

Henryk Banaszak  
Zakład Statystyki, Demografii i Socjologii Matematycznej,  
Instytut Socjologii UW  
[habanasz@is.uw.edu.pl](mailto:habanasz@is.uw.edu.pl)

Wyjaśnienie komunikatu

<https://twitter.com/NaStrazySondazy/status/425331348639735808>

Poniższe wyjaśnienia odnoszą się do notatek zamieszczonych na portalu tygodnika Newsweek:

1	12-12-2013 , ostatnia aktualizacja 12-12-2013 13:56	<a href="http://polska.newsweek.pl/sondaz-wyборы-sejm-newsweek-pl,artykuly,276794,1.html">http://polska.newsweek.pl/sondaz-wyборы-sejm-newsweek-pl,artykuly,276794,1.html</a> <b>Sondaż: Partia Gowina powyżej progu wyborczego</b>
2	14-12-2013 , ostatnia aktualizacja 14-12-2013 09:35	<a href="http://polska.newsweek.pl/bronislaw-komorowski-sondaz-wyборы,artykuly,276928,1.html">http://polska.newsweek.pl/bronislaw-komorowski-sondaz-wyборы,artykuly,276928,1.html</a> <b>Sondaż „Newsweeka”: Komorowski już w pierwszej turze</b>
3	20-01-2014 , ostatnia aktualizacja 20-01-2014 10:23	<a href="http://polska.newsweek.pl/sondaz-newsweeka-platforma-wyprzedza-pis-czytaj-na-newsweek-pl,artykuly,279158,1.html">http://polska.newsweek.pl/sondaz-newsweeka-platforma-wyprzedza-pis-czytaj-na-newsweek-pl,artykuly,279158,1.html</a> <b>Sondaż Newsweeka: Platforma wyprzedza PiS</b>
4	21-01-2014 , ostatnia aktualizacja 21-01-2014 09:45	<a href="http://polska.newsweek.pl/sondaz-prezydencki-newsweeka-,artykuly,279206,1.html">http://polska.newsweek.pl/sondaz-prezydencki-newsweeka-,artykuly,279206,1.html</a> <b>Sondaż prezydencki „Newsweeka”: Komorowski w pierwszej turze, Kaczyński traci</b>

W powyższych publikacjach interpretowane, objaśniane i komentowane są wyniki badań wykonanych dla Newsweeka przez firmę Estymator. Wyniki te traktowane są poważnie, jako prawdziwy, wiarygodny i precyzyjny opis rzeczywistości. Autor notatek (Paweł Szaniawski) twierdzi, że „na PiS zagłosowałoby 31 proc. Polaków” (12-12-2013), „na Bronisława Komorowskiego zagłosowałoby ponad połowa Polaków” (14-12-2013), że „PO wyprzedziła w sondażach PiS” (20-01-2014) a jego rozmówcy, uniwersyteccy politolodzy, twierdzenia te przyjmują jako opis stanu faktycznego a następnie interpretują je i komentują. Objasnia je również dr Jacek Chołoniewski, z wykształcenia astrofizyk.

Niestety, jak wynika z not metodologicznych, związek z rzeczywistością liczb dostarczonych Newsweekowi przez Estymator jest nieznan, stąd wygłaszanie na ich podstawie uogólnień (jak cytowane wyżej) jest operacją wielce ryzykowną. Zastosowana przez firmę procedura wyłaniania osób odpowiadających na pytania ankietera nie gwarantuje bowiem trafnego, wiarygodnego i dokładnego oszacowania aktualnych preferencji wyborczych Polaków.

Aby preferencje te poprawnie oszacować, trzeba poświęcić znacznie więcej czasu i środków niż zużyto na zamówione przez Newsweek badanie. Produkt dostarczony redakcji przez Estymator nie pochodzi bowiem z badania sondażowego, tak jak zbior ankierowanbych osób nie jest – wbrew deklaracjom właściciela firmy – „próbą reprezentatywną” uprawnionych do głosowania Polaków.

Jako uzasadnienie stawianej tu tezy o niskiej wartości poznawczej liczb dostarczanych Newsweekowi przez Estymator przytoczę fakty znane wszystkim zajmującym się zawodowo badaniami sondażowymi i wykładane na wstępnych kursach statystyki czy też metod doboru dla początkujących sondażystów. Służę oczywiście wyjaśnieniami tak redakcji jak i właścicielowi Estymatora.

## Po co się robi sondaż

Celem sondażu jest udzielenie odpowiedzi na pytanie o aktualny stan populacji. W sondażu preferencji wyborczych pytanie to brzmi: jaki jest odsetek zwolenników poszczególnych partii w populacji uprawnionych do głosowania Polaków?

Odpowiedzi na tak postawione pytanie mogą być prawdziwe lub fałszywe, dokładne lub przybliżone, uzyskane dużym lub małym kosztem, a czas ich uzyskania może być krótki lub długi.

Na poważne traktowanie zasługuje tylko jeden ze sposobów odpowiadania na tego typu pytania - szacowanie odsetków na podstawie rozkładów uzyskanych w poprawnie wylosowanej i rzetelnie zrealizowanej próbie pobranej z populacji. Tylko wyniki poprawnie przeprowadzonego badania sondażowego mogą być podstawą odpowiedzi, której związek z rzeczywistością (tu: preferencjami wyborczymi elektoratu) jest znany i dla której jest również znana (bo daje się wyliczyć) tej odpowiedzi precyzja (dokładność oszacowania).

Nie oznacza to bynajmniej, że przepowiednie wróżek, poglądy ekspertów czy też wyniki przedsięwzięć sondażo-podobnych (sond ulicznych, plebiscytów audio-tele, ankiet internetowych itp.) są zawsze nietrafne - różnie to bywa. Sondażowe sposoby odpowiadania na pytania o cechy populacji mają jednak cechy, których wróżby, poglądy ekspertów i wyniki przedsięwzięć sondażo-podobnych nie mają, bo mieć nie mogą.

Poprawnie wykonany sondaż daje oszacowania tak samo trafne i tak samo precyzyjne **zawsze**, bez względu na to, czy dotyczy poglądów politycznych czy też ulubionego proszku do prania, czy pogoda w „miejscach publicznych” sprzyjała konwersjacom z ankierem czy nie i czy wróżka lub ekspert mieli dobry czy zły dzień. I tym sondaż różni się od wróżb i przedsięwzięć sondażo-podobnych. Niestety, nie za darmo.

## Dobry sondaż kosztuje

Poprawnie zaprojektowany i uczciwie zrealizowany sondaż musi kosztować.

Próba, na podstawie której będą szacowane parametry populacji (jak odsetki czy średnie) musi zostać z populacji wylosowana. Kosztuje operat losowania i projekt schematu doboru, łatwy i tani w realizacji badania (mało dojazdów) a zarazem statystycznie efektywny (wystarczająca dokładność oszacowań).

Jeśli precyzja oszacowań ma być wysoka, próba musi być duża i trzeba ją rzetelniej zrealizować w terenie, co kosztuje dodatkowo.

Najwięcej kosztuje dotarcie do osoby wylosowanej. W poprawnie prowadzonym badaniu sondażowym ankieter musi dotrzeć do każdego z wylosowanych i przekonać go do udziału w badaniu. To kosztuje - przeciętny koszt dojazdu do 1 osoby wylosowanej w badaniu na próbie ogólnopolskiej to kwota rzędu 20-40 złotych.

Kosztuje również wynagrodzenia ankietera. Jeśli jest zbyt niskie, jego skuteczność docierania do osób wylosowanych i uzyskiwania od nich odpowiedzi spada, odsetek jednostek wylosowanych lecz niedostępnych wzrasta a dokładność oszacowań drastycznie się pogarsza.

Wbrew temu, co wiele osób sądzi, tak w mediach jak i agencjach badawczych, nie da się zrobić dobrego jakościowo sondażu szybko i za niewielką kwotę. Jeśli ktoś oferuje ogólnopolskie badanie 1200 osób zrealizowane w dwa dni za kilkanaście tysięcy złotych, to oferuje niby-sondaż.

Oferta obiecująca wyniki niby-sondażu może wyglądać atrakcyjnie, gdy tak jak w przypadku Estymatora znajdzie się w niej wzmianka o „*reprezentatywnej*” próbie „*osób dorosłych o liczebności 1200 osób*” (20-01-2014). Szczegóły pojawiające się dalej dowodzą jednak, że Estymator próby wcale nie losował i w związku z tym nie wiadomo, czy zbadane 1200 osób to zbiorowość „reprezentatywna” dla populacji dorosłych Polaków czy też nie. Ponieważ nie była losowana, mogła być reprezentatywna do pewnego stopnia i w pewnym ograniczonym zakresie. Do jakiego stopnia i ze względu na jakie zmienne – nie wiadomo.

## Co to znaczy że próba jest reprezentatywna i jak to sprawdzić

Pojęcie „reprezentatywności” ma w badaniach sondażowych dwa znaczenia, „przedmiotowe” i czysto statystyczne, „proceduralne”.

W pierwszym wypadku reprezentatywność jest cechą konkretnej, pojedynczej, zrealizowanej już próby, w drugim – własnością sposobu jej utworzenia (losowania). Między tymi dwoma znaczeniami zachodzi relacja „statystycznego wynikamia”.

### Reprezentatywność pojedynczej próby

Próba „reprezentatywna” w przedmiotowym sensie, to taka, która ma strukturę zgodną ze strukturą badanej populacji. Zgodność ta jest stopniowalna.

Stopień zgodności pojedynczej próby z populacją, którą ma „reprezentować”, można czasami ustalić lecz jeśli jest to możliwe, to tylko ze względu na pewien skończony zestaw cech. Ta sama próba może zatem być wysoce reprezentatywna ze względu na płeć i wiek badanych a jednocześnie wysoce niereprezentatywna ze względu na poziom ich wykształcenia.

Reprezentatywność jest cechą konkretnej próby. Oznacza to, że próba, o ile nie jest losowana, w jednym badaniu może być dla badanej populacji wystarczająco reprezentatywna a w następnym badaniu już nie.

Stopień reprezentatywności konkretnej próby można sprawdzić tylko ze względu na te cechy, których populacyjne rozkłady są znane. O konkretnej próbie można więc powiedzieć, że ma strukturę zgodną z populacyjną (jest dla niej przedmiotowo reprezentatywna) ze względu na płeć, wiek, miejsce zamieszkania i wykształcenie, gdyż rozkłady populacyjne tych cech uzyskuje się w wyniku Narodowego Spisu Powszechnego. Nie da się natomiast ustalić czy zrealizowana próba ma strukturę zgodną z populacyjną ze względu na te cechy

jednostek, których Narodowy Spis Powszechny nie rejestruje: preferencji wyborczych, poglądów, wartości i tak dalej.

Wobec tego jakie mamy racjonalne podstawy aby twierdzić, że wyniki badania sondażowego dobrze reprezentują stan, w jakim znajduje się populacja? Co pozwala nam traktować je jako trafne, wiarygodne i wystarczająco dokładne oszacowanie tych cech populacji, którymi nie zajmuje się Narodowy Spis Powszechny – na przykład preferencjami wyborczymi? Odpowiedź brzmi – procedura doboru jednostek do badania - losowy dobór próby.

## Reprezentatywność to własność procedury doboru próby

Próba dobrana jest w sposób losowy, gdy prawdopodobieństwo dostania się do próby jest znane (nie musi być dla wszystkich identyczne) dla każdego elementu populacji. Wtedy wiadomo, w jaki sposób na podstawie próby najlepiej oszacować rozkłady cech w populacji oraz wiadomo jaka jest tych oszacowań dokładność.

Losowy dobór próby ma trzy ważne praktyczne konsekwencje:

1. Przy ustalonym schemacie losowania i liczebności próby procedura optymalnego szacowania oraz jego dokładność są takie same dla wszystkich cech jednostek w populacji, zarówno takich jak płeć, wiek czy wykształcenie jak i tych, którymi Narodowe Spisy Powszechne się nie zajmują, jak na przykład preferencje wyborcze. Oznacza to, że każda wylosowana w ten sam sposób próba jest tak w takim samym stopniu reprezentatywna (przedmiotowo) dla populacji ze względu na wszystkie cechy jej jednostek. Jeśli próba jest losowana, to stopień jej reprezentatywności i dokładność oszacowań oznaczają de facto to samo.
2. Z poprawnego schematu losowania wynika funkcyjna zależność dokładności oszacowań od liczebności próby. Wiadomo więc, bo daje się to policzyć, jaką dokładność oszacowań populacyjnych parametrów (na przykład odsetka zamierzających głosować na partię X) daje się uzyskać dla próby o liczebności 500 a jaką dla próby liczącej 1600 osób. Pozwala to zamawiającemu sondaż racjonalnie wybierać między kosztem sondażu a jego dokładnością. Jeśli próba nie ma losowego charakteru, zamawiający nie ma wyboru - może tylko wierzyć agencji badawczej, że dostarczane przez nią wyniki mają jakiś związek ze stanem populacji, bez żadnej jednak gwarancji co stopnia ich zgodności ze stanem faktycznym.
3. Jeśli próby są dobierane w sposób losowy, można na ich podstawie wnioskować o istotności (statystycznej) różnic między odsetkami w populacji. W konsekwencji interpretatorzy wyników mają szansę uniknięcia bezpodstawnych twierdzeń w rodzaju „Partia Gowina powyżej progu wyborczego” (12-12-2013) czy też „PO wyprzedza PiS” (20-01-2014). Gdyby Estymator losował próby, oba twierdzenia byłyby ewidentnie nieuzasadnione, gdyż obie różnice, różnica między notowaniami partii Gowina a poziomem 5% jak i różnica notowań PO i PiS byłyby wtedy statystycznie nieistotne. W obu przypadkach próby losowane nie były a dokładności oszacowań Estymator nie podał, więc pytania o istotność różnic nie warto nawet zadawać.

## Dlaczego zalecamy ostrożność

Omawiane tu „sondaże” Estymatora nie zasługują na podawanie ich wyników opinii publicznej, gdyż związek tych wyników ze stanem faktycznym (preferencjami wyborczymi Polaków) nie daje się ustalić. Jest to skutek karkołomnej procedury rekrutacji respondentów, którą Estymator zastosował zamiast losowego ich doboru. Procedura ta ma charakter nielosowy, w związku z tym rozkłady odpowiedzi otrzymane w tych przedsięwzięciach są wątpliwymi przesłankami wnioskowania o stanie i zmianach w populacji polskich wyborców.

Gdyby „sondaże” te ograniczały się do badania metodą CATI, gdyby numery telefonów były generowane losowo, gdyby rejestrowano odmowy i połączenia nieudane, gdyby kontrolowano przedmiotową reprezentatywność zrealizowanej tą metodą próby, o wiarygodności badania można by dyskutować, a reprezentatywność (przedmiotową) zrealizowanych prób można by próbować wyznaczyć.

Jednak uznawanie osób rekrutowanych w „miejscach publicznych” za reprezentantów populacji dorosłych Polaków a sklejęnej z dwóch części „próby” za „reprezentatywną” jest zabiegiem, który wyniki dostarczone Newsweekowi czyni bezwartościowymi.

Komentując na Tweeterze casus Estymatora moi byli uczniowie z zespołu „Na Straży Sondaży” wykazali się nadmierną tolerancją dla szumu informacyjnego, który właściciel firmy swoimi niby-sondażami wygenerował a następnie redakcji Newsweeka sprzedał.

Proceder uprawiany przez Estymator zasługuje na reakcję bardziej adekwatną do jego charakteru. Oceny nie zmieniają przy tym wyjaśnienia właściciela firmy, który na Facebooku zespołu „Na Straży Sondaży” napisał:

*Jestem szefem Estymatora. Odpowiadam na pytanie o zasadność mieszania metodologii. W CATI procent odmów jest bardzo znaczny co oczywiście bardzo zmniejsza dokładność wyników. W badaniach typu ON-STREET procent odmów jest mniejszy, no i, co ważne, ma także inną „strukturę” - inne osoby odmawiają niż w CATI. Mierzenie tej samej, nieznannej wielkości różnymi metodami ma sens, bo można przypuszczać, że każda z tych metod jest obciążona inaczej i „biorąc średnią” mniej się mylimy. Pozdrawiam wszystkich zainteresowanych sondażami. Jacek Chołoniewski.”*

Argumentacja Jacka Chołoniewskiego jest nietrafna i wręcz absurdalna.

Z faktu iż w badaniu typu ON-STREET niedostępne są inne segmenty populacji niż w badaniu typu CATI nie wynika bynajmniej że dokładność oszacowania na podstawie połączonych „prób” jest lepsza, niż w każdej z nich z osobna.

Po pierwsze, skąd wiadomo i jak to można sprawdzić, że „*„biorąc średnią” mniej się mylimy*”? Skąd wiadomo, że jednostki niedostępne w badaniu typu CATI pojawiają się na ulicy i wtedy stają się dostępne? Skąd wiadomo, że takie zjawisko „kompensacji prób” zachodzi? Osobiste doświadczenie i przekonanie autora nie ma tu nic do rzeczy, gdyż wątpliwości statystyczne wyjaśnić można tylko przy użyciu statystycznych argumentów – danych i wyników obliczeń, dowodu formalnego.

Po drugie, nawet jeśli efekt kompensacji ma miejsce, to można go rozpoznać w odniesieniu do niektórych cech respondentów (na przykład gdy w badaniu CATI częściej odmawiają ludzie młodzi zaś w badaniu ). Nie można natomiast wiedzieć jakie preferencje wyborcze mają osoby niedostępne w obu typach badań. Czy więc można wiedzieć, że szacując rozkład preferencji wyborczych dorosłych Polaków „*„biorąc średnią” mniej się mylimy*”? Nie można, bo skąd.

Po trzecie, pojęcie „odmowy” jako przyczyny niezrealizowania wywiadu dotyczy tylko jednostek wylosowanych do badania. Tylko wtedy bowiem można ustalić (wyliczyć) jak odmowa „*zmniejsza dokładność wyników*”. W badaniu typu ON-STREET osoby zapraszane do udziału w badaniu nie są losowane. Losowania takiego nie da się przeprowadzić, gdyż dla zbiorowości osób znajdujących się w jakimś „miejscu publicznym” nie da się wyznaczyć ich prawdopodobieństw dostania się do próby, nie bardzo wiadomo, ile tych osób w sumie jest. Wobec tego liczenie i używanie odsetka odmów („*mniejszego*” niż dla CATI !) jako drogi do wyznaczenia dokładności oszacowań w badaniu typu ON-STREET jest zabiegiem magicznym a nie statystycznym. Z tego właśnie powodu w metodologii badań sondażowych pojęcie badania typu ON-STREET nie występuje. Z faktu, że do mediów przedostają się wyniki „sondy ulicznej” nie wynika bynajmniej, iż zawierają one treści odnoszące się do stanu populacji w znany sposób. Nie są więc warte publikacji.

Aby wynik badania załugiwał na publikację musi zawierać liczby, których relacja ze stanem faktycznym (na przykład z preferencjami wyborczymi) populacji (na przykład populacji dorosłych Polaków) jest znana.

Nie wystarczy zadeklarować, że liczby te są rezultatem badania próby reprezentatywnej aby można było je traktować jako trafne, wiarygodne i wystarczająco dokładne źródło wniosków na temat badanego społeczeństwa. Trzeba tę reprezentatywność zapewnić losując próbę a dokładność oszacowań wyliczyć. Obu tych warunków nie spełniają niby-sondaże oferowane przez Estymator.

Estymator nie jest niestety jedynym przedstawicielem firm „badawczych”, które sprawiają wrażenie profesjonalnych lecz oferują szum informacyjny pozbawiony wartości. Firmy takie można poznać po tym, że nie są w stanie podać dokładności swoich oszacowań a jeśli je podają, nie chcą pokazać, jak tę dokładność wyliczyły.

Korzystając z pozornie atrakcyjnych ofert tego typu firm media ryzykują utratę swojej wiarygodności. Firmy tego typu mogą bowiem od czasu do czasu dostarczać wyniki zbliżone do stanu faktycznego, mogą też bezkarnie produkować dobrze sprzedawalne sensacje artefakty, gdyż nikt nie jest w stanie definitywnie wykazać zgodności lub sprzeczności ich ustaleń (reprezentatywności przedmiotowej) ze stanem faktycznym.

Od czasu do czasu zdarzają się jednak wybory i wówczas okazuje się kto robi badania sondażowe a kto wykonuje usługi sondażo-podobne. Niestety, fakty wykazują, że Estymator należy do tej drugiej kategorii i pozostaje mi wyrazić ubolewanie, iż redakcja szacownego tygodnika nie zorientowała się w jakości produktu oferowanego przez tę firmę. Bez względu na to ile kosztują Newsweek produkty Estymatora, redakcja z pewnością za nie przepłaciła. Ich wartość informacyjna, jak to dla białego szumu, jest bliska zera. A sondaży tanich i dobrych po prostu nie ma.

Henryk Banaszak  
nauczyciel wielu członków zespołu „Na Straży Sondaży”,  
którzy notatkę na Tweeterze opublikowali