

**FRANK A. GEARSON, KENNETH R. BENNET — STATISTICAL METHODS  
APPLIED TO AGRICULTURAL ECONOMICS**

Wydanie piąte, N. York 1955, str. 451.

Książka została napisana, co autorzy stwierdzają we wstępie, jako podręcznik dla studentów oraz jako pomoc dla osób interesujących się zastosowaniem metod statystycznych do badań ekonomicznych w rolnictwie. Ten cel zdecydował o układzie i sposobie przedstawienia omawianych zagadnień. Wychodząc z założeń, że czytelnik nie ma żadnego lub tylko nikie wiadomości z zakresu statystyki, autorzy rozpoczynają omawianie poszczególnych mierników statystycznych od zagadnień podstawowych, a dopiero dalej pokazują ich zastosowanie w dziedzinie badań ekonomicznych w rolnictwie.

Praca została napisana bardzo przystępnie, szczególnie dla osób, które niechętnie czytają książki z dużą ilością wzorów matematycznych. Omawianie każdego miernika statystycznego autorzy rozpoczynają od zwięzłego określenia jego treści. Następnie w oparciu o konkretny przykład liczbowy pokazują sposób jego obliczania. Zazwyczaj dopiero na zakończenie lub przy tablicach, jako wyjaśnienie techniki obliczeń, dają odpowiedni wzór czy wzory matematyczne. Sposób obliczeń poszczególnych mierników autorzy podają ponadto, zwłaszcza w początkowych rozdziałach książki, przy pomocy „wzorów tekstowych”.

Nawet mierniki dosyć skomplikowane w obliczaniu, jak na przykład współczynnik korelacji<sup>1</sup> przy badaniu związku dwóch zmiennych, a nawet współczynnik korelacji wielorakiej<sup>2</sup> są podawane najpierw przy pomocy wzorów tekstowych, a dopiero później przy pomocy wzorów matematycznych. Wzór matematyczny poszczególnych mierników jest więc nie punktem wyjścia, jak to ma miejsce w wielu podręcznikach, ale podsumowaniem tego, co czytelnik już się poprzednio dowiedział zarówno o istocie jak i technice jego obliczeń.

Na zakończenie działu traktującego o danym mierniku statystycznym autorzy wydzielają punkty: „Cechy charakterystyczne” oraz „Zastosowanie”. Jeżeli dla oświetlenia danego zagadnienia można stosować różne mierniki, to po omówieniu każdego z nich osobno autorzy podają na ogół dosyć szeroko ujęte ich porównanie.

Mimo tak przystępnego i dostosowanego w zasadzie poziomu do czytelnika o żadnym lub minimalnym zasobie wiadomości ze statystyki, autorzy potrafili uniknąć nadmiernego balastu objaśnień i przykładów. Książka jest zwięzła i przejrzysta, a tym samym łatwa do czytania i zrozumiała. Można ją śmiało polecić również tym osobom, które dość dobrze znają statystykę.

Zagadnienia omawiane w kolejnych rozdziałach można podzielić na trzy większe grupy. Pierwsza grupa (7 rozdziałów) obejmuje zagadnienia związane z analizą statystyczną jednej zmiennej. W grupie drugiej (8 rozdziałów) autorzy omawiają metody badania zależności pomiędzy zmiennymi. Trzecia część (7 rozdziałów) jest poświęcona zagadnieniu pewności („reliability”<sup>3</sup>) wartości miar statystycznych, gdy chcemy wyniki przeprowadzonych badań uogólnić.

W pierwszym rozdziale omówiono zasady budowy szeregu rozdzielczego oraz jego ogólną analizę. W oparciu o konkretne dane liczbowe podano najważniejsze typy rozkładów.

<sup>1</sup> Str. 151.

<sup>2</sup> Str. 167.

<sup>3</sup> W słowniku polsko-rosyjsko-angielskim statystyki matematycznej i statystycznej kontroli jakości. Polska Akademia Nauk, Instytut Matematyczny: PWN Warszawa 1958 nie zamieszczono tego terminu.

Dalej, zgodnie z klasycznym układem podręczników statystyki, pokazano zastosowanie oraz sposób obliczania średnich oraz miar dyspersji. Porównanie zastosowania, zalet i wad różnych średnich oraz dalej różnych miar zmienności porządkowej nabyte przez czytelnika wiadomości.

Przy omawianiu zagadnienia badania tendencji rozwojowej (rozdz. 5) autorzy obok metody analitycznej (metody najmniejszych kwadratów) znajdowania tendencji prostoliniowej podają kilka metod uproszczonych wyznaczania prostej tendencji. Metody te pozwalają na znaczną oszczędność czasu, a wyniki liczbowe jakie otrzymali oni w przytoczonym przez siebie przykładzie są podobne do wyników otrzymanych przy zastosowaniu metody analitycznej. Wydaje się jednak, że w przypadkach bardziej skomplikowanych niektóre metody uproszczone mogą dać wyniki obciążone większym błędem.

Metodę matematycznego znajdowania krzywej tendencji autorzy podali tylko dla linii tendencji o stałym przyroście względnym. Metoda średnich ruchomych zapoznaje czytelnika z badaniem tendencji krzywoliniowej, szczególnie w przypadkach jej bardziej złożonego przebiegu.

Dwa następne rozdziały (6 i 7) zajmują się metodami badania wahań w szeregach czasowych. Autorzy podają zarówno metody pozwalające badać sezonowość bez eliminowania tendencji rozwojowej, jak i metody które eliminują wpływ tendencji na obliczoną wartość wskaźników sezonowych. Osobny punkt tego rozdziału jest poświęcony zagadnieniu często w polskich podręcznikach nie uwzględnianemu — mianowicie eliminacji wahań sezonowych (tylko sezonowych a nie łącznie z przypadkowymi czy cyklicznymi). Metody badania wahań cyklicznych (rozdz. 7) są dla nas mniej interesujące. Analiza nawet najbardziej wnikliwa jednej tylko cechy bez powiązania jej z innymi nie daje pełnego obrazu badanego zjawiska.

Wielkie znaczenie przywiązują autorzy do problemu badania zależności. Świadczą o tym następujące słowa: „Každy człowiek od świtu do nocy dostrzega zależności, przyjmuje ich istnienie i podejmuje decyzje na ich podstawie. Ze względu na to, że istniejące zależności wpływają na wszelką ludzką działalność, są one najważniejszym zagadnieniem w statystyce”<sup>1</sup>. Odpowiednio do tych stwierdzeń ilość miejsca poświęconego metodom badania zależności w omawianym podręczniku jest proporcjonalnie bardzo duża.

Zależności możemy badać przy pomocy metody tabelarycznej i przy pomocy metody opierającej się na matematyce — metody korelacyjnej. Szczególnie tej pierwszej metodzie autorzy poświęcają bardzo wiele miejsca i uwagi. Już we wstępie podkreślali oni, że „Książka różni się od większości innych podręczników tym, że zawiera dwa rozdziały traktujące o tabelarycznej metodzie badania zależności. Metoda ta jest pomijana w większości podręczników mimo tego, że była ona i będzie dalej przy badaniu zależności najpowszechniej stosowaną”<sup>2</sup>.

Autorzy dali przejrzysty przegląd zasad stosowania metody tabelarycznej tak do badania zależności pomiędzy dwoma zmiennymi jak i zależności pomiędzy kilkoma czynnikami. Metoda tabelaryczna pozwala nie tylko stwierdzić czy zależność istnieje i jaki jest jej kierunek, ale pozwala również stwierdzić, w sposób łatwy i prosty czy regresja ma charakter prosty czy krzywoliniowy.

Na przykład autorzy wykazują na liczbach empirycznych, że przy różnym poziomie plonów wpływ obszaru na dochód ogólny jest zupełnie różny. Przy plonach bardzo niskich w miarę wzrostu obszaru dochód spada (a raczej rośnie strata), przy wysokich plonach wzrost obszaru wpływa bardzo silnie na wzrost dochodu<sup>3</sup>. Zjawisko jest oczywiste i zrozumiałe, interesująca natomiast ze względu na jej prostotę jest metoda tabelaryczna jego zbadania. Wymaga ona najogólniej biorąc podzielenia zbiorowości na kilka grup i obliczenia w tych grupach średniej wartości dochodu, a następnie przeniesienia otrzymanych danych na wykres. Szczególnie przydatna dla ekonomistów rolników jest metoda tabelaryczna ze względu na to, że pozwala jako cechę wyróżniającą jednostki stosować również cechy jakościowe (np. typ gospodarstwa, kierunek produkcyjny itp.).

Mimo znaczenia jakie autorzy przywiązują do metody tabelarycznej badania zależności, omawiają oni również w szerokim zakresie matematyczną metodę badania zależności — metodę korelacyjną. Obok metody badania zależności dwóch zmiennych w przypadku regresji krzywoliniowej, omówiono również metody wyznaczenia krzy-

<sup>1</sup> Str. 20.

<sup>2</sup> Str. V.

<sup>3</sup> Str. 133.

wych regresji i znajdowania wielkości zależności pomiędzy zmiennymi w przypadku korelacji wielorakiej. Najtrudniejsze bodaj przy badaniu zależności zagadnienie — metoda badania korelacji sprężonej zostało również dosyć przystępnie omówione.

Dalej przeprowadzone porównanie metody tabelarycznej i metody matematycznej badania zależności wyraźnie podnosi zalety tej pierwszej. Wyniki jakie autorzy otrzymali rozwiązując konkretne zagadnienie przy pomocy obu metod są prawie zbieżne.

Metoda tabelaryczna, mimo jej niewątpliwie licznych zalet, nie daje niestety odpowiedzi na jedno z podstawowych pytań przy badaniu zależności: jaka jest ścisłość zależności. Wprawdzie w ostatnich czasach większe znaczenie przywiązuje się do badania regresji niż korelacji<sup>1</sup>, niemniej wielkość współczynnika korelacji jest istotna ze względu chociażby na to, że wskazuje nam jak duże możemy przywiązywać znaczenie do znalezionej współczynnika regresji.

Dlaczego metoda tabelaryczna badania zależności nie jest bardziej powszechnie stosowana? Wydaje się, że istnieją tu dwie związane zresztą ze sobą zasadnicze przyczyny (nie mówiąc o zagadnieniu mniejszej precyzji metody tabelarycznej bo to jest znakomicie zrekomensowane przez prostotę obliczeń). Pierwszą przyczyną jest trudność techniczna: przy pomocy metody tabelarycznej trudno badać zależność większej liczby zmiennych, gdyż tablica rozrosłaby się do zbyt dużych rozmiarów, a liczba jednostek, z których obliczamy średnie w poszczególnych pozycjach tablicy mogłaby zbyt szybko zmniejszyć się. Wpłynęłoby to w konsekwencji na mniejsze zaufanie jakie moglibyśmy mieć do poszczególnych wartości.

Ostatnie siedem rozdziałów książki omawia zagadnienia związane z oceną prawdziwości wyników otrzymanych przy pomocy różnych miar statystycznych. Wprowadzenie do tych zagadnień jest tak charakterystyczne dla bezpośredniości i prostoty z jaką jest napisana cała książka, że nie możemy się oprzeć chęci jego zacytowania: „wiele doktryn na tym świecie nie jest zgodnych z rzeczywistością. Czy jakies mniemanie jest słuszne czy niesłuszne zależy częściowo od metody na podstawie której to mniemanie ugruntowało się. Większość ludzi dochodzi do danych przekonań na podstawie rozważań teoretycznych, na podstawie obserwacji albo na podstawie badań liczbowych. Ze wszystkich doktryn rozwiniętych przy pomocy rozważań teoretycznych jedynie niewielki procent okazał się być prawdziwy. Ze wszystkich zjawisk zaobserwowanych wysoki procent dowiódł swojej prawdziwości. Równie wysoka część doktryn opartych na badaniach liczbowych okazała się być prawdziwa... Statystyka zajmuje się mierzaniem — mierzaniem różnych wielkości charakteryzujących zjawiska, jak przeciętny poziom, dyspersję i współzależność. Teorie oparte na mierzaniu są częściej prawdziwe niż utworzone przy pomocy innych metod. Jednak, przy poszukiwaniu prawdy, doskonałość nie jest nawet tutaj osiągnana. Przypadek bowiem wpływa zarówno na wielkość mierników liczbowych jak na obserwację. Różnica jest tylko w stopniu jego oddziaływania”<sup>2</sup>.

Przy wszelkich badaniach statystycznych, których wyniki chcemy w jakiegokolwiek mierze uogólnić musimy więc brać pod uwagę oddziaływanie czynników przypadkowych.

Miarami, które pozwalają wyznaczyć z określonym prawdopodobieństwem granice w jakich leży prawdziwa wartość, są błędy standardowe poszczególnych mierników statystycznych. Autorzy omawiają sposoby obliczania i znaczenie błędów standardowych różnych parametrów statystycznych. Możliwość oceny z punktu widzenia teorii prawdopodobieństwa wyników otrzymywanych przy zastosowaniu metody tabelarycznej pozwala na szersze i „bezpieczniejsze” stosowanie tej posiadającej liczne zalety metody.

Drugą metodą umożliwiającą ocenę istotności otrzymanych wyników jest analiza wariancyjna. Zasadnicze podstawy przy analizie wariancyjnej są te same co przy błędach standardowych, technika obliczeń jest jednak inna. Analiza wariancyjna (której rozwój nastąpił później niż teoria błędów standardowych) może być stosowana do szerszego kręgu zagadnień niż metoda błędów standardowych. Autorzy pokazują jej zastosowanie do oceny istotności szeregu miar statystycznych, nie zapominając również o pokazaniu jej zastosowania do tabelarycznej metody badania zależności. Zagadnieniu badania zgodności rozkładu z rozkładem teoretycznym przy pomocy testu „Chikwadrat” autorzy poświęcili cały rozdział, omawiając szczegółowo jego

<sup>1</sup> Por. M. Ezekiel, K. A. Fox. *Methods of correlation and regression analysis*. Third Edition, New York 1959.

<sup>2</sup> Str. 300.

stosowanie na licznych przykładach empirycznych, w szczególności przy badaniu tabelarycznym zależności.

Wyniki badania otrzymane przy zastosowaniu metody korelacyjnej wymagają również przeprowadzenia oceny ich prawdziwości („reliability”). W rozdziale ostatnim autorzy przedstawili metodę przeprowadzania korekty miar korelacji otrzymanych z próbki, metodę oceny czy korelacja rzeczywiście istnieje czy nie, ocenę różnic w wartości współczynników. Ocena krzywoliniowości regresji oraz badanie istotności wyników otrzymanych na podstawie równań regresji zamyka ostatni rozdział pracy.

Obszerny indeks alfabetyczny podany na końcu książki ułatwia bardzo korzystanie z niej jako z podręcznika w pracy statystyczno-badawczej.

*Teresa Marszałkiewicz*