

FILIP CHYBALSKI
Katedra Zarządzania
Politechnika Łódzka

O UOGÓLNIANIU WYNIKÓW ANALIZ ILOŚCIOWYCH W NAUKACH O ZARZĄDZANIU

W artykule omówiono problematykę uogólniania wyników analiz ilościowych z próby na populację, szczególną uwagę kierując na wykorzystanie metod wnioskowania statystycznego w naukach o zarządzaniu. Dyskusji poddano uwarunkowania i specyfikę procesu uogólniania wyników, losowe i nielosowe schematy doboru próby oraz różnice pomiędzy wynikami z próby oraz wynikami z populacji. Wyjaśniono źródło ryzyka popełnienia błędu na etapie wnioskowania statystycznego oraz omówiono jego prawdopodobieństwo. Zwrócono uwagę na możliwość uogólniania wyników z próby nielosowej na populację, ale w oparciu o prawdopodobieństwo subiektywne. Omówiono również problematykę uogólniania wyników w przypadku analiz mechanizmów ekonomicznych.

1. Wstęp

Metody statystyczne są szeroko stosowane w naukach ekonomicznych, przy czym wydaje się, że dotychczas znajdowały o wiele większe zastosowanie w takich dyscyplinach, jak ekonomia czy finanse, mniejsze zaś w dyscyplinie nauki o zarządzaniu. Można jednak dostrzec tendencję do coraz powszechniejszego stosowania tych metod również przez naukowców zajmujących się badaniem organizacji i procesów zarządzania. A warto w tym miejscu podkreślić, że nauka o zarządzaniu, jako dyscyplina z dziedziny nauk ekonomicznych, obejmuje inny zakres badawczy aniżeli ekonomia czy finanse. Jak wskazuje bowiem Centralna Komisja ds. Stopni i Tytułów Naukowych w Komunikacie nr 7/2010: „Nauki o zarządzaniu zajmują się powstawaniem, funkcjonowaniem, przekształcaniem, rozwojem oraz współdziałaniem organizacji gospodarczych, przede wszystkim przedsiębiorstw, lecz również instytucji sektora publicznego (administracji rządowej i samorządowej, instytucji naukowych, oświatowo-wychowawczych, służby zdrowia, opieki społecznej, instytucji kulturalnych i innych). Nauki o zarządzaniu obejmują zagadnienia, które mają wpływ na podejmowanie w organizacjach racjonalnych decyzji i sprawne ich funkcjonowanie, przyczynia-

jąc się do osiągnięcia ich celów. Dotyczą one sprawności działań zespołowych i do ich istoty należą problemy pracy kierowniczej” [3, s. 2]. W odniesieniu do aparatu metodycznego, Centralna Komisja oprócz podobieństwa dostrzega również dość istotne różnice pomiędzy analizowanymi trzema dyscyplinami z dziedziny nauk ekonomicznych [3, s. 3]:

- „ekonomia posługuje się w większym stopniu kategoriami syntetycznymi i abstrakcyjnymi, we wnioskowaniu stosuje głównie narzędzia matematyczne, statystyczne i logiczne,
- finanse posługują się w dużym stopniu metodami matematycznymi i statystycznymi oraz narzędziami analizy informacji finansowych w odniesieniu do konkretnych podmiotów i segmentów rynku,
- nauki o zarządzaniu posługują się w większym stopniu opisem konkretnych sytuacji, a we wnioskowaniu stosują głównie metody indukcyjne i metody badań operacyjnych”.

Przywołany Komunikat Centralnej Komisji potwierdza to, co można dostrzec w przeglądzie literatury z obszaru ekonomii, finansów i nauk o zarządzaniu. Mianowicie, w ramach nauki o zarządzaniu w mniejszym stopniu aniżeli w przypadku pozostałych dyscyplin z dziedziny nauk ekonomicznych, w prowadzonych badaniach bazuje się na zjawiskach masowych, na prawie wielkich liczb. Przez to mniejsze są możliwości wykorzystania rachunku prawdopodobieństwa w uogólnianiu wyników z próby na populację.

W niniejszym artykule, mającym przede wszystkim charakter metodyczny, podjęto próbę usystematyzowania wiedzy z zakresu statystyki pod kątem możliwości wykorzystania jej metod do uogólniania wyników analiz ilościowych, odnosząc się przede wszystkim do badań naukowych prowadzonych w dyscyplinie nauk o zarządzaniu. Aczkolwiek problematyka ta jest szersza, i dotyczy również innych dyscyplin ekonomicznych, a także innych dziedzin naukowych, czy obszarów wiedzy. Dyskusji poddano uwarunkowania i specyfikę procesu uogólniania wyników, schematy doboru próby oraz możliwość stosowania narzędzi wnioskowania statystycznego, także w odniesieniu do analiz mechanizmów ekonomicznych. W ostatniej części artykułu sformułowano wnioski oraz podsumowujące rekomendacje, odnoszące się do wykorzystania metod statystycznych do uogólniania wyników prowadzonych badań.

2. Wnioskowanie na podstawie próby losowej

Uogólnianie wyników z próby na populację jest bardzo „kuszące”, ale nie zawsze uzasadnione. Warunkiem koniecznym, by można było wnioski sformułowane na podstawie próby uogólnić z określonym prawdopodobieństwem na populację, z której ta próba została pobrana, jest reprezentatywność tej próby.

Owa reprezentatywność oznacza, że próba dobrze odzwierciedla rozkład badanej cechy w populacji generalnej. Sokołowski próbę nazywa reprezentatywną, jeśli jej struktura jest identyczna lub bardzo zbliżona do struktury populacji. Mechanizm prawa wielkich liczb sprawia, że ta reprezentatywność „zapewni się sama” [16, s. 8]. Zatem, aby zapewnić próbie atrybut reprezentatywności, powinna ona zostać pobrana z zastosowaniem losowego schematu i jednocześnie być dostatecznie liczna. W pierwszej kolejności omówimy schematy doboru próby, a następnie jej wymaganą liczebność.

Schematy doboru próby dzielimy na losowe i nielosowe. Te pierwsze umożliwiają dobór próby w drodze losowania, co oznacza, że przypadek, i tylko przypadek decyduje o tym, że dana jednostka populacji wejdzie lub nie wejdzie do próby [6, s. 107]. Szreder tę definicję próby losowej uzupełnia o równość szans wszystkich jednostek populacji co do możliwości wylosowania – każda jednostka ma jednakowe prawdopodobieństwo trafienia do próby [17, s. 168]. Wszystkie schematy doboru próby, które nie mają charakteru losowego, nazywamy schematami nielosowego doboru próby.

Wśród losowych schematów doboru próby należy wymienić [18, s. 46-51]:

- schemat losowania warstwowego, który warto stosować do pobierania próby z populacji charakteryzującej się silnym zróżnicowaniem badanej cechy. Stosując ten schemat, badaną populację dzieli się na warstwy, z których dokonuje się następnie oddzielnie losowania. Przy czym wydzielone warstwy powinny być możliwie jednorodnie wewnętrznie i jak najbardziej różnić się między sobą;
- dobór wielostopniowy, który jest stosowany w sytuacji, gdy możliwie jest podzielenie populacji na kolejne, coraz mniej liczne grupy. W pierwszej kolejności wybierane są jednostki losowania pierwszego stopnia, po czym jednostki te dzieli się na mniejsze, dochodząc na końcu do podstawowych jednostek badania;
- schemat doboru systematycznego, w przypadku którego z uporządkowanego zbioru jednostek populacji dokonuje się doboru odpowiedniej liczby jednostek w równych odstępach, czyli „co któreś”, aż do uzyskania wymaganej liczebności próby;
- schemat doboru z wykorzystaniem tablic liczb losowych, w przypadku którego odczytywane są kolejne liczby zamieszczone w tych tablicach. Pomija się jednak te spośród nich, które są większe od liczebności populacji.

Próba dobrana za pomocą wyżej opisanych schematów ma charakter losowy. Jeśli dodatkowo jest dostatecznie liczna, umożliwia stosowanie metod wnioskowania statystycznego w postaci estymacji przedziałowej lub testów statystycznych. Tym samym umożliwia uogólnianie wniosków, sformułowanych na podstawie wyników z takiej próby, na populację generalną.

Oczywiście, jak najbardziej zasadne w tym miejscu jest pytanie o liczebność próby, która wraz z losowością nadaje jej atrybut reprezentatywności. Przy czym

nie ma ściśle określonej w sposób arbitralny i uniwersalny liczebności próby, która umożliwi bezpieczne uogólnianie wniosków. Bowiem, po pierwsze to uogólnianie z próby na populację zawsze jest obarczone ryzykiem popełnienia błędu. To ryzyko oczywiście maleje wraz ze wzrostem liczebności próby, czemu towarzyszy w sposób naturalny upodabnianie się próby do populacji pod względem rozkładu badanej cechy. Znika ono dopiero wtedy, gdy próba przestaje być próbą, a staje się populacją. Ale wówczas badanie nie jest już częściowe, lecz całkowite. Po drugie, wymagana minimalna liczebność próby zależy od estymowanego parametru. Dla przykładu, w przypadku wskaźnika struktury (częstości) nie można szacować przedziału ufności, gdy próba liczy poniżej 100 elementów [12, s. 252]. Dla parametrów położenia (wartość przeciętna) czy zmienności (odchylenie standardowe, wariancja) możliwe jest oszacowanie przedziału ufności dla prób liczących mniej niż 100 elementów. Po trzecie, liczebność próby zależy od zadanego z góry prawdopodobieństwa (zwanego poziomem ufności i oznaczanego przez $p = 1 - \alpha$), z jakim oszacowany przedział ufności ma pokrywać nieznaną wartość estymowanego parametru w populacji generalnej. Po czwarte, ważny jest założony dopuszczalny maksymalny błąd szacunku, który jest równy połowie długości przedziału ufności. I po piąte, liczebność próby zależy od schematu losowania. Ma znaczenie, czy próba jest pobrana według schematu losowania ze zwracaniem, czy według schematu losowania bez zwracania [zob. 11, s. 34-36]. Im mniejszy poziom wiarygodności i im większy dopuszczalny maksymalny błąd szacunku, tym minimalna liczebność próby jest mniejsza. I odwrotnie. Szeroki przegląd formuł wyznaczania minimalnej wymaganej liczebności próby można znaleźć w [11; 13].

Losowość próby i prawo wielkich liczb, objawiające się w odpowiedniej liczebności próby, pozwalają z dużym prawdopodobieństwem odzwierciedlić w tej próbie rozkład badanej cechy w populacji, z której ta próba została pobrana. Przy czym dobór losowy sprawia, że niedoskonałości aktu losowania objawiającego się błędem losowania można przypisać prawdopodobieństwo, które oznacza się przez α [17, s. 170]. I właśnie możliwość odniesienia się do tego prawdopodobieństwa warunkuje wykorzystanie w odniesieniu do próby losowej metod wnioskowania statystycznego w postaci estymacji przedziałowej oraz testów statystycznych w celu uogólnienia wyników na populację generalną. Witkowska możliwość określenia błędu uogólniania wyników wskazuje jako podstawową zaletę wnioskowania statystycznego [19, s. 13]. W przypadku estymacji przedziałowej jesteśmy w stanie zbudować wokół uzyskanej w próbie oceny (szacunku) estymatora określonego parametru statystycznego (wartości przeciętnej, wariancji, częstości) przedział, który z zadanym z góry prawdopodobieństwem zawiera nieznaną wartość tego parametru w populacji. Przy czym prawdopodobieństwo to wynosi $(1 - \alpha)$, co oznacza, że w przypadku stukrotnego losowania próby i szacowania na jej podstawie przedziału ufności, w $(1 - \alpha) \times 100$ przypad-

kach przedział ten zawierałby nieznaną wartość parametru w populacji generalnej. Natomiast w $\alpha \times 100$ przypadkach przedział ten nie pokrywałby nieznaną wartość estymowanego parametru w populacji. α jest zatem miarą ryzyka, że szacując przedział ufności na podstawie próby losowej nie „trafimy” w nieznaną rzeczywistą wartość tego parametru w populacji. Przy czym jedynym źródłem tego ryzyka jest „zła” próba, to znaczy taka, która nie reprezentuje dobrze rozkładu badanej cechy w populacji. Wówczas próba nie odzwierciedla dobrze rzeczywistości populacji [1]. Natomiast w przypadku testów statystycznych, weryfikacji poddajemy hipotezę zerową, co do której stwierdzimy brak podstaw do odrzucenia albo ją odrzucimy na rzecz hipotezy alternatywnej. Przy czym poziom istotności α oznacza prawdopodobieństwo, że popełniony zostanie błąd polegający na odrzuceniu hipotezy zerowej w sytuacji, gdy jest ona prawdziwa. Źródłem tego błędu, podobnie jak w przypadku estymacji przedziałowej, będzie „zła” próba, to znaczy taka, która nie reprezentuje dobrze populacji generalnej. Zawsze bowiem istnieje ryzyko wylosowania takiej niereprezentatywnej próby i wynosi ono właśnie α . Prawdopodobieństwo to przekłada się na ryzyko oszacowania nietrafnego przedziału ufności lub ryzyko podjęcia niewłaściwej decyzji odnośnie hipotezy zerowej, ponieważ granice tego przedziału, jak i wartość statystyki testowej przy weryfikacji określonej hipotezy, zależą wyłącznie od wyników z próby.

3. Wnioskowanie na podstawie próby nielosowej

Błąd losowania występuje tylko tam, gdzie jest dobór losowy (losowanie), natomiast tam, gdzie nie ma tego doboru, czyli w przypadku doboru nielosowego, ten błąd nie występuje [17, s. 170]. W konsekwencji, nie można w odniesieniu do prób uzyskanych za pomocą schematów nielosowych posługiwać się prawdopodobieństwem błędu losowania, czyli wartością α . Ma to zasadnicze konsekwencje dla możliwości stosowania w odniesieniu do tak uzyskanej próby metod wnioskowania statystycznego. Mianowicie, tę możliwość eliminuje. Skoro nie jest znane prawdopodobieństwo pobrania „złej” próby, nie jest możliwe określenie ryzyka niepokrycia przez przedział ufności nieznaną wartość estymowanego parametru w populacji generalnej, jak i ryzyko odrzucenia hipotezy zerowej w sytuacji, gdy jest ona prawdziwa. Powyższe wynika zresztą w sposób bezpośredni zarówno z definicji estymacji, jak i z definicji testu statystycznego. Estymacją nazywa się bowiem szacowanie nieznaną wartość parametrów lub nieznaną rozkładów zmiennej losowej w populacji na podstawie rozkładu empirycznego (czyli zaobserwowanego) w próbie losowej, z tej populacji pobranej [12, s. 233]. Warunek losowości próby wskazują również Białek i Depta, definiując estymację przedziałową [2, s. 154]. Testem statystycznym jest natomiast reguła decyzyjna, czy metoda postępowania, która na podstawie wyni-

ków z próby losowej pozwala podjąć decyzję o przyjęciu (stwierdzeniu braku podstaw do odrzucenia, w przypadku testów istotności) lub odrzuceniu hipotezy zerowej [zob. 12, s. 256; 8, s. 101]. Bez założenia o losowości próby nie ma wnioskowania statystycznego.

Na co zatem pozwalają nielosowe schematy doboru próby? A wśród nich można wymienić np. [18, s. 52-53]:

- schemat doboru jednostek typowych, w przypadku którego do próby są dobierane jednostki uważane przez badacza za typowe dla badanej populacji;
- schemat doboru proporcjonalnego, zwany też doбором kwotowym, w przypadku którego wybiera się liczbowo określone segmenty próby odpowiadające proporcjom strukturze populacji, z której próba jest pobierana;
- schemat doboru na zasadzie eliminacji, w przypadku którego eliminacji podlegają jednostki nietypowe dla populacji.

Generalnie, badanie przeprowadzone na podstawie tego typu prób uniemożliwia stosowanie estymacji przedziałowej czy też testów statystycznych. Tym samym nie ma możliwości wyników uzyskanych na podstawie tak pobranej próby uogólnić na populację generalną z określonym ryzykiem popełnienia błędu, wyrażonego prawdopodobieństwem α , ponieważ to prawdopodobieństwo nie jest znane. To oczywiście nie wyklucza możliwości ostrożnego, pozbawionego odwołania do rachunku prawdopodobieństwa, przenoszenia wniosków z próby na populację generalną, jednak w oparciu o przesłanki wskazujące na to, że próba, mimo że dobrana w sposób nielosowy, ma określone cechy reprezentatywności populacji. Dotychczas jednak cały czas odnosiliśmy się do prawdopodobieństwa obiektywnego, czyli opartego na definicji aksjomatycznej lub częstościowej. Jednak badacz może dysponować wiedzą na temat populacji, którą to wiedzę wykorzystuje w sposób uzasadniony przy doborze próby z tejże populacji. Wówczas uznaje za zasadne ingerowanie w proces wyboru jednostek populacji, które podda badaniu, na podstawie którego uzyskane wyniki będzie chciał w jakiś sposób uogólnić na populację generalną. Z wiadomych nam już powodów nie uczyni tego za pomocą metod wnioskowania statystycznego. Jednak może się oprzeć na tzw. prawdopodobieństwie subiektywnym, które wyraża jego przekonanie czy też stopień pewności co do prawdziwości określonego sądu (tezy). Ta wyższość prawdopodobieństwa subiektywnego nad obiektywnym może być przez badacza uzasadniona celem ostatecznym prowadzonych badań, którym nie jest przecież doskonałość próbkowania, lecz możliwie dobre poznanie badanej populacji [17, s. 172]. Warto nadmienić, że prawdopodobieństwo subiektywne znajduje też zastosowanie w budowaniu prognoz i procesie podejmowania decyzji [5, s. 174-177; 7; 9; 10].

Warto również wspomnieć w tej części o pewnych technikach pozyskiwania danych do badań, które mogą sprawiać wrażenie spełniających kryterium losowości, a tym samym umożliwiającymi stosowanie narzędzi wnioskowania staty-

stycznego, podczas gdy faktycznie tego kryterium nie spełniają. Technikami takimi, coraz częściej stosowanymi w naukach o zarządzaniu, są m.in. CATI (wspomagany komputerowo wywiad telefoniczny) oraz CAWI (wspomagany komputerowo wywiad przy pomocy strony www). Dla przykładu, przyjmijmy że przeprowadzono tymi obiema technikami wywiady na wystarczająco licznej próbie konsumentów z zamiarem ich uogólnienia na populację. Tymczasem z powodu zastosowanej techniki próba ta nie miała charakteru reprezentatywnego względem populacji konsumentów. Wynika to z tego, że nie każdy konsument ma telefon i nie każdy konsument ma dostęp do internetu. Tym samym każdy konsument, który nie ma telefonu lub dostępu do internetu z góry był skazany na niepobranie do próby. Jego szanse pod tym względem były zerowe. Chyba, że badaczowi chodziło wyłącznie o osoby posiadające telefon lub posiadające dostęp do internetu i to one stanowiły populację. Z drugiej strony, łagodząc nieco wymagane założenia statystyczne na rzecz użyteczności przeprowadzonego badania, w sytuacji, gdy szacowany odsetek konsumentów w populacji (np. w oparciu o prawdopodobieństwo subiektywne, co nie znaczy dobrze nieuzasadnione) jest znikomy i pominięcie takich konsumentów w badaniu istotnie nie wpłynie na jego wyniki, można podjąć uzasadnioną próbę uogólnienia wniosków na populację, ale bez stosowania metod wnioskowania statystycznego.

4. Wyniki z próby a wyniki z populacji

Badanie całkowite tym różni się od badania częściowego, że w przypadku tego pierwszego badana jest każda jednostka populacji, a w przypadku drugiego tylko wybrane jednostki populacji. Jeśli te jednostki zostały pobrane w sposób losowy i jest ich dostatecznie wiele (próba jest reprezentatywna), można stosować metody wnioskowania statystycznego, co już szeroko omówiono w pkt 2. A co można zrobić w przypadku badania całkowitego? Można obliczyć wybrane, interesujące nas, parametry statystyczne dla badanych cech, zbadać ich rozkłady empiryczne i dokonać interpretacji tak uzyskanych wyników. Następnie na tej podstawie można sformułować wnioski odnoszące się np. do struktury badanej populacji czy zależności pomiędzy zachodzącymi w niej zjawiskami. Trzeba bowiem w tym miejscu podkreślić, że w przypadku badania całkowitego dotyczącego określonych cech populacji generalnej w danym momencie/okresie czasu, nie ma sensu i potrzeby stosowania metod wnioskowania statystycznego. Wynika to z oczywistego faktu, jakim jest znajomość całej populacji, a dokładnie rozkładu empirycznego badanej cechy w populacji w danym momencie/okresie czasu, na podstawie czego można obliczyć wartość przeciętną, wariancję czy wskaźniki struktury dla całej populacji. Zakładając, że dane o badanej cesze zebrano w sposób rzetelny oraz, że nie popełniono błędów w obliczeniach staty-

stycznych, uzyskane wyniki np. w postaci wartości przeciętnej, wariancji czy też częstości, charakteryzują się 100-procentową wiarygodnością.

Podobnie będzie w przypadku współczynnika korelacji liniowej Pearsona, obliczonego dla dwóch wybranych cech, zmierzonych w populacji, a nie w próbach. Wówczas, chcąc ocenić siłę związku korelacyjnego pomiędzy tymi cechami w danym momencie/okresie czasu (bez uogólniania na szerszą perspektywę czasową), należy się oprzeć wyłącznie na uzyskanej wartości współczynnika korelacji i dokonać jej interpretacji. Jak wskazują Ostasiewicz, Rusnak i Siedlecka, wartość bezwzględna współczynnika korelacji mniejsza niż 0,2 oznacza praktycznie brak związku liniowego, z przedziału 0,2-0,4 wyraźną, ale niską zależność, 0,4-0,7 zależność umiarkowaną, 0,7-0,9 zależność znaczącą, a powyżej 0,9 zależność bardzo silną [12, s. 311]. Nie można natomiast zastosować w tym przypadku testu istotności dla współczynnika korelacji. Test bowiem nie służy temu, by ocenić siłę związku, lecz tylko jej istotność statystyczną. Różnica jest zasadnicza. Jeśli przykładowo dla dwóch cech, zmierzonych dla wszystkich jednostek w populacji generalnej, uzyskano współczynnik korelacji liniowej Pearsona równy 0,5, to wiadomo jest, że ten współczynnik dla badanych cech w tej populacji faktycznie tyle wynosi. Zależność jest zatem umiarkowana. Co więcej, wiadomo to na pewno, że 100-procentową wiarygodnością. Nie ma potrzeby potwierdzania tego za pomocą testu statystycznego, tym bardziej, że wówczas zawsze istnieje większe od 0 (równe α) prawdopodobieństwo popełnienia błędu podczas wnioskowania. Test statystyczny służy bowiem udzieleniu odpowiedzi na pytanie, czy określona wartość współczynnika korelacji uzyskana w próbie (np. równa 0,5), przy uwzględnieniu wielkości tej próby oraz założonego prawdopodobieństwa α popełnienia błędu odrzucenia prawdziwej hipotezy zerowej, może być uznana za istotnie różną od zera w populacji generalnej. Podobnie nie ma sensu stosować testu istotności dla dwóch wartości przeciętnych – dla przeciętnego wynagrodzenia w zakładzie A i w zakładzie B w danym momencie/okresie czasu (np. na koniec roku, albo średnie wynagrodzenie za dany rok). Jeśli na podstawie badania całkowitego (dane w postaci wynagrodzeń wszystkich pracowników obu zakładów) uzyskano wartości przeciętne równe odpowiednio 3100 PLN i 3110 PLN, to ze 100-procentową wiarygodnością można stwierdzić, że w zakładzie B przeciętne wynagrodzenie jest wyższe niż w zakładzie A. Można ewentualnie dodać, że różnica w tych wartościach przeciętnych nie jest znaczna. Gdyby natomiast te wyniki pochodziły z próby losowej, wówczas wnioski nie byłyby tak oczywiste w odniesieniu do populacji generalnej. Mianowicie, przy odpowiedniej wartości odchyłeń standardowych wynagrodzeń w obu populacjach (lub w obu próbach) lub przy odpowiednio małej liczebności prób, mogłoby się okazać, że dla zadanego poziomu istotności α brak jest podstaw do odrzucenia hipotezy zerowej mówiącej o równości wartości przeciętnej tych wynagrodzeń w obu populacjach. Dlaczego? Bo to są „tylko”

wyniki z próby losowej, co oznacza, że dla jednej próby test statystyczny mógłby wskazać na odrzucenie hipotezy zerowej, a dla innej próby na brak podstaw do jej odrzucenia.

Można tutaj, na zasadzie analogii, zastosować porównanie do sondażu wyborczego oraz wyniku wyborów. W dniu wyborów, po zamknięciu lokali wyborczych, poznajemy wyniki *exit poll*, czyli uzyskane na podstawie sondażu przeprowadzonego wśród losowo wybranych wyborców, opuszczających lokal wyborczy pod oddaniem głosu. W tym przypadku istnieje ryzyko błędu losowania, skutkującego obciążeniem próby. To ryzyko przekłada się na poziom wiarygodności, z jakim wyniki tego badania można uogólnić na populację wszystkich głosujących. Tu jest miejsce na testy statystyczne i weryfikowanie przykładowej hipotezy, że na partię A głosował większy odsetek wyborców niż na partię B. Gdyby przyjąć, że partia A uzyskała 4% głosów w badaniu *exit poll*, a partia B 5%, to biorąc pod uwagę błąd statystyczny (np. na poziomie 3%) nie można tak jednoznacznie stwierdzić, że partia B wygrała z partią A w wyborach. Dlatego właśnie należałoby zweryfikować hipotezę zerową o równości odsetka oddanych głosów na obie partie wobec hipotezy alternatywnej, że odsetek głosów oddanych na partię A jest niższy niż odsetek głosów oddanych na partię B. Jeśli teraz przyjąć, że we wtorek Państwowa Komisja Wyborcza ogłasza wyniki ze wszystkich lokali wyborczych i partia A uzyskuje 4,75% głosów a partia B 4,65%, to ze 100-procentową wiarygodnością wiadomo jest, że w populacji generalnej (wszyscy głosujący obywatele) partia A uzyskała wyższe poparcie niż partia B. I nie ma tutaj już miejsca na testy statystyczne. A przeprowadzone badanie *exit poll* nie ma już żadnego znaczenia (chyba, że ma służyć poprawie jakości doboru próby).

W naukach o zarządzaniu badania często są prowadzone na jednej lub na kilku organizacjach. Pojawia się zatem pytanie, czy w takich sytuacjach mogą znaleźć zastosowanie jakiegokolwiek metody wnioskowania statystycznego, czyli probabilistycznego uogólniania wyników z próby na populację. Oczywiście wszystko zależy od tego, jakie zjawisko badamy i z czego na co chcemy uogólniać. Co jest badaną populacją, a co jest próbą. Naturalnie, w przypadku badania przeprowadzonego w jednym lub kilku przedsiębiorstwach nie będzie mowy o jakimkolwiek uogólnianiu wyników na populację przedsiębiorstw. To nie oznacza, że jeśli zbadamy np. 1000 przedsiębiorstw z województwa łódzkiego, czyli całkiem liczną grupę podmiotów, wyniki będziemy mogli uogólniać z wykorzystaniem narzędzi wnioskowania statystycznego. Jeśli wspomniane 1000 przedsiębiorstw stanowi próbę reprezentatywną ze względu na badaną cechę (np. wielkość zatrudnienia) dla województwa łódzkiego, wówczas jak najbardziej można wyniki uogólnić, ale tylko na to województwo, a nie np. na cały kraj. Są to rzeczy oczywiste, ale jednocześnie bardzo ważne. Nie można uznać przeprowadzonego badania za obejmujące przedsiębiorstwa zlokalizowane w całej

Polsce, jeśli próba, bez względu na liczebność, została pobrana w jednym województwie. Chyba, że w oparciu o przywołane wcześniej prawdopodobieństwo subiektywne badacz jest w sposób wiarygodny, merytoryczny i przekonujący uzasadnić tezę, że to wybrane województwo jest pod względem badanej cechy „Polską w soczewce”. Tak jak w latach 1993-2009 uznawano za taką „Polskę w pigułce” Wrześnię, organizując w niej prawyборы.

Czy istnieje w ogóle możliwość stosowania narzędzi wnioskowania statystycznego, gdy badanie jest przeprowadzone metodą studium przypadku i obejmuje jedno przedsiębiorstwo produkcyjne? Jeśli badanym zjawiskiem jest jakość wyrobów w tym przedsiębiorstwie i jest ona analizowana na podstawie próby losowej uzyskanej np. w drodze losowania systematycznego (co 20. wyrób schodzący z linii produkcyjnej jest poddawany statystycznej kontroli jakości), to wówczas jak najbardziej mamy do czynienia z próbą losową. Próba ta została pobrana z populacji, którą są wszystkie wytworzone w danym okresie wyroby. Zatem wyniki z tak pobranej i dostatecznie licznej (czyli łącznie spełniającej kryterium reprezentatywności) próby będzie można uogólnić metodami probabilistycznymi na całą populację, np. poprzez oszacowanie przedziału ufności dla średniej czy odchylenia standardowego (wagi wyrobu, jego rozmiarów), czy też dla odsetka wyrobów niespełniających nałożonych norm jakościowych. Można powiedzieć, że wówczas zastosowaną metodą badawczą będzie studium przypadku (tego przedsiębiorstwa) z „zagnieżdżoną” metodą statystyczną w postaci metod wnioskowania statystycznego. Jednak z oczywistych względów sformułowanych w ten sposób wniosków nie będzie można uogólnić na szerszą populację przedsiębiorstw, wytwarzających np. tego samego rodzaju wyroby. Masowość zjawiska i prawo wielkich liczb będą się bowiem odnosić do zjawiska (procesu) zachodzącego w tym przedsiębiorstwie, a nie np. na całym rynku czy na danej przestrzeni geograficznej.

5. Uogólnianie wyników w przypadku analiz mechanizmów ekonomicznych

Na koniec rozważań podejmiemy jeszcze jeden problem, wydaje się, że najtrudniejszy. Mianowicie, można znaleźć w empirycznych badaniach ekonomicznych testy statystyczne przeprowadzane lub modele ekonometryczne szacowane dla całej populacji lub dla próby nielosowej, przy czym w tym drugim przypadku wyniki nie są uogólniane na szerszą populację, lecz odnoszone nadal do tej grupy. Dla przykładu, oszacowano model ekonometryczny dla rocznego wyniku finansowego stu celowo wybranych przedsiębiorstw na podstawie szeregu przekrojowo-czasowego z lat 2010-2015. Albo na podstawie rocznych danych w postaci szeregu przekrojowo-czasowego dla wszystkich krajów UE za lata

2005-2015 oszacowano model ekonometryczny stopy wzrostu PKB. Jaki sens w takim przypadku ma stosowanie metod wnioskowania statystycznego (bo przecież model ekonometryczny to estymacja, a jakość tego modelu ocenia się za pomocą testów statystycznych)? Jak wskazuje Sokołowski, w takich sytuacjach bywa, że negowana jest relacja pomiędzy populacją a próbą. Jednak wówczas celem badania empirycznego jest mechanizm ekonomiczny [16, s. 8]. W takich sytuacjach trudno jest faktycznie ocenić, co jest próbą (czy w ogóle występuje próba), a co jest populacją. Przy czym jako przedmiot badania należy traktować w takim przypadku przede wszystkim analizowany mechanizm ekonomiczny, a nie populację, w której on działa. Wydaje się, że w odniesieniu do działania określonego mechanizmu ekonomicznego odpowiednikiem populacji jest ogólna zasada działania tego mechanizmu w całym „możliwym” okresie, w jakim on działa, a przynajmniej w dłuższym okresie, niż ten który badamy. Próba jest zaś właśnie ten wycinek czasu (moment/okres), na podstawie którego to badanie jest przeprowadzone. W takich sytuacjach uogólnianie wyników nie dotyczy relacji próba (100 przedsiębiorstw, kraje UE) – populacja (wszystkie przedsiębiorstwa zlokalizowane na danej przestrzeni, z której pobrano 100-elementową próbę, wszystkie kraje europejskie, czy nawet szerzej, na globie). Uogólnianie dotyczy raczej relacji postaci: działanie mechanizmu w badanym momencie/okresie czasu – ogólny *modus operandi* określonego mechanizmu ekonomicznego. Należy jednak zachować zdrowy rozsądek przy tym uogólnianiu, tzn. dane za lata 1990-2010 służące za odpowiednik próby nie mogą służyć uogólnieniu wyników np. na cały XX i XXI wiek, ale na okres nieco szerszy aniżeli ten, z którego te dane pochodziły.

Zazwyczaj chodzi nam oczywiście o uogólnienie na moment/okres przyszły, po to bowiem szacujemy modele ekonometryczne oraz chcemy poznawać sposób działania określonych mechanizmów ekonomicznych, by poznawać przyszłość, a nie modelować przeszłość. Modelowanie przeszłości nie jest celem samym w sobie. Służy ono natomiast, przy spełnieniu określonych przesłanek dotyczących przekonania o działaniu zidentyfikowanych mechanizmów również w przyszłości, prognozowaniu. Takie podejście jest zgodne z hipotezą adaptacyjnych oczekiwań, według której prognoza zmiennej opiera się tylko na podstawie historycznych danych o tej zmiennej [zob. 15, s. 202] i odpowiada pasywnej postawie wobec przyszłości [zob. 4, s. 43]. Wówczas model ekonometryczny jest szacowany na podstawie jakiejś „próbki” z przeszłości po to, by uzyskane wyniki w postaci ocen parametrów modelu, opisujących zależności pomiędzy zmienną objaśnianą a zmiennymi objaśniającymi, uogólnić na zdefiniowaną w sposób *explicite* lub *implicite* przyszłość. Jeżeli dodatkowo statystyk uwzględnia w modelu informację spoza danych historycznych i dokonuje korekty ocen parametrów lub koryguje samą prognozę uwzględniając zmiany jakościowe w prognozowanym zjawisku i kształtującym go mechanizmie ekonomicznym, odnosi się

wówczas również do hipotezy racjonalnych oczekiwań. Zgodnie z nią, prognozy są formułowane w oparciu o najlepszą dostępną informację, a nie tylko informację historyczną [zob. 14, s. 510]. Takie podejście odpowiada postawie aktywnej wobec przeszłości [zob. 4, s. 43-44].

W przypadku definiowania okresu lub momentu prognozy w sposób *explicite*, ten moment lub okres są ściśle określone, natomiast w przypadku definiowania w sposób *implicite*, przeszłość, jakiej dotyczy prognoza, pozostaje w domyśle [zob. 4, s. 19]. I analogicznie, wydaje się, że o ile w przypadku badania, którego celem jest poznanie określonej populacji, zarówno ta populacja, jak i próba, jak również relacja między nimi, są zdefiniowane w sposób *explicite*. O tyle w przypadku badań, których celem jest poznanie mechanizmu ekonomicznego, zarówno odpowiednik próby, jak i populacji, są zdefiniowane w sposób *implicite*.

6. Podsumowanie

Badania naukowe w dziedzinie nauk ekonomicznych, w tym także w dyscyplinie nauki o zarządzaniu, coraz częściej wymagają zastosowania metod statystycznych. Bez względu na to, jaki stopień zaawansowania lub trudności towarzyszy tym metodom, kwestią kluczową pozostaje umiejętność ich dobrania, zastosowania i interpretacji uzyskanych rezultatów. Trzeba bowiem zdać sobie sprawę z tego, że współcześnie, kiedy do dyspozycji mamy mniej lub bardziej zaawansowane oprogramowanie statystyczne, nawet to, zawarte w pakiecie MS Excel (funkcje, narzędzia analizy danych), obliczenia oparte na skomplikowanych wzorach nie są już wyzwaniem. Wyzwaniem jest odpowiedni dobór metody, świadomość jej ograniczeń i umiejętność wnioskowania na podstawie uzyskanych wyników, przy zachowaniu odpowiedniej ostrożności. Dobór metod statystycznych każe pamiętać, że statystyka dzieli się na opisową i matematyczną (wnioskowanie statystyczne). Statystyka opisowa pozwala badać populację i formułować wnioski w odniesieniu do tej populacji, ale na podstawie badania całkowitego. Pozwala ona również zbadać próbę, ale nie umożliwia uogólnienia wyników z tej próby na szerszą populację z wykorzystaniem metod probabilistycznych. Brak możliwości uogólniania wniosków z próby na populację czyni tę próbę faktycznie jakąś „wąską” populacją. Natomiast statystyka matematyczna umożliwia uogólnianie wniosków z dostatecznie licznej próby losowej na populację, z której ta próba została pobrana. Pozwala również na poznawanie działania mechanizmów ekonomicznych poprzez badanie określonego wycinka rzeczywistości (działania mechanizmu w określonym momencie lub okresie czasu) i uogólnienie zauważonych prawidłowości na szerszą, określoną raczej w sposób *implicite* niż *explicite*, perspektywę czasową, zazwyczaj na zasadzie kontynuowania w przyszłości prawidłowości zidentyfikowanych w przeszłości.

Literatura

- [1] **Banerjee A., Chitnis U.B., Jadhav S.L., Bhawalkar J.S., Chaudhury S.:** *Hypothesis testing, type I and type II errors*, Industrial Psychiatry Journal, 18(2), 2009, pp. 127-131.
- [2] **Bialek J., Depta A.:** *Statystyka dla studentów z programem STAT_STUD 1.0.*, Wydawnictwo C.H. Beck, Warszawa 2010.
- [3] **Centralna Komisja ds. Stopni i Tytułów Naukowych:** Komunikat nr 7/2010 [file:///C:/Users/user/Downloads/sd_komunikat_CK.pdf, dostęp 15.11.2016].
- [4] **Cieślak M. (red.):** *Prognozowanie gospodarcze. Metody i zastosowanie*, Wydawnictwo Naukowe PWN, Warszawa 2001.
- [5] **Dittmann P.:** *Prognozowanie w przedsiębiorstwie*, Oficyna Ekonomiczna, Kraków 2003.
- [6] **Domański Cz. (red.):** *Metody statystyczne. Teoria i zadania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2001.
- [7] **Engelberg J., Manski Ch. F., Williams J.:** *Comparing the point predictions and subjective probability distributions of professional forecasters*, NBER Working Paper No. 11978, 2006.
- [8] **Gajek L., Kaluszka M.:** *Wnioskowanie statystyczne. Modele i metody*, Wydawnictwo Naukowo-Techniczne, Warszawa 2000.
- [9] **Hampton J.M., Moore P.G., Thomas H.:** *Subjective probability and its measurement*, Journal of the Royal Statistical Society, Series A (General), Vol. 136, No. 1, 1973, pp. 21-42.
- [10] **Lahiri K., Wang J.G.:** *Subjective probability forecasts for recession*, Business Economics, April 2006, pp. 1-12.
- [11] **Miszczak W.:** *Projektowanie próby*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2004.
- [12] **Ostasiewicz O., Rusnak Z., Siedlecka U.:** *Statystyka. Elementy teorii i zadania*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2003.
- [13] **Rószkiewicz M., Perek-Białas J., Węziak-Białowolska D., Zięba-Pietrzak A.:** *Projektowanie badań społeczno-ekonomicznych, Rekomendacje i praktyka badawcza*, Wydawnictwo Naukowe PWN, Warszawa 2013.
- [14] **Samuelson P.A., Nordhaus W.D.:** *Ekonomia*, t. 1, Wydawnictwo Naukowe PWN, Warszawa 2003.
- [15] **Snowdon B., Vane H., Wynarczyk P.:** *Współczesne nurty teorii makroekonomii*, Wydawnictwo Naukowe PWN, Warszawa 1998.
- [16] **Sokołowski A.:** *O niewłaściwym stosowaniu metod statystycznych*, Statsoft Polska, 2004.
- [17] **Szreder M.:** *Losowe i nielosowe próby w badaniach statystycznych*, Przegląd statystyczny, RLVII, Zeszyt 4, 2010, ss. 168-174.
- [18] **Witkowska D. (red.):** *Statystyka w zarządzaniu*, AND, Łódź 2004.
- [19] **Witkowska D.:** *Podstawy ekonometrii i teorii prognozowania*, Oficyna a Wolters Kluwer business, Warszawa 2012.

SOME REMARKS ABOUT INFERENCE OF RESULTS OF QUANTITATIVE ANALYSES IN MANAGEMENT SCIENCE

Summary

The paper addresses a very important topic of the inference from a sample to population, with a special attention paid to the application of statistical methods of inference in management science. The conditions and specifics of inference, random and non-random sampling techniques, the difference between results obtained from the population or from the sample as well as inference in the case of the studies on economic mechanisms are discussed. The source of an error corresponding to statistical inference as well as its probability are explained. The author addresses also the problem of the inference from a non-random sample to population with the application of subjective probability.