



## ZASTOSOWANIA STATYSTYKI W BADANIACH NAUKOWYCH – POTRZEBA CZY KONIECZNOŚĆ?

*Janusz Wątroba, StatSoft Polska Sp. z o.o.*

Przy próbie oceny korzyści i ograniczeń, jakie występują w przypadku stosowania statystyki na różnych etapach badania naukowego, warto rozważyć analogię do systemów sprawowania władzy. System demokratyczny jest dość powszechnie krytykowany, a jego przeciwnicy przytaczają wiele argumentów potwierdzających ich tezy. Z drugiej jednak strony można również spotkać się z opinią, że demokracja ma oczywiście swoje wady, ale nic lepszego do tej pory nie wymyślono. W przypadku wykorzystywania podejścia statystycznego badacz świadomie lub nieświadomie godzi się na cały szereg uproszczeń ze względu na brak lepszej alternatywy. Jednocześnie w bardzo wielu badaniach można zaobserwować dość swobodne traktowanie zarówno wymogów statystycznych, jak i metodologicznych. W niniejszym opracowaniu zwrócono uwagę na niektóre wybrane aspekty planowania i interpretacji wyników badań empirycznych, które powinny być brane pod uwagę przy ich podejmowaniu.

### **Zalety i ograniczenia podejścia statystycznego w badaniach naukowych**

Zastosowanie statystyki w badaniach empirycznych jest przez wielu badaczy negowane. Można chyba zaryzykować twierdzenie, że przynajmniej częściowo jest to spowodowane niewystarczającą wiedzą na temat rzeczywistych zalet i ograniczeń tego podejścia, jak również niedobłą praktyką, którą obserwuje się w opracowaniach publikowanych w różnych dziedzinach badań empirycznych. Jednym z głównych zarzutów jest twierdzenie, że podejście statystyczne prowadzi do zbytniego upraszczania rzeczywistego przebiegu badanych zjawisk oraz powiązań występujących pomiędzy czynnikami, które je determinują.

Trudno przynajmniej częściowo nie zgodzić się z takim stanowiskiem. Jednakże warto podkreślić, że przy empirycznej obserwacji zjawisk niemal zawsze dochodzi do różnego rodzaju uproszczeń i nieścisłości. Dotyczą one niemal każdego etapu badań. Zwrócimy uwagę na niektóre z nich, wybierając te, które wydają się mieć największe konsekwencje. W pierwszej kolejności zajmiemy się zagadnieniem możliwości uogólnienia wyników uzyskanych w pojedynczym badaniu na szerszą populację obiektów [4]. Kolejny problem to możliwość oszacowania liczebności obiektów, które należałoby objąć badaniami, aby



upewnić się, że obserwowany efekt nie jest dziełem przypadku. Dla każdego niemal badacza ważnym praktycznym problemem jest dobór odpowiednich metod opracowania danych pod kątem stawianych w badaniu hipotez lub pytań badawczych. W przypadku stosowania metod analizy statystycznej niezwykle ważnym zagadnieniem jest umiejętne przełożenie wniosków statystycznych na wnioski merytoryczne.

W niektórych dziedzinach badań empirycznych dostosowano zasady podejścia statystycznego do specyfiki badanych zjawisk (np. w doświadczalnictwie rolniczym, psychologii, pedagogice czy ekonometrii). W niektórych innych dziedzinach, ze względu na różne trudności i ograniczenia, wydaje się, że wspomniane zasady wymagają dalszych wysiłków w celu stworzenia bardziej spójnego podejścia do planowania, prowadzenia i interpretacji wyników badań (np. w medycynie i dziedzinach pokrewnych).

## **Czy statystyk może pomóc przy planowaniu badania?**

Do niedawna przynajmniej wśród części badaczy panowało przekonanie, że potrzeba zastosowania statystyki pojawia się dopiero po zgromadzeniu empirycznych danych. Obecnie coraz powszechniej zwraca się uwagę na potrzebę albo wręcz konieczność uwzględnienia podejścia statystycznego już na etapie planowania badania. Chodzi przede wszystkim o dwie kwestie: zapewnienie możliwości uogólnienia wyników badania na szerszą niż rzeczywiście poddana badaniom zbiorowość obiektów oraz określenie liczebności badanych obiektów niezbędnej do realizacji postawionych celów. Obydwa zagadnienia zostaną nieco dokładniej omówione poniżej.

### ***Zagadnienie trafności zewnętrznej badania***

W większości projektów badawczych nie ma możliwości objęcia badaniami całej interesującej badacza docelowo populacji obiektów. Wynika to z kilku zasadniczych powodów. Po pierwsze populacja taka jest zazwyczaj nieskończona. Po drugie, nawet jeśli populacja jest skończona, to koszt takich badań, czas ich trwania oraz inne trudności techniczno-organizacyjne uniemożliwiają w praktyce przeprowadzenie badania wyczerpującego. Ponadto w przypadku niektórych badań sama procedura badawcza prowadzi do przynajmniej częściowego zniszczenia obiektu.

W takiej sytuacji jedynym rozsądnym rozwiązaniem jest przeprowadzenie badania częściowego, które obejmie obiekty wybrane z populacji docelowej. Aby takie badanie spełniało standardy badania reprezentacyjnego, musi istnieć tzw. operat losowania, czyli lista wszystkich obiektów należących do populacji, a ponadto należy zapewnić, że każdy obiekt ma takie samo prawdopodobieństwo wybrania do zbiorowości, która zostanie objęta badaniami. W praktyce jest to zazwyczaj trudne do zrealizowania. W związku z tym rozważa się podzbiór obiektów, które spełniają postawione cele badania, a jednocześnie są dostępne do przebadania. W takiej sytuacji oczywiście dobór obiektów do badania musi zostać określony jako dobór celowy. Dla przybliżenia zbiorowości obiektów, która została poddana badaniom, określa się jej wyjściową charakterystykę. Dzięki temu jest możliwość jej porównania do innych projektów badawczych prowadzonych w tym samym obszarze.

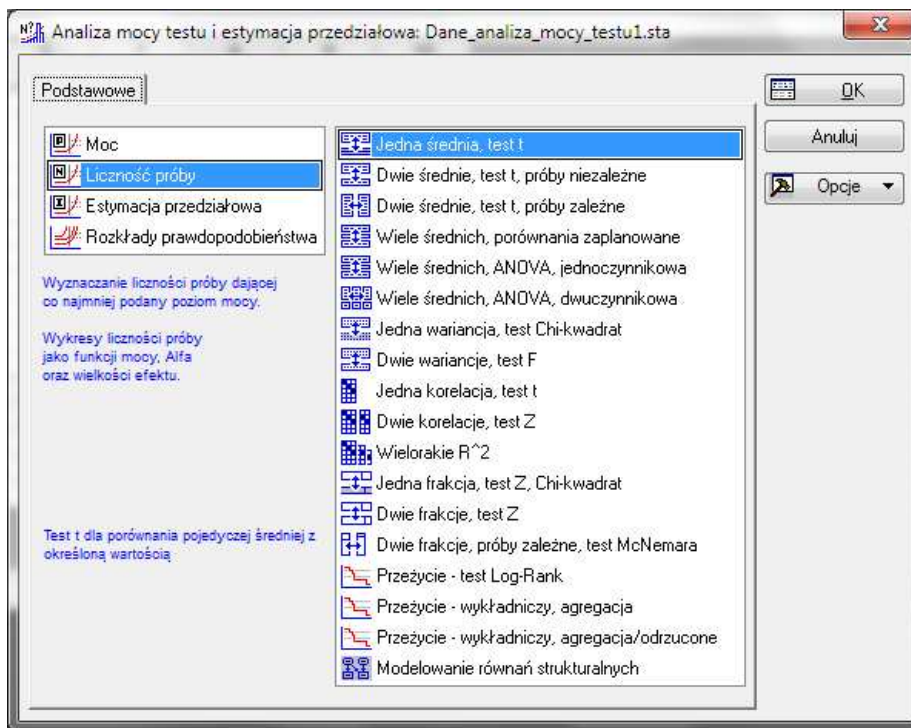


Oczywiście na badaczu ciąży obowiązek uzasadnienia możliwości uogólnienia wyników badań na populację docelową. W przypadku jeśli w danym obszarze badawczym zostało przeprowadzonych wiele badań o podobnej wyjściowej charakterystyce obiektów istnieje możliwość przeprowadzenia tzw. metaanalizy, która polega na połączeniu wyników podobnych badań dla uzyskania bardziej wiarygodnych wniosków. Z analizami tego typu można spotkać się w przypadku badań klinicznych prowadzonych dla oceny skuteczności i bezpieczeństwa stosowania leków lub procedur medycznych [2].

Kryteriabrane pod uwagę przy wyborze docelowej populacji obiektów powinny zostać dokładnie określone, gdyż to one wpływają na *trafność zewnętrzną* badania, czyli możliwość uogólnienia wyników danego badania na szerszą populację.

### ***Czy przed rozpoczęciem badania jest możliwość oszacowania liczebności obiektów, które powinny zostać uwzględnione w projekcie?***

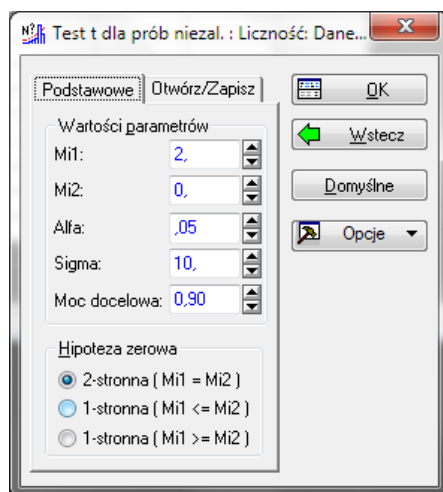
Jest to jedno z tych zagadnień, z którymi badacz najczęściej zwraca się o pomoc do statystyka. Aby można było rozwiązać ten problem, muszą zostać spełnione określone warunki. Przede wszystkim badacz musi bardzo precyzyjnie zdefiniować cel badania, uszczegółowić go w postaci hipotez badawczych i wstępnie dobrać metody ich statystycznej weryfikacji. W przypadku niektórych badań do oszacowania niezbędnej liczebności badanych obiektów wystarcza zastosowanie stosunkowo prostych formuł. Często jednak jest to związane z koniecznością uwzględnienia pojęcia mocy testu [1]. W praktyce oznacza to konieczność wskazania odpowiedniej techniki statystycznej, która będzie wykorzystywana przy testowaniu hipotezy badawczej.



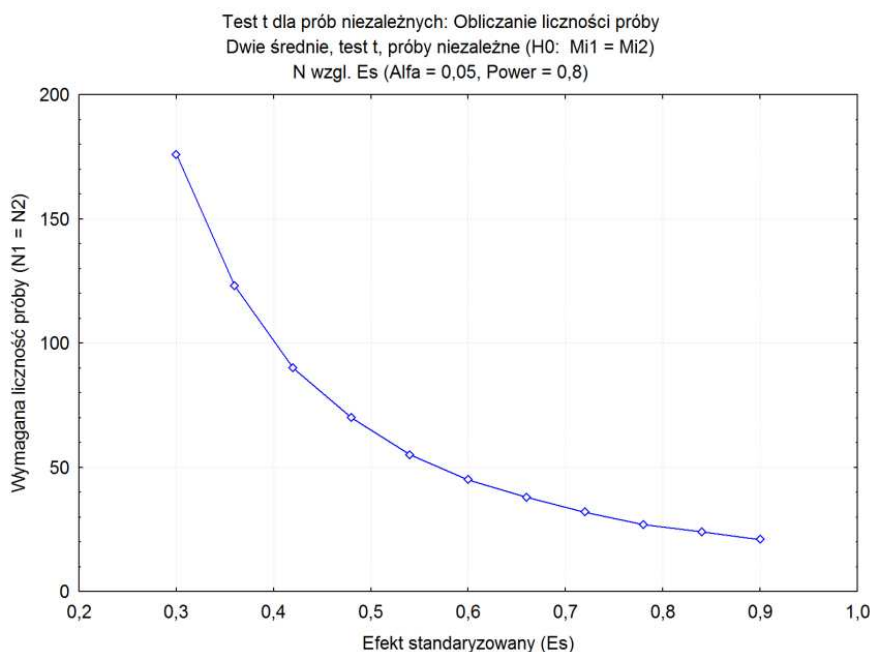
Rys. 1. Metody szacowania liczebności próby dostępne w programie *STATISTICA*.

Zamieszczony powyżej rysunek przedstawia wybór odpowiednich narzędzi statystycznych przeznaczonych do szacowania liczebności dostępnych w module *Analiza mocy testu* programu *STATISTICA*.

W zależności od postawionego w badaniu problemu badawczego oraz wybranej metody jego statystycznej weryfikacji oszacowanie niezbędnej liczebności obiektów wymaga jeszcze określenia pewnych dodatkowych parametrów. Na przykład jeśli postawiona hipoteza badawcza wymaga użycia testu *t-Studenta* dla prób niezależnych, wówczas dla oszacowania liczebności porównywanych grup należy podać oczekiwane wartości średnich, wielkość odchylenia standardowego w połączonych próbach, zakładany poziom istotności testu  $\alpha$  oraz moc testu.



Rys. 2. Przykładowe okno definiowania parametrów analizy mocy testu.



Rys. 3. Zależność pomiędzy wielkością efektu standaryzowanego a wymaganą liczebnością porównywanych prób.



Zamieszczony powyżej wykres pokazuje zależność pomiędzy wielkością standaryzowanego efektu (jest on wyliczany w postaci ilorazu różnicy zakładanych średnich i wspólnego odchylenia standardowego) a wymaganą liczebnością badanych prób.

Dzięki wykresowi tego typu jest możliwe oszacowanie liczebności w sytuacji, gdy badacz nie wie dokładnie, jakiej różnicy pomiędzy badanymi grupami może się spodziewać. W takiej sytuacji należy rozważyć kilka różnych wielkości efektu standaryzowanego i na podstawie wykresu oszacować zakres wymaganej liczebności.

## **Dobór odpowiedniej metody opracowania wyników badań**

Kluczowym zagadnieniem przy doborze odpowiednich metod statystycznego opracowania zebranych danych jest zrealizowanie postawionych celów badawczych. Jak to zostało już wcześniej wspomniane przy okazji omawiania problemu szacowania liczebności obiektów, które powinny zostać poddane badaniom, cele badawcze powinny zostać przełożone na hipotezy i pytania badawcze. Dopiero wtedy jest możliwe dobranie określonych metod statystycznych, które zostaną zastosowane.

W przypadku dobrze zaplanowanego badania wybór odpowiednich metod analizy statystycznej powinien być dokonywany wstępnie w fazie jego projektowania. Jednakże ostateczny wybór powinien być przeprowadzony po wcześniejszym sprawdzeniu założeń wymaganych przez poszczególne metody. W związku z tym zazwyczaj jest bardzo ważne, aby rozpocząć analizę od eksploracyjnej analizy danych, która wymaga utworzenia wykresów i obliczenia statystyk opisowych dla oceny rozkładów zmiennych, źródeł zmienności, powiązań pomiędzy zmiennymi oraz niezależności obserwacji. Wstępne badanie dotyczy zazwyczaj rozkładu zmiennej zależnej, a także badania pod kątem odstających obserwacji. Jeśli założenia nie są spełnione, wówczas plan analizy powinien uwzględniać metody alternatywne. Dane o charakterze dychotomicznym lub skategoryzowanym powinny być analizowane przy pomocy metod przeznaczonych dla danych jakościowych. Podobnie metody „klasyczne” są generalnie odpowiednie dla danych, które w przybliżeniu podlegają rozkładowi normalnemu oraz charakteryzują się wystarczająco dużą liczebnością. W niektórych przypadkach, np. dla danych dotyczących zliczeń lub przeżyć z obserwacjami uciętymi, wymagane są specjalne techniki modelowania. Każdy model lub metoda statystyczna mając swój odrębny zestaw założeń, które powinny być sprawdzane łącznie z dopasowaniem modelu. Wymaga to zazwyczaj badania odchyleń pomiędzy wartościami rzeczywistymi i przewidywanymi przez model (odchylenia te są nazywane resztami). Za pomocą wykresów często udaje się zidentyfikować trendy lub inne problemy z rozkładem reszt.

Wybór stosownych metod statystycznego opracowania danych zależy po części od przyjętego przez badacza schematu badawczego, który uwzględnia takie kwestie jak: charakter badania (konfirmacyjne lub eksploracyjne), typ problemu badawczego (ocena różnicowania czy modelowanie zależności), liczbę i charakter grup porównawczych oraz brane pod uwagę zmienne (jakościowe, ilościowe). Niektóre kryteria decydujące o wyborze są możliwe do sprawdzenia dopiero po zebraniu danych (np. sprawdzenie, czy zebrane dane spełniają założenia wybieranych do zastosowania metod). W związku z tym badacz powinien



także uwzględnić możliwość zastosowania alternatywnych sposobów analizy danych (np. metod nieparametrycznych, które mają zazwyczaj mniejsze wymagania formalne).

Przy wyborze odpowiednich metod analizy statystycznej zalecane jest skorzystanie z porady doświadczonego statystyka, który zagwarantuje wzięcie pod uwagę problemów statystycznych pojawiających się w praktyce badań empirycznych. Dobrze jest, jeśli taki konsultant ma również praktyczne doświadczenie w dziedzinie, której dotyczą prowadzone badania. Pozwala to na uniknięcie wielu możliwych komplikacji.

## Interpretacja wyników analiz statystycznych

Przystępując do interpretacji wyników przeprowadzonych badań, powinno się rozpocząć od statystycznej interpretacji wyników analiz. Wiele ze stosowanych metod analizy danych pozwala na oszacowanie wielkości efektu badania (skuteczności zastosowanego zabiegu, metody nauczania czy siły bodźca) w postaci różnicy przeciętnego poziomu branej pod uwagę zmiennej zależnej pomiędzy grupami porównawczymi. Zaleca się podawanie wielkości efektu wraz z odpowiednim przedziałem ufności dla szacowanego efektu. Przedział ten może być interpretowany jako zakres wartości efektu wynikający z zebranych danych. Zazwyczaj podaje się granice 95% przedziału ufności. Niezależnie od podania przedziałów ufności wyniki badania powinny zostać także ocenione w kategoriach ich statystycznej istotności. Standardem jest podanie wartości prawdopodobieństwa określającego stopień przekonania badacza, że otrzymane wyniki (albo wyniki jeszcze bardziej ekstremalne) zostały uzyskane na drodze czysto losowej, przy założeniu braku efektu zróżnicowania porównywanych grup obiektów.

Duża ostrożność jest zalecana w przypadku przeprowadzania interpretacji wyników analiz statystycznych, które nie zostały wcześniej przez badacza zaplanowane w szczegółowym protokole badania. Chodzi tu o tzw. *analizy w podgrupach* oraz metody korygowania wyników analizy przy uwzględnieniu zróżnicowania wyjściowych charakterystyk porównywanych grup obiektów. Każda z zastosowanych metod powinna zostać dokładnie opisana i wyjaśniona.

## Przełożenie wniosków statystycznych na wnioski merytoryczne

Jest to etap badań, który wymaga od badacza szczególnego zachowania ostrożności. Zawsze przy takiej interpretacji badacz musi przede wszystkim mieć na uwadze różnicę pomiędzy istotnością statystyczną otrzymanych wyników a ich istotnością (ważnością, znamiennością) związaną z merytoryczną oceną zaobserwowanych różnic. Problem ten dobrze ilustruje sytuacja, która ma często miejsce w przypadku wyników badań klinicznych w medycynie. Chodzi o to, że wynik statystycznie istotny niekoniecznie oznacza efekt, który ma znaczenie dla praktyki klinicznej lub oznacza zauważalną dla pacjenta poprawę jego stanu zdrowia. Wnioski merytoryczne powinny dotyczyć możliwości wyjaśnienia otrzymanych wyników na bazie postulowanego mechanizmu. W miarę możliwości



otrzymane wyniki powinny zostać skonfrontowane z wynikami uzyskanymi przez innych autorów w podobnych badaniach (najbardziej zalecanym sposobem jest tutaj przegląd systematyczny lub metaanaliza wyników innych badań). Przy interpretacji merytorycznej należy też zwrócić uwagę na ograniczenia, które wystąpiły w zrealizowanych badaniach.

## **Potrzeba dalszego rozwoju zastosowań podejścia statystycznego w różnych dziedzinach badań naukowych**

Obecny stan rozwoju zastosowań metod statystycznych do wspomagania różnych etapów badań naukowych pokazuje, że nie jest możliwe opracowanie jednolitego podejścia, niezależnego od specyfiki szczegółowych dziedzin. W niektórych obszarach badań empirycznych podstawy takiego podejścia zostały już opracowane. Sytuacja taka ma miejsce w doświadczalnictwie rolniczym czy psychologii eksperymentalnej. W niektórych innych dziedzinach powstały odrębne dyscypliny zajmujące się specyficznymi metodami stosowanymi w wybranych obszarach badawczych (np. ekonometria czy chemometria). Oczywiście dyscypliny te opierają się na metodach, które pierwotnie zostały wprowadzone w statystyce.

Chyba najbardziej dynamiczny rozwój zastosowań metod statystycznych obserwuje się obecnie w szeroko pojętym obszarze nauk biomedycznych. Jest on spowodowany niezwykle trudnymi warunkami do prowadzenia badań empirycznych, jakie występują w medycynie i naukach pokrewnych. Projektowanie badań z udziałem pacjentów wymaga bowiem uwzględnienia różnych ograniczeń metodologicznych, medycznych, prawnych i etycznych.

Oprócz tego w przypadku badań klinicznych instytucje, które są odpowiedzialne za dopuszczanie do obrotu produktów leczniczych i wyrobów medycznych w Europie (*EMEA, European Medicines Agency*) i Stanach Zjednoczonych (*FDA, Food and Drug Administration*), stawiają coraz wyższe wymagania [2].

Jednocześnie coraz wyższe standardy są wymagane przez redakcje czasopism naukowych, które najczęściej odwołują się do określonych zaleceń tworzonych przez specjalnie dobrane gremia doświadczonych badaczy [3].

## **Zastosowania statystyki - potrzeba czy konieczność?**

Biorąc pod uwagę korzyści z poprawnego stosowania metod statystycznych przy opracowywaniu wyników badań naukowych oraz związane z nimi ograniczenia, można zauważyć zarówno potrzebę wykorzystywania narzędzi statystycznych, jak i w niektórych obszarach wręcz konieczność. W dużej mierze zależy to od tego, na jakim etapie, jeśli chodzi o rozwój badań empirycznych, znajduje się określona dziedzina. Tam, gdzie metody badań empirycznych są rozwijane od dłuższego czasu, zazwyczaj obowiązują wyższe standardy, zarówno jeśli chodzi o zakres stosowanych narzędzi, jak i poprawność ich stosowania. Z kolei w dziedzinach, w których badacze zaczynają prowadzić badania empiryczne, standardy na razie nie są postawione zbyt wysoko. Należy jednak sądzić, że z czasem sytuacja



ulegnie zmianie w takim samym kierunku jak w tych obszarach badań, w których badania empiryczne są prowadzone od dłuższego czasu.

## Podsumowanie

W opracowaniu omówiono pewne wybrane aspekty zastosowania metod statystycznych w badaniach naukowych. Wskazano na korzyści, jakie może czerpać badacz na różnych etapach prowadzenia badań empirycznych ze znajomości szeroko pojętego podejścia statystycznego. Jednocześnie zwrócono również uwagę na cały szereg ograniczeń, które należy brać pod uwagę przy stosowaniu metod statystycznych.

Obserwowany obecnie stan zaawansowania zastosowań metod statystycznych, jak i związane z nim standardy różnią się w zależności od dojrzałości metodologicznej danej dziedziny badań. Można jednak zaobserwować wyraźną tendencję do coraz powszechniejszego wykorzystania statystyki w kolejnych obszarach badawczych.

## Literatura

1. Bausell R.B., Li Y.-F. (2002) Power Analysis for Experimental Research. A Practical Guide for the Biological, Medical and Social Sciences, Cambridge University Press.
2. Chow, S.C., Liu, J. P. (2004) Design and Analysis of Clinical Trials. Concepts and Methodologies, (2<sup>nd</sup> ed.), John Wiley and Sons, NJ: Erlbaum.
3. Moher D., Hopewell S., Schulz K.F., Montori V., Gøtzsche P.C., Devereaux P.J., Elbourne D., Egger M., Altman D.G., CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomized trials, *at all. CONSORT, J Clin Epidemiol.* 2010 Aug; 63(8).
4. Podstawy EBM, czyli medycyny opartej na danych naukowych dla lekarzy i studentów medycyny, pod red. P. Gajewskiego, R. Jaeschke i J. Brożka, Medycyna Praktyczna, Kraków 2008.
5. Rao C., R., 1994, Statystyka i prawda, PWN, Warszawa.