

Może warto bliżej zapoznać się z metodą randomizacji odpowiedzi?

May one learn more about the randomized response method?

Zofia Mielecka-Kubień

Uniwersytet Ekonomiczny w Katowicach

Przedstawiając zasady i przykłady zastosowania różnych modeli metody (czy techniki) randomizacji odpowiedzi często spotykam się z nieprzychylną oceną tej metody ze strony socjologów, pedagogów czy lekarzy (ogólnie *nie-statystyków*), prowadzących badania ankietowe dotyczące używania substancji psychoaktywnych innymi, (nierandomizowanymi) metodami. Postawa ta jest dla mnie po części zrozumiała – posługiwanie się metodą randomizacji odpowiedzi, a nawet docenienie potrzeby i zasadności jej zastosowania, wymaga dobrej znajomości rachunku prawdopodobieństwa i statystyki matematycznej. Nie jest więc przypadkiem, że modele randomizacji odpowiedzi są prezentowane przede wszystkim w czasopismach statystycznych, a nie w czasopismach dotyczących uzależnień czy szerzej – zdrowia. Metody te nie są oprogramowane w popularnych statystycznych pakietach komputerowych, a obliczenia są często niełatwe; dodatkowo zastosowanie randomizacji odpowiedzi w badaniach ankietowych jest trudniejsze, bardziej czasochłonne i bardziej kosztowne niż wiele innych metod ankietowania.

Z drugiej strony, zarówno wyniki eksperymentów statystycznych prowadzonych przeze mnie lub pod moim kierunkiem w Polsce (1, 2, 3), jak i badań w innych krajach, dotyczących wielkości spożycia napojów alkoholowych czy udziału w populacji osób używających narkotyków, otrzymane metodą randomizacji i porównane z wynikami otrzymanymi tradycyjnymi metodami ankietowania, stawiają pod znakiem zapytania wiarygodność tych ostatnich.

Trudno mi się zgodzić ze stwierdzeniami, że *metoda randomizacji odpowiedzi nie znalazła powszechnego zastosowania w badaniach ankietowych na temat używania substancji psychoaktywnych* czy też, że *otwartą kwestią pozostaje, w jakim zakresie pytania o używanie substancji psychoaktywnych rzeczywiście spełniają ten warunek* [tj. są pytaniami kłopotliwymi ZMK] (4). Chociaż metoda randomizacji nie jest powszechnie stosowana w naszym kraju, jest ona dobrze opracowana od strony metodologicznej (5–10), a w literaturze światowej odnotowano już wiele badań tą metodą (np. 11–19).

Oczywiście, można zadawać pytania bardziej kłopotliwe dla respondentów, niż pytania o używanie substancji psychoaktywnych. Oprócz dochodów – wspomnianych przez J. Sierosławskiego (4) – do listy można dołączyć pytania o oszustwa podatkowe, różnego rodzaju przestępstwa, dewiacje seksualne, aborcję itp. W badaniach

ankietowych, poza rozważanymi w moim artykule błędami wynikającymi z odmów odpowiedzi i udzielania odpowiedzi niezgodnych z prawdą, występują zazwyczaj również inne błędy nielosowe. Nie oznacza to przecież, że należy zaniechać wysiłków w ograniczaniu zakresu niektórych z nich, jeżeli taka możliwość istnieje, nawet jeśli są one większe w badaniach innych zagadnień.

Mylny jest pogląd, że metodę randomizacji odpowiedzi można stosować *jedynie w przypadku prostych wskaźników, które możemy otrzymać zadając jedno pytanie*. Można zadać dowolną liczbę pytań (uruchamiając każdorazowo mechanizm losowy), co zresztą pokazano w badaniu, które prowadziłam (2), następnie oszacować np. rozkład spożycia każdego z napojów alkoholowych i parametry tych rozkładów. Można także posłużyć się taką formułą tej metody, gdzie respondent będzie odpowiadał na więcej pytań (np. o częstotliwość spożycia i średnią wielkość spożycia przy jednej okazji) przy jednym uruchomieniu mechanizmu losowego, a więc odpowiedzi będą dotyczyć tej samej osoby (20, 21).

Nie mogę również zgodzić się z opinią, że metoda randomizacji odpowiedzi jest nieprzydatna *do dużych badań populacyjnych oraz tych, których celem jest nie tylko ocena skali zjawiska, ale też analiza jego złożonych uwarunkowań* (22), trudno bowiem analizować uwarunkowania czy skutki np. spożycia alkoholu, dysponując zaniżonymi ocenami poziomu tego spożycia. Metoda randomizacji odpowiedzi (lub inna podobnie efektywna) daje przynajmniej możliwość oceny rzędu wielkości tego niedoszacowania i powinna być, w moim przekonaniu, stosowana jako uzupełniająca właściwe badanie (np. 23, 24).

Przy stosowaniu metody randomizacji odpowiedzi proces wypełniania kwestionariusza nie musi zostać zaburzony, jest to kwestia organizacji badania – badanie można przeprowadzić innego dnia i niekoniecznie metodą *papier-ołówek*, może bowiem być prowadzone telefonicznie, za pośrednictwem poczty lub przy użyciu komputera czy Internetu; opisy takich badań można znaleźć w literaturze przedmiotu (np. 25, 26, 27).

Nie widzę potrzeby porównywania moich skromnych eksperymentów statystycznych, prowadzonych na podstawie prób o minimalnych (w praktyce stosowania randomizacji odpowiedzi) liczebnościach z badaniami prowadzonymi na szeroką skalę w ramach międzynarodowej sieci badawczej HBSC. Celem moich badań było przede wszystkim stwierdzenie, czy i jak metoda randomizacji odpowiedzi funkcjonuje w różnych populacjach. Zastosowanie jej na szerszą skalę wymaga wielu prób podobnych do moich, prowadzonych po to, aby wybrać odpowiedni model tej metody, mechanizm losowy, określić pytania właściwe i neutralne, odpowiednio zorganizować badanie itp.

Niestety, nie podzielam optymizmu co do zaufania respondentów do ankieterów. Jeśli już prowadzimy takie dywagacje to, ze względu na możliwość utraty prestiżu przez respondenta, łatwiej jest – moim zdaniem – otrzymać wiarygodną odpowiedź dotyczącą używania narkotyków od więźniów w zakładach karnych, niż od ogólnie szanowanych menadżerów, nauczycieli czy lekarzy. Zgadza się, że w przypadku dzieci i młodzieży stopień tego zaufania jest przypuszczalnie wyższy, niż w innych populacjach, chociaż w przypadku np. uczniów gimnazjum już niekoniecznie (3).

Nie propaguję zresztą w jakiś szczególny sposób stosowania metody randomizacji odpowiedzi wśród dzieci i młodzieży. Celem badania prowadzonego przeze mnie i współautorkę wśród uczniów gimnazjum było głównie sprawdzenie, czy tak młodzi respondenci potrafią zrozumieć zasadę, w oparciu o którą zbudowano metodę randomizacji odpowiedzi i to, że zapewnia im ona niepodważalną anonimowość, co jest warunkiem powodzenia w stosowaniu tej metody. W moim przekonaniu w obu przypadkach odpowiedź jest twierdząca.

W kwestii zaufania respondentów co do zapewnienia im anonimowości w tradycyjnie przeprowadzanej ankiecie uważam, że możliwość odmowy odpowiedzi nie niweluje prawdopodobieństwa uzyskania nieprawdziwej odpowiedzi: *W związku z tym trudno jest założyć, że ankietowany, mając do wyboru ominięcie danego pytania, będzie świadomie udzielał odpowiedzi sprzecznej z prawdą, a społecznie oczekiwanej* (22). Załóżmy, że pytanie dotyczy używania „twardych” narkotyków. Czy brak odpowiedzi na to pytanie nie będzie skłaniał do podejrzewania respondenta o używanie takich narkotyków? Bo przecież gdyby nie używał, odpowiedź na pytanie nie stanowiłaby dla niego problemu. Trudno oczekiwać, by młodzi ludzie nie dostrzegali możliwości takiej interpretacji braku odpowiedzi na pytania i nie reagowali w odpowiedni sposób.

Podsumowując, uważam metodę randomizacji odpowiedzi za jedną z metod o większej efektywności niż tradycyjne metody ankietowania i widzę możliwości jej stosowania, zwłaszcza w formie badania uzupełniającego, do szacowania niektórych wielkości. Może jednak warto bliżej zapoznać się z tą metodą i próbować zastosować ją w badaniach, z korzyścią dla poprawy ich jakości?

PIŚMIENNICTWO

1. Mielecka-Kubień Z (2004) Zastosowanie metody randomizacji odpowiedzi w badaniach poziomu i jakości życia. W: Ostasiewicz W (red.) *Ocena jakości życia*. Wrocław: Wydawnictwo AE im. Oskara Lange, 214–256
2. Mielecka-Kubień Z (1997) *Ocena rozkładu spożycia napojów alkoholowych za pomocą metody randomizacji odpowiedzi wśród pracowników województwa katowickiego*. Materiały z XXXII Konferencji Statystyków, Ekonometryków i Matematyków Akademii Ekonomicznych Polski Południowej i XIV Seminarium Ekonometrycznego im. Profesora Zbigniewa Pawłowskiego, Osieczany, Akademia Ekonomiczna w Krakowie.
3. Mielecka-Kubień Z, Fugiel A (2010) Zastosowanie metody randomizacji odpowiedzi do badania kontaktu uczniów gimnazjum z narkotykami, papierosami i alkoholem. W: Barczak AS (red.) *Badania ekonometryczne w teorii i praktyce*. Prace naukowe Uniwersytetu Ekonomicznego w Katowicach.
4. Sierosławski J (2012) Czy metoda randomizacji odpowiedzi ogranicza margines błędu w badaniach ankietowych nad używaniem substancji psychoaktywnych? *Alkoholizm i Narkomania*, 25, 1, 109–111.
5. Warner SL (1965) Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 60, 309, 63–69.
6. Fox JA, Tracy PE (1986) *Randomized Response. A Method for Sensitive Surveys*. Beverly Hills, London, New Delhi: SAGE Publications.

7. Chaudhuri A, Mukerjee R (1988) *Randomized Response, Theory and Techniques*. New York and Basel: Marcel Dekker, INC.
8. Duffy JC, Waterton JJ (1984) Randomized Response Models for Estimating the Distribution Function of a Quantitative Character. *International Statistical Review*, 52, 2, 165–171.
9. Bellhouse DR (1995) Estimation of correlation in randomized response. *Survey Methodology*, 21, 13–19.
10. Van der Heijden PGM, van Gils G (1996) *Some logistic regression models for randomized response data*. *Statistical Modeling*. Proceedings of the 11th International Workshop on Statistical Modeling, Orvieto, Italy.
11. Duffy J (1990) Data Collection: A Methodological Response. W: Plant M, Goos C et al. (red.) *From Alcohol and Drugs Research and Policy*. National Criminal Justice 160822.
12. Fisher M, Kupferman LB, Lesser M (1992) Substance use in a school-based clinic population: Use of the randomized response technique to estimate prevalence. *Journal of Adolescent Health*, 13, 281–285.
13. Goodstadt MS, Cook G, Grudson V (1978) The Validity of Reported Drug Use: the Randomized Response Technique. *Substance Use & Misuse*, 13, 359–367.
14. Nordlund S, Holme I, Tamsfoss S (1994) Randomized response estimates for the purchase of smuggled liquor in Norway. *Addiction*, 89 (4), 401–405.
15. Sandler HM, Barth JT (1976) Evaluation of the Randomized Response Technique in a Drinking Survey. *Journal of Studies on Alcohol*, 37, 690–693.
16. Striegel H, Ulrich R, Simon P (2010) Randomized Response Estimates for Doping and Illicit Drug Use in Elite Athletes. *Drug and Alcohol Dependence*, 106, 230–232.
17. Volicer BJ, Cahill MH, Neuberger E, Arnt G (1983) Randomized Response Estimates of Problem Use of Alcohol Among Employed Females. *Alcoholism, Clinical and Experimental Research*, 7, 321–326.
18. Volicer BJ, Volicer L (1982) Randomized Response Technique for Estimating Alcohol Use and Noncompliance in Hypertensives. *Journal of Studies on Alcohol and Drugs*, 43, 7, 739.
19. Weissman AN, Steer RA, Lipton DS (1986) Estimating illicit drug use through telephone interviews and the randomized response technique. *Drug and Alcohol Dependence*, 18, 225–233.
20. Christofides TC (2005) Randomized response technique for two sensitive characteristics at the same time. *Metrika*, 62, 53–63.
21. Tamhane AC (1981) Randomized Response Techniques for Multiple Sensitive Attributes. *Journal of the American Statistical Association*, 76, 916–923.
22. Mazur J (2012) Perspektywa dużych badań populacyjnych młodzieży szkolnej. *Alkoholizm i Narkomania*, 25, 1, 113–116.
23. Arnab R (2004) Optional randomized response techniques for complex survey designs. *Biometrical Journal*, 46, 114–124.
24. Shimizu IM, Bonham GS (1978) Randomized Response Technique in a National Survey. *Journal of the American Statistical Association*, 73, 35–39.
25. Boeijs H, Lensvelt-Mulders G (2002) Honest by chance: A qualitative interview study to clarify respondents' (non-) compliance with computer assisted randomized response. *Bulletin de Methodologie Sociologique*, 75, 1, 24–39.
26. Lensvelt-Mulders G, van der Heijden PGM, Laudy O, van Gils G (2006) A Validation of a Computer-Assisted Randomized Response Survey to Estimate the Prevalence of Fraud in Social Security. *Journal of the Royal Statistical Society, Series A*, 169, 2, 305–318.
27. Stem DE Jr., Steinhurst RK (1984) Telephone Interview and Mail Questionnaire Applications of the Randomized Response Model. *Journal of the American Statistical Association*, 79, 555–564.