształcenie, status zawodowy i inne).

Jak to zrobić w praktyce? Jak dokonać losowego doboru próby?

W naszym przypadku badaną zbiorowość (populację generalną) najczęściej stanowią będą instytucjonalni i indywidualni uczestnicy i beneficjenci projektu. Każdy projektodawca, zwłaszcza realizujący działania dofinansowane ze środków unijnych, prowadzi bazę danych o uczestnikach, w których - głównie na potrzeby sprawozdawczości - gromadzi różne informacje. Taka baza danych, gdzie każdy rekord dotyczący poszczególnych uczestników ma swój numer, może być świetnym operatem losowania. Jak przeprowadzić samo losowanie? Można w ramach zabawy losować „szczęśliwe numerki” spośród wszystkich numerków przypisanych elementom naszej populacji (wszystkim uczestnikom projektu figurującym w bazie danych), można też posłużyć się komputerowym generatorem liczb losowych, który „wylosuje” dla nas potrzebną ilość numerów jednostek populacji z operatu losowania. Prosty generator liczb losowych dostępny jest np. w programie Excel – jest to funkcja matematyczna oznaczona jako LOS.

Oprócz opisanego powyżej prostego doboru jednostek do badanej grupy, możliwe są też inne sposoby stworzenia próby losowej. Do najbardziej popularnych należą: dobór warstwowy, zespołowy, wielostopniowy i systematyczny. Spośród wymienionych, w badaniach społeczno-gospodarczych szczególnie często stosowany jest schemat losowania warstwowego. Zbiorowość generalną dzielimy według przyjętych kryteriów na mniejsze grupy, zwane warstwami, następnie z każdej warstwy niezależnie losujemy jednostki do próby. Losowanie warstwowe uważane jest za efektywniejsze od schematu losowania prostego, bo dostarcza bardziej precyzyjnych informacji. Wymaga jednak znajomości struktury populacji (czyli struktury zbioru wszystkich uczestników projektu). Załóżmy, że nasz projekt realizujemy na terenie powiatu, w skład którego wchodzi 10 gmin. Uczestnikami są kobiety i mężczyźni pochodzący ze wszystkich gmin. Interesuje nas, jak z projektu skorzystali mieszkańcy poszczególnych gmin oraz jak skorzystały z niego kobiety, a jak mężczyźni. Mamy 10 gmin, 2 płcie, w sumie będziemy mieć 20 warstw, z których następnie będziemy losować jednostki do próby. Ze względu na prostotę schematu warto też zwrócić uwagę na losowanie systematyczne. Jest ono zalecane m.in. w przypadkach badania bardzo dużej populacji, kiedy numerowanie wszystkich jednostek populacji może być dość uciążliwym zadaniem. Jeśli np. w naszym projekcie uczestniczyło 1000 osób, a do próby potrzebnych jest nam 100 osób, to do badań wybieramy elementy odległe od siebie o 1000 : 100 = 10 jednostek (dobieramy co 10-tą osobę, począwszy od pierwszego elementu próby, który wyznaczamy w sposób losowy, na „chybił-trafił”).

Próby nielosowe

Ponieważ nie wszystkie populacje są jednakowo dostępne, czasem musimy zdecydować się na dobór nielosowy, dokonany na podstawie z góry ustalonych przesłanek. Powoduje to niereprezentatywność próby i co za tym idzie niemożność określenia błędu wnioskowania. Uogólnianie wyników na całą populację jest w tym przypadku statystycznie nieuprawnione. Dwie najpopularniejsze metody nielosowego doboru próby to dobór celowy i dobór kwotowy.

W przypadku doboru celowego eksperci, na podstawie własnej wiedzy i doświadczenia, ustalają, kto w danym badaniu powinien znaleźć się w próbie. Wskazanie na jednostki wchodzące w skład próby może być oparte na przesłankach merytorycznych lub na chęci współpracy badanych z badającymi. Możemy np. zdecydować, że do próby wejdą uczestnicy projektu, którzy uczestniczyli we wszystkich działaniach – wszystko widzieli, wszystko wiedzą, stąd ich opinia o projekcie będzie najbardziej przekrojowa, oparta na pełnych informacjach.

Dobór kwotowy opiera się na znajomości struktury populacji generalnej według przyjętych cech i narzuceniu tej struktury na skład próby. Stosowanymi kryteriami podziału mogą być: wiek, płeć, wielkość rodziny, dochód, przynależność do grupy społecznej, rodzaj aktywności zawodowej. Liczebność grup (segmentów) w próbie ustala się na podstawie przemnożenia rozkładu procentowego wybranych cech w populacji generalnej, przez ogólną liczebność próby.

Jak duża powinna być próba badawcza?

Niestety, na to pytanie nie ma prostej odpowiedzi, choćby dlatego, że pojęcie dużej – małej próby jest w statystyce dość umowne. Tzw. prawo wielkich liczb mówi, iż błąd popełniany przy szacowaniu wielkości pewnej cechy na podstawie badania próby będzie tym mniejszy, im większa będzie liczebność próby. Jednak duża próba sama z siebie nie gwarantuje dużej dokładności. Klasyczną ilustracją takiego poglądu jest przypadek sondaży przedwyborczych prowadzonych przez wiele lat w USA przez *Literary Digest*, które badaniem obejmowało 10 milionów respondentów. Wszyscy zachwycali się wielką dokładnością badań do czasu, kiedy wynik sondażu okazał się o 30% różny od rzeczywistego wyniku wyborów. Duża próba – to także duże koszty badania.

Niektórzy badacze twierdzą, iż w prostych badaniach z rzetelną kontrolą sposobu doboru jednostek do badanej grupy dobre wyniki uzyskać przy próbie liczącej sobie już od 10 do 20 jednostek. Inni¹ uważają, że w większości badań minimum – to 30-40 jednostek i więcej, zaś w przypadku bardziej złożonych badań w próbie powinniśmy mieć kilkaset jednostek. Klasyczne podejście statystyczne² jako graniczną liczbę podaje 120. Zależnie od tego, czy nasza próba jest większa, czy mniejsza niż 120 jednostek stosować będziemy różne wzory przy wnioskowaniu o dokładności przenoszenia uzyskanych wyników na całą zbiorowość.

W przypadku oceny projektów całe zadanie upraszcza się przez to, że - po pierwsze ewaluacja nie jest badaniem naukowym i choćby z tego tytułu nie musi spełniać bardzo ścisłych

¹ Np. Mark Balnaves i Peter Caputi w swoim podręczniku „Introduction to Quantitative Research Methods”, Sage Publications 2001.

² Prezentowane np. w podręcznikach profesora Luszczewicza – „Statystyka stosowana”, PWE 1996; „Statystyka nie jest trudna. Metody wnioskowania statystycznego”, PWE 2001

rygorów badawczych (nie tworzymy nowych teorii naukowych, ale opisujemy pewien dynamicznie zmieniający się wycinek społeczno-gospodarczej rzeczywistości); po drugie – projektodawca na ogół dobrze zna populację uczestników projektu. Są to osoby, które już na etapie rekrutacji muszą spełnić pewne kryteria i z którymi potem realizator projektu ma wiele różnorodnych kontaktów. Wszystkie te czynniki sprzyjają właściwemu (czyli dość reprezentatywnemu) doborowi próby, co w praktyce ma większe znaczenie dla uprawnionego przenoszenia wniosków z badania próby na całą populację uczestników projektu niż liczebność badanej grupy.

Autorka: Lidia Kuczmierowska