

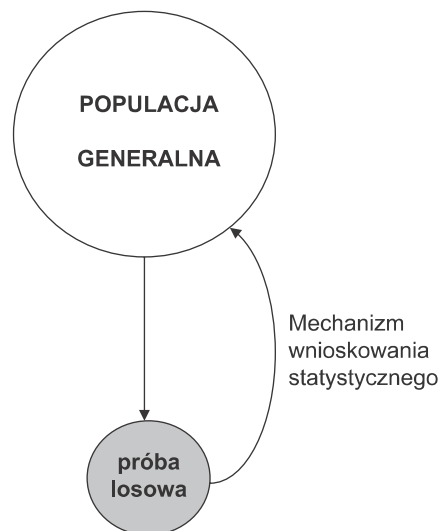
# Wprowadzenie

Statystyka jest nauką o metodach wykrywania i badania prawidłowości występujących w zjawiskach masowych. W ekonomii, zarządzaniu, socjologii, psychologii, medycynie i wszędzie tam, gdzie operuje się danymi (nie tylko ilościowymi, ale również jakościowymi) wykorzystywane są metody statystyczne. Pozwalają one zidentyfikować zaobserwowane prawidłowości oraz odpowiednio je wyrazić. Określenie prawidłowości może sprowadzać się do rozpoznania *rozkładu* interesującej nas cechy w badanej zbiorowości (np. rozkładu płac w populacji), znalezienia *parametrów* tego rozkładu (np. średniej płacy) czy też zbadania *współzależności* między cechami (np. związku wysokości płacy ze stażem pracy, poziomem wykształcenia).

Statystyka, jako dyscyplina naukowa, dzieli się na dwa podstawowe działy: statystykę opisową i matematyczną (inaczej indukcyjną). **Statystyka opisowa** zajmuje się zagadnieniami związanymi z gromadzeniem, opracowaniem, prezentacją i analizą danych. Metody statystyki opisowej prowadzą do syntetycznego opisu badanej zbiorowości i wykorzystywane są wtedy, gdy obserwacja statystyczna obejmuje całą badaną zbiorowość. **Statystyka matematyczna (indukcyjna)** zajmuje się badaniem zbiorowości (populacji generalnej) na podstawie analizy jej części (próby). Metody statystyki matematycznej stosowane są wówczas, gdy nie jest możliwe zebranie informacji o wszystkich elementach populacji interesującej badacza i należy wypowiadać się o badanym zjawisku na podstawie danych częściowych. Statystyka matematyczna jest więc nauką o wnioskowaniu. **Wnioskowanie** przebiega „od próby do populacji”, co należy rozumieć w ten sposób, że na podstawie prawidłowości zaobserwowanych w próbie wypowiadamy się na temat prawidłowości populacji (rys. 1).

Należy w tym miejscu zaznaczyć, iż jedynie próba pobrana w sposób **losowy** upoważnia nas do stosowania metod statystyki matematycznej i uogólniania wyników pochodzących z próby na całą populację generalną. **Losowy dobór próby** (losowanie) oznacza, że jednostki trafiające do próby zostały pobrane za pomocą pewnego mechanizmu losowego. Przykładem takiego mechanizmu jest generator liczb losowych (funkcja LOS w arkuszu kalkulacyjnym Excel). Tak więc

w przypadku doboru losowego, to „los”, a nie osoba przeprowadzająca badanie decyduje, które jednostki wejdą do próby. Skład próby jest całkowicie niezależny od woli badacza. Próba pobrana metodą losową nazywa się **próba losowa**.



Rysunek 1. Schemat mechanizmu wnioskowania statystycznego

Źródło: opracowanie własne.

Warunkiem pobrania próby losowej jest posiadanie tzw. **operatu losowania**, czyli *wykazu* ponumerowanych jednostek losowania. Operatem losowania może być na przykład:

- lista studentów (ponumerowanych według nazwisk lub numeru albumu),
- rejestr ludności (ponumerowanych według numeru PESEL),
- wykaz podatników (ponumerowanych według numeru NIP),
- wykaz przedsiębiorstw (ponumerowanych według numeru REGON),
- Krajowy Rejestr Sądowy,
- ogólnodostępne komputerowe bazy danych (spisy miejscowości, wykazy ulic, Panoramy Firm, książki telefoniczne).

Ważne jest, aby operat losowania był rzetelny i aktualny w momencie badania. Należy zaznaczyć, iż operat losowania odpowiedni dla celów jednego badania, może być nieodpowiedni dla celów innego badania. Na przykład rejestr ludności PESEL – odpowiedni dla celów zbadania poziomu płac ludności, jest nieodpowiedni w badaniach dochodów gospodarstw domowych. Przyczyną jest fakt, że gospodarstwa o większej liczbie osób mają większe szanse dostania się

do próby. Trudności w uzyskaniu rzetelnego i aktualnego operatu losowania są jednym z powodów uniemożliwiających losowy dobór próby. W takich przypadkach stosowane są *metody nielosowe*. Nielosowe metody doboru próby charakteryzują się tym, że wybór jednostek do próby całkowicie zależy od osoby przeprowadzającej badanie. Jest to więc wybór subiektywny, uzależniony od wiedzy i poglądów badacza. Uzyskana w ten sposób próba nie jest *próbą reprezentatywną* i nie można w tym przypadku stosować metod wnioskowania statystycznego.

Metody wnioskowania statystycznego wymagają losowego doboru próby, gdyż tylko wówczas można uzyskać **próbę reprezentatywną**. Za próbę reprezentatywną uważa się taką, która jest „miniaturką” populacji. Innymi słowy, odzwierciedla strukturę populacji i wszystkie prawidłowości w niej występujące.

Szczególne znaczenie we wnioskowaniu statystycznym ma tzw. **prosta próba losowa**. Jest to taka próba, która została pobrana za pomocą losowania *indywidualnego, nieograniczonego i niezależnego* (tj. ze zwracaniem). Losowanie spełniające te trzy warunki nazywa się **losowaniem prostym**. Tak więc próbę prostą uzyskuje się w wyniku losowania prostego, tzn. losując do próby pojedyncze elementy, każdorazowo z całej populacji. Jeśli nie jest spełniony którykolwiek z trzech warunków definiujących losowanie proste, wówczas mówi się o **losowaniu złożonym**, a próbę otrzymaną w wyniku takiego losowania określa się mianem **próby złożonej**.

W przypadku *n*-elementowej *prostej próby losowej*, każdy zbiór *n* elementów ma takie same szanse znalezienia się w próbie. W kolejnych rozdziałach książki zakładamy, że wszystkie analizowane próby są prostymi próbami losowymi (w skrócie: *próbami losowymi*). **Omawiane metody wnioskowania dotyczą prostych prób losowych.**

Wśród metod wnioskowania statystycznego wyróżnia się dwie zasadnicze grupy:

- **estymację**, czyli szacowanie nieznanymi parametrów populacji. Na przykład możemy być zainteresowani oszacowaniem średniego czasu obsługi klientów w sklepie i w tym celu przeprowadzimy stosowne obliczenia na podstawie pobranej próby klientów;
- **weryfikację**, czyli testowanie hipotez statystycznych (tj. przypuszczeń dotyczących właściwości populacji). Na przykład możemy postawić hipotezę, że występuje silny związek między czasem emisji reklamy pewnego produktu a zyskiem uzyskanym z jego sprzedaży, a następnie w wyniku zastosowania określonych procedur podjąć decyzję, czy hipotezę można uznać za słuszną, czy też fałszywą.

Wymienione metody wnioskowania przybliżymy w dalszej części książki.