

O znaczeniu tajemnicy statystycznej dla jakości badań ilościowych

Mirosław Szreder

Wprowadzenie

Rosnąca rola informacji we współczesnych społeczeństwach (społeczeństwach informacyjnych), służąca zarówno instytucjom publicznym, jak i przedsiębiorstwom na rynku do podejmowania decyzji sprawia, że wzrasta cena wiarygodnej i dobrej jakościowo informacji. Rzetelna informacja sprzyja bowiem podejmowaniu właściwych decyzji. Błędna, lub słabej jakości informacja daje decydentom fałszywy obraz rzeczywistości, prowadzi do marnotrawienia wysiłku, czasu i pieniędzy. Dzieje się tak zarówno w świecie wielkiej polityki, gdzie błędne informacje służb wywiadowczych mogą być powodem niepotrzebnych wojen i nieszczęść, ale także w skali mniejszej – społeczeństwa, lokalnej społeczności, grupy konsumentów narażonych na błędne decyzje tych, którzy podejmują je na podstawie fałszywych przesłanek lub słabej jakości informacji. W większości sytuacji chodzi tu nie o pojedyncze rejestrowane dane, lecz o zbiory liczb, a więc informacje o charakterze statystycznym. Obecnie właśnie zbiory liczb, informacje o różnych charakterystykach badanych populacji uznaje się za najbardziej precyzyjny opis znacznych obszarów rzeczywistości społecznej i ekonomicznej. Rośnie zarówno znaczenie, jak i cena wiarygodnej informacji statystycznej, bo na przeciwnym biegunie rosną koszty złych decyzji (lub zaniechania decyzji) jako konsekwencji słabej jakości informacji otrzymanych przez decydentów.

1. Rola respondenta w badaniach statystycznych

Po to, aby dane statystyczne przedstawiały prawdziwy i możliwie pełny obraz rzeczywistości, potrzebne są zarówno dobrze przygotowany projekt badania, wraz z odpowiednim narzędziem pomiarowym (kwestionariuszem, ankietą), jak i pozytywny stosunek oraz zaufanie badanych osób (konsumentów, przedsiębiorców) do realizatorów badania. Chęć do współpracy wytypowanych do badania osób (respondentów) jest szczególnie ważna we wszelkiego rodzaju zagadnieniach społecznych i ekonomicznych. Pomiaru w tych badaniach – w przeciwieństwie do dziedzin eksperymentalnych – nie dokonuje sam statystyk, odpowiedzialny za właściwe przygotowanie narzędzia pomiarowego, lecz respondent, dostarczający statystykowi potrzebnych informacji. Statystyk nie liczy pracujących w zakładzie osób, nie szacuje wartości majątku przedsiębiorcy, nie sprawdza kwoty pieniędzy wydanej na zakupy przez konsumenta. Informacje tego rodzaju pochodzą od samego respondenta. Statystyk jest zaś odpowiedzialny za to, aby skłonić respondenta do udzielenia informacji i aby były to informacje rzetelne i precyzyjne, mogące służyć podejmowaniu właściwych decyzji. Od respondenta natomiast zależy, czy zechce on podzielić się z ankierem swoją wiedzą, czy poda prawdziwe informacje, czy też będzie zmierzał do zafałszowania znanego mu fragmentu rzeczywistości, albo w ogóle odmówi udziału w badaniu. Wieloletnie doświadczenia statystyków, socjologów, badaczy opinii publicznej pokazują, że kluczowe znaczenie dla postawy respondenta ma zaufanie do badacza, wewnętrzne przekonanie, że ujawnione w badaniu fakty i opinie pozostaną na zawsze anonimowe. Respondent chce mieć pewność, że żadna osoba i żadna instytucja nie będą miały dostępu do podanych przez niego indywidualnych informacji. Również w opracowaniach z tego typu badań nie może być zagrożenia dla respondenta, iż zostanie on przez kogokolwiek zidentyfikowany. Przekonanie respondenta o tym, że wszystkie podane przez niego informacje są chronione tajemnicą statystyczną, stanowi fundament współpracy

między badającym a badanym. Od tego w największym stopniu zależy jakość zebranych w badaniu informacji. Jako statystycy przekonujemy się wielokrotnie o tym, iż respondenci nie życzą sobie ujawniania jakichkolwiek z podanych faktów lub wyrażonych opinii. Obawiają się o swój wizerunek w środowisku, o zachowania konkurentów na rynku, o ingerencję instytucji publicznych, a czasami też o swoje bezpieczeństwo.

2. Tajemnica statystyczna a jakość badań ilościowych

Każdy incydent złamania tajemnicy statystycznej może poważnie zaszkodzić zaufaniu respondentów do wszelkiego typu badań masowych. Lęk respondenta przed wyrażeniem swoich opinii lub informacji o faktach prowadzi w prosty sposób albo do odmowy udziału w badaniu, albo do podania nierzetelnych lub fałszywych danych. Jedna i druga postawa respondentów pogarsza jakość badań statystycznych i bezpośrednio wpływa na tworzenie nieprawdziwego obrazu rzeczywistości dla odbiorców wyników tych badań. Pogorszenie jakości badań jest w tych okolicznościach tym bardziej dotkliwe, że trudno jest oszacować błąd, jakim wymienione postawy respondentów obciążają ostateczne wyniki badania.

Jako odbiorcy badań opinii publicznej jesteśmy przyzwyczajeni do otrzymywania szacunku błędu losowego (statystycznego), zwykle równego 3 proc., jaki przypisać należy wynikom sondażu. W rzeczywistości błąd ten bywa często najmniejszym składnikiem całkowitego błędu, którym obciążone są wyniki wielu badań ilościowych. Poza tym błędem istnieją 4 inne kategorie błędów. Są to:

- błąd pokrycia jednostek badanej zbiorowości przez operat losowania (ang. *coverage error*), który może być wynikiem wykorzystania w badaniu nieaktualnego lub niekompletnego spisu jednostek populacji;
- błąd spowodowany brakiem odpowiedzi respondentów (ang. *nonresponse error*);
- błąd pomiaru (ang. *measurement error*), związany z zarejestrowaniem nieprawdziwych informacji o badanym respondencie;
- błąd przetwarzania danych zebranych w trakcie badania (*postsurvey processing error*).

Aż dwie spośród wymienionych czterech kategorii błędów są bezpośrednio związane z postawą respondentów. Błąd braku odpowiedzi może wynikać z utraty zaufania respondenta do realizatorów badania, albo z niewystarczającego przekonania o gwarancji anonimowości. Jest to ta kategoria błędów, która mimo coraz bardziej zaawansowanych technik imputacji brakujących danych, wciąż mocno wpływa na ogólną jakość badań. Przy obserwowanym w wielu krajach nasilającym się zmęczeniu społeczeństwa liczbą różnych badań ankietowych, ważne jest, aby zjawiska tego nie pogłębiać, m.in. poprzez brak konsekwencji w zachowaniu tajemnicy statystycznej, a przez to podważanie ogólnego zaufania do badań statystycznych. Raz utracone zaufanie w tej kwestii może być trudne do odzyskania. Dbać o to powinny nie tylko instytucje statystyki publicznej, ale całe środowiska statystyków, socjologów, badaczy rynku i opinii publicznej. Drugą kategorią błędów, których wielkość kształtuje postawa respondentów, jest błąd pomiaru. Jest to błąd, którego źródłem może być źle sformułowane lub źle przez respondenta zrozumiane pytanie w kwestionariuszu, ale także umyślne podanie ankietarowi nieprawdziwej informacji. Motywem tego ostatecznego zachowania bywa obawa respondenta przed ujawnieniem prawdziwych informacji osobom i podmiotom trzecim, czyli niedostatek zaufania do realizatorów badania.

W Polsce, podobnie jak w innych krajach, istnieje duża liczba instytucji i organizacji wykonujących różnego rodzaju badania statystyczne. Wszystkie one powinny dbać o to, aby

wzmacniać i utrzymywać zaufanie społeczne do realizatorów badań, a także pogłębiać przekonanie o przydatności tego typu badań. Poufność indywidualnych danych zebranych w trakcie badań i przestrzeganie zasady anonimowości respondentów stanowią jedne z najważniejszych zasad, jakimi powinny się kierować. Główny Urząd Statystyczny (GUS) oraz urzędy statystyczne są na mocy ustawy zobowiązane do zachowania tajemnicy statystycznej. Pełne przestrzeganie tej zasady buduje zaufanie społeczeństwa do tych instytucji. Łamanie jej, albo choćby pojedyncze odstępstwa od niej, niszczą delikatną więź między badaczem a respondentem, która jest konieczna dla pozyskania od respondentów pełnej i prawdziwej informacji na dany temat.

3. Zainteresowani poszanowaniem zasady poufności w badaniach statystycznych

Kto przede wszystkim powinien zabiegać o pełne respektowanie ustawowo zagwarantowanej tajemnicy statystycznej? Zainteresowani są tym oczywiście statystycy, ale wydaje się, że w takim samym, albo i większym stopniu, dbać o to powinni odbiorcy wyników badań, którym najbardziej zależy na wysokiej ich jakości. Danymi statystyki publicznej posługują się w dużej mierze ministerstwa, inne urzędy centralne, urzędy administracji lokalnej i wiele innych podmiotów, w których podejmowane są najważniejsze dla społeczeństwa lub lokalnej społeczności decyzje. To te organizacje, w przekonaniu o tym, że dobre decyzje można podejmować tylko na podstawie rzetelnych i możliwie pełnych informacji, powinny upominać się o ochronę tajemnicy statystycznej, gdy jest ona zagrożona. Będzie to równoznaczne z upominaniem się o otrzymywanie wysokiej jakości danych statystycznych. Zasada poufności – zgodnie z podstawowymi zasadami statystyki oficjalnej, ogłoszonymi przez Organizację Narodów Zjednoczonych – powinna obowiązywać wszystkie organy statystyczne. Zasada ta głosi: „Dane indywidualne zbierane przez organy statystyczne do opracowań zbiorczych dotyczące zarówno osób fizycznych, jak i prawnych są ściśle poufne i mogą być wykorzystane wyłącznie dla celów statystycznych”.^[1] Wszystkie instytucje w państwie demokratycznym powinny zasadę tę respektować, mając na względzie dobrze pojęty interes społeczny w krótszej i dłuższej perspektywie czasowej.

Sopot, 7 maja 2008r.

prof. dr hab. Mirosław Szreder

^[1] Wyciąg z Podstawowych Zasad Statystyki Oficjalnej uchwalonych przez Komisję Statystyczną Rady Ekonomiczno-Społecznej ONZ w dniach 11-15 kwietnia 1994 r.