PRZEGLĄD STATYSTYCZNY
R. LVII – ZESZYT 4 – 2010

GRZEGORZ KRZYKOWSKI

FILOZOFICZNE KONCEPCJE ESTYMACJI STATYSTYCZNEJ

Estymacja jako obszar analiz statystycznych danych pomiarowych jest wyznaczona trzema głównymi nurtami opisanymi w monografiach E.L. Lehmanna [9], J.W. Tukey [15] i A. Gelmana ze wsp. [5]. Zagadnienia estymacji występujące w tych monografiach pojawiają się w wielu wcześniejszych pracach, ale pełne prezentacje kierunków estymacji zostały szczegółowo opisane przez wyżej wymienionych autorów.

1. REDUKCJA – TEORIA ESTYMACJI PUNKTOWEJ

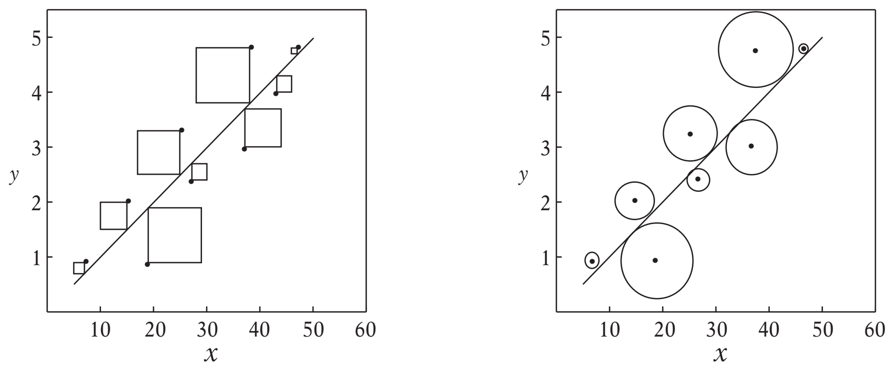
Na wstępie postaramy sie odpowiedzieć na pytanie, jakie główne koncepcje prowadzą do rozróżnienia podejścia do estymacji. Przyczyn rozróżnienia należy szukać w filozoficznej teorii nauki. W książce E.L. Lehmanna [9] zaprezentowano teorię estymacji punktowej (porównaj również [2]). W tej koncepcji przyjmujemy, że rzeczywistemu zjawisku badawczemu przypisujemy model statystyczny składający się z rodziny rozkładów. Zakładamy, że rodzina rozkładów jest indeksowana parametrem. Zwykle przyjmujemy techniczne założenie dotyczące własności tej parametryzacji, na przykład identyfikowalność parametryzacji (przykładowo porównaj [3]). Celem estymacji w tym podejściu jest wybór rozkładu (lub rozkładów) z przyjętej rodziny, zgodnie z różnymi kryteriami optymalizacji. Estymatorem jest rozkład lub zbiór rozkładów.

Wybranie rozkładu sprowadza się na ogół do wybrania parametru rozkładu. W najprostszym przykładzie przyjmujemy, że badanemu zjawisku eksperymentalnemu przypisujemy rodzinę rozkładów normalnych $\mathcal{P} = \mathcal{N}(\mu, \sigma^2)\}$, gdzie μ i σ^2 są parametrami identyfikującymi rozkład z tej rodziny ($\mu \in (-\infty, \infty)$, $\sigma^2 \in (0, \infty)$). Jeśli $\underline{x}$ stanowi realizacje próby losowej prostej $(X_i)_{i=1}^n$ z tej rodziny, to uznajemy, że średnia z próby $\bar{x}$ identyfikuje grupę rozkładów normalnych o wartości oczekiwanej $\bar{x}$ i nieznanej wariancji.

W podejściu praktycznym, wartości estymatora nadajemy stosowną interpretację (jest to wartość oczekiwana), ale w sensie teoretycznym wybór estymatora w tym przypadku jest sprecyzowaniem podzbioru rozkładów w rodzinie rozkładów normalnych. Zwróćmy uwagę, że wynikiem estymacji jest podzbiór rodziny rozkładów.

W niektórych zagadnieniach jest to podzbiór jednoelementowy, w niektórych (jak w powyższym przykładzie) bardziej rozbudowany. Często staramy się nadać wynikowemu podzbiorowi rozkładów pewną strukturę funkcyjną.

Na przykład przyjmujemy, że zbiór parametrów jest podzbiorem przestrzeni $\Theta \in \mathbb{R}^k$ i wynik estymacji traktujemy jako podzbiór rodziny rozkładów wyznaczony przez podzbiór w Θ. Sytuacje takie, w szczególności gdy wyposażymy zbiór parametrów Θ w strukturę przestrzeni liniowej, nazywamy estymacją parametryczną. Jeśli nie zadajemy parametryzacji lub nie uwzględniamy jej struktury (na ogół liniowej), to estymację traktujemy jako nieparametryczną. Rozróżnienie estymacji parametrycznej i nieparametrycznej nie jest do końca jednoznaczne, jednak nie to jest głównym zagadnieniem w tym artykule. Przejdźmy do dalszych zagadnień.



Rysunek 1. Regresja średniokwadratowa i „średniokołowa"

Metod estymacji punktowej, czyli kryteriów optymalizacyjnych implikujących wybór rodziny rozkładów, jest bardzo dużo i każdy badacz indywidualnie decyduje o wyborze metody estymacji. Przytoczymy przykład z zakresu znajdowania linii regresji. Przyjmujemy, że pomiary ($\underline{x}$, $\underline{y}$) stanowią realizację próby losowej prostej $(X_i, Y_i)_{i=1}^{n}$ z rodziny rozkładów dwuwymiarowych oznaczonej $\mathcal{P}$, o której zakładamy tylko tyle, że wartości oczekiwane rozkładów brzegowych są skończone. Dla dowolnego rozkładu dwuwymiarowego z tej rodziny istnieją stałe $(a, b) \in R^2$ takie, że $\mathbb{E}Y_i = a + b\mathbb{E}X_i$. Problem znalezienia oszacowania dla parametrów a i b jest przedmiotem estymacji punktowej. Zauważmy jednak, że w zagadnieniu tym rodzina rozkładów jest nieidentyfikowalna i *de facto* prowadzimy estymację nieparametryczną.

Wróćmy jednak do metody estymacji punktowej. Klasyczne podejście pochodzące od A.M. Legendrea[1] opiera się na metodzie optymalizacji średniokwadratowej, w której estymatory $\hat{a}$ i $\hat{b}$ są zadane klasycznymi wzorami Legendrea.

Pewne zdziwienie wywołałoby napisanie estymatorów dla współczynników a i b w postaci (patrz na przykład [11])

$$\hat{a} = \bar{y} - \hat{b}\bar{x}, \ \hat{b} = \frac{-S_x S_y + \sqrt{(S_x S_y)^2 - 4S_x S_y}}{S_{xx}^2}.$$

[1] Ostatecznie po dyskusjach historyków statystyki, za twórcę metody najmniejszych kwadratów uznano Legendrea, a nie jak wcześniej sądzono Gaussa [14]

Najczęściej recenzenci publikacji, w której zostałaby podana taka estymacja nie uznaliby jej za właściwą. Czy do końca mieliby rację?

Spójrzmy na rysunek 1. Okazuje się, że estymacja klasyczna (Legendrea) opiera się na minimalizacji średniokwadratowej odległości na osi Y, a przytoczony estymator parametrów *a* i *b* na minimalizacji średniokwadratowej odległości punktu od prostej (patrz na przykład [11]). Estymacja klasyczna jest zdecydowanie prostsza w obliczeniach analitycznych, ale można jej wiele zarzucić merytorycznie. Współcześnie, gdy praktycznie wszystkie obliczenia statystyczne są wykonywane w pakietach komputerowych, zastosować można bardziej skomplikowane obliczenia. Badacz powinien na tyle zapoznać się z zagadnieniem, nad którym pracuje, by móc określić jakie przesłanki leżą u podstaw tej czy innej metody estymacji punktowej. Podstawą wyboru określonej metody w estymacji punktowej powinno być merytoryczne uzasadnienie wyboru procedury optymalizacyjnej, a nie przywołanie klasycznych monografii, w których opisano konkretne standardowe metody estymacji punktowej.

W sensie filozofii nauki estymację punktową, polegającą na wyborze podzbioru rozkładów z zadanej rodziny rozkładów, należy uznać za formę rozumowania nazwanego redukcją. Redukcja jest metodą wnioskowania, która była analizowana i rozwijana przez polskich filozofów tworzących szkołę lwowsko-warszawską (patrz [16]).

W ogólnej metodologii nauk wyróżniamy pojęcia „racja” i „następstwo”. W obszarze znaczeniowym „racji” znajduje się rodzina rozkładów, a w obszarze znaczeniowym „następstwa” zespół danych pomiarowych. Według jednego z twórców szkoły, Jana Łukasiewicza, rozumowanie redukcyjne polega na znalezieniu „racji” na bazie „następstwa”. Rozumowanie to odpowiada procesowi estymacji punktowej[2].

2. INDUKCJA – EKSPLORACJA DANYCH

O ile metody estymacji punktowej tworzone były przez dziesiątki, a nawet setki lat, o tyle estymacja, którą można nazwać eksploracją danych, powstała dość raptownie. Powstanie metodologii EDA (Exploratory Data Analysis) wiążemy z pojawieniem sie książki J.W. Tukeya o tym samym tytule[3]. Pojawienie się tego podręcznika uświadomiło statystykom podział estymacji na metodologię EDA i wcześniejszą CDA (Confirmatory Data Analysis). Przed publikacją monografii Tukeya [15] nie uważano za konieczne wyróżnienie tego podziału. Metoda eksploracji danych stosowana była prywatnie, niejako lokalnie, w etapie wstępnym poprzedzającym metodę estymacji punktowej. Takie wykorzystanie eksploracji danych wynika ze złożoności postępowania badawczego. Metoda EDA polega na analizie wyników pomiarów bez wchodzenia w naturę probabilistyczną pozyskanych danych badawczych. Można więc zapytać, co jest wynikiem takiej estymacji? W sensie teoretycznym odpowiedź nie jest oczywista, ale w sensie praktycznym owszem. Wynikiem estymacji są wszelkiego rodzaju pod-

[2] Konsekwencje koncepcji Łukasiewicza prowadziły do trudności w klasyfikacji nauk, co było przedmiotem dyskusji w ramach szkoły lwowsko-warszawskiej (patrz [1]).

[3] Książka [15] powstaje po wielokrotnych wcześniejszych publikacjach stanowiących pierwotne składowe ostatecznej publikacji. Fragmenty te funkcjonowały w środowisku statystycznym w postaci niepublikowanych odbitek przez ponad 20 lat.

sumowujące wykresy. Należy podkreślić, że stosowanie eksploracji danych kończy się podsumowaniem w postaci opisu graficznego.

Wszelkie dalsze spostrzeżenia badawcze są już konsekwencją przyjrzenia się wykresom zagregowanych danych. Niektóre formy graficzne zaproponowane przez Tukeya nie utrzymały się w takiej postaci, jak to zostało zaproponowane w jego monografii. Na przykład wykresy typu „gałęzie i liście" („stem-and-leaf") nie wytrzymują konkurencji ze współczesnymi histogramami, ale wykresy skrzynkowe (box-plot) zrewolucjonizowały prezentację statystyczną. Dodatkowym czynnikiem, który wpłynął na efektywne wykorzystanie metod eksploracji danych, jest rozwój techniki obliczeniowej. Współczesne pakiety statystyczne oferują całą gamę graficznej prezentacji wyników pomiarów, z których duża część ma swoje źródło w prezentacjach Tukeya. Wykorzystanie pięciu statystyk porządkowych[4] i ich prezentacja graficzna pozyskała dodatkowe określenia[5] i nowatorską współczesną oprawę graficzną. Idea pozostaje jednak przy koncepcji Tukeya.

J.W. Tukey był genialnym dzieckiem, genialnym studentem i genialnym naukowcem. Przez 40 lat pracował w Bell Telephone Laboratories i jednocześnie przez tych 40 lat był członkiem wspólnoty akademickiej Princeton. Jego myśl widać we wszystkich znaczących osiągnięciach matematycznych, poczynając od Pewnika Wyboru, a kończąc na procedurze FFT i analizie spektralnej[6]. Współuczestniczył w niemal wszystkich znaczących projektach badawczych tamtych czasów – począwszy od badań nad bombą wodorową[7] i prac nad konstrukcją samolotów U2, poprzez analizę raportu Kinsey'a, udział w opracowaniu spisów ludności, wkład w podstawy rozwoju komputerów i wiele, wiele innych. Jego wyniki badawcze, jak zresztą wielu statystyków, mają charakter pomocniczy i na ogół są prezentowane w postaci wyników grupowych. Pozostawał w przyjaźni z największymi umysłami końca XX wieku, poczynając od Feymanna, Ulama, Halmosa czy Fellera. Trudno przecenić jego wpływ na rozwój statystyki. W pracy [4] D.R. Brillinger podaje 72 sformułowania statystyczne, które uzyskały nowe znaczenie w wyniku wpływu podejścia J.W. Tukeya, między innymi, takie jak „ANOVA", „jackknife", „software", „Lemat Zorna".

To symboliczne zaprezentowanie sylwetki J.W. Tukeya należy zderzyć z analizą eksploracyjną. Tukey proponuje prostą, żeby nie powiedzieć prymitywną analizę jako fundament wnioskowania. Jest w tym coś, co cechuje wielkich ludzi. Ma świadomość, że dla ogromnej rzeszy badaczy, nie zawsze genialnych i często wykorzystujących statystykę jako narzędzie, należy mówić obrazkami, na których znajduje się prezentacja zagregowanych danych w postaci kropek, kwadracików czy kreseczek. Wiele osób stosowało i stosuje taką wstępna analizę, ale trzeba było wielkiej osobowości Tukeya, żeby świat nauki zaakceptował to podejście jako wartościowe działanie badawcze. Nie stało to się od razu i do dzisiaj mamy pewien opór w akceptacji opracowania wyników pomiarów, w których nie przeprowadzono procedury testowania (w szczególności w publikacjach medycznych). Nie zawsze słusznie.

[4] Statystyki Tukeya: Min, Q_{25}, Q_{50}, Q_{75}, Max.

[5] Na wykresach skrzynkowych prezentujemy dodatkowo pomiary „odstające" i „ekstremalne".

[6] Fast Fourier Transform.

[7] Tukey uczęszczał na wykłady Ulama z topologii, a potem już w Princeton blisko współpracowali.

Szereg publikacji, w szczególności raportów dotyczących gospodarki, rozwoju biznesu czy produkcji firmy, nie potrzebuje testowania. Celem tych raportów jest prezentacja obserwacji bez wchodzenia w strukturę losową przedmiotu obserwacji. Prezentacje przeznaczone są dla szerokiego kręgu odbiorców, zatem muszą być dla nich zrozumiałe i czytelne. Stąd mamy niezwykłą wprost popularność metod analiz wskaźników giełdowych opisanych w książce J.J. Murphy'ego [10]. Metody zawarte w podręczniku Murphy'ego nie opierają się na strukturze losowej. Polegają na wyobrażeniu zachowania się wartości finansowych w postaci wykresów wzbogaconych o linie ograniczeń, trendów czy prezentację zmienności. Przypomnijmy, że w 2003 roku Nagrodę Nobla w zakresie ekonomii otrzymali Robert F. Engle i Clive W.J. Granger. Autorzy ci, w swoich analizach statystycznych cen na giełdzie opierają się na wyjątkowo zaawansowanych modelach losowych i ich wyniki stanowią najbardziej ewidentne zastosowanie CDA. Jednak „świat finansowy" ogląda wykresy świecowe lub fale Elliotta, które są w sferze opracowań EDA i te wyniki są podstawą niezliczonej ilości cząstkowych decyzji finansowych.

Na koniec powróćmy do metodologii badań naukowych. Metodę EDA należy lokować w sferze myśli wywodzących się od Sokratesa[8].

Jest to typowa postać indukcji rozumianej jako wnioskowanie od „szczegółu" do „ogółu". Dane empiryczne traktujemy jako przesłanki szczegółowe, a wnioski statystyczne dotyczące analizowanego zjawiska badawczego określamy jako „ogół". Na przykład pod filozoficznym pojęciem „ogółu" możemy rozumieć sentencję „cena waloru wzrośnie".

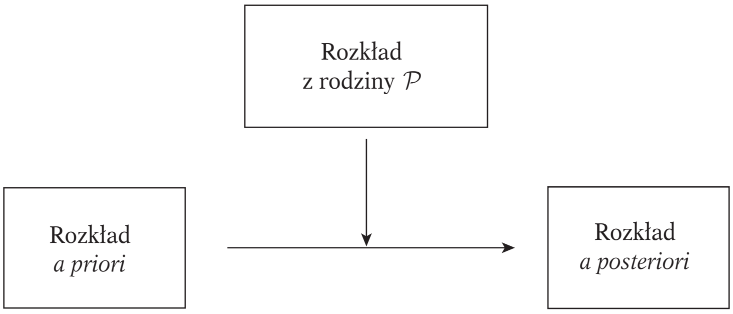
Biorąc pod uwagę, jak Sokrates w Dialogach platońskich dochodził do definicji danego „przedmiotu", „cechy", zauważamy zbieżność z analizą eksploracyjną. Stanowisko badawcze określone na przykład w sformułowaniu „cena waloru wzrośnie", jest konsekwencją wyabstrahowania analogii z szeregu obserwowanych zjawisk (zjawisk dotyczących dotychczasowego zachowania się rynków finansowych). Formalnie rozumując, tak właśnie postępujemy. Dysponujemy zbiorem wyników badawczych i poprzez analizę tych wyników wnioskujemy o konsekwencjach badawczych (patrz [12]).

3. DEDUKCJA – ESTYMACJA BAYESOWSKA

Estymacja bayesowska wymaga znacznej wiedzy matematycznej i intuicji badawczej. Jest najtrudniejsza w stosowaniu, ale chyba najbardziej efektywna. Zobowiązuje do najwyższego poziomu dyscypliny i solidności badawczej.

Samą procedurę postępowania można zapisać w postaci schematu przedstawionego na rysunku 2. Podejście Bayesa do estymacji ma charakter decyzyjny.

[8] „Sokrates słusznie poszukiwał istoty, poszukiwał bowiem sylogizmów i zasady sylogizmów, bo «to, czym rzecz jest» [istota], jest punktem wyjściowym sylogizmów. [...] Dwa bowiem odkrycia można słusznie przypisać Sokratesowi: wnioskowanie indukcyjne i ogólną definicję, jedno i drugie dotyczy początku wiedzy" (Arystoteles, Metafizyka, A 6, 987 b 1 nn).



Rysunek 2. Schemat procedury Bayesa

Źródło: wyniki własne [7].

Podobnie jak w przypadku estymacji punktowej, próba losowa prosta pochodzi z rozkładu z rodziny rozkładów $\mathcal{P}$. W przypadku podejścia Bayesa, zawsze rozkłady z rodziny $\mathcal{P}$ są indeksowane parametrem. Jest to więc sparametryzowana rodzina rozkładów. Akcent badawczy jest przeniesiony z wyboru rozkładu z rodziny $\mathcal{P}$ na wybór parametru. Posługując się prostym przykładem powiemy, że dążymy do oszacowania wartości oczekiwanej z próby losowej prostej, czy znajdujemy rozkład z rodziny rozkładów o wybranej wartości oczekiwanej. W sensie praktycznym są to identyczne informacje, ale w sensie teoretycznym rozpatrujemy różne przestrzenie. W pierwszym przypadku jest to zbiór parametrów, zaś w drugim jest to zbiór rozkładów. Z tym problemem spotykamy się już w przypadku estymacji punktowej, jednak w podejściu bayesowskim rozróżnienie zbioru parametrów i rodziny rozkładów jest daleko bardziej istotne. Można przyjąć, że skupiamy się głównie na parametryzacji rodziny. Oczywiście, nadajemy jej sens empiryczny i staramy się szczegółowo zinterpretować znaczenie i wartość parametru identyfikującego rodzinę.

Kluczowym założeniem w metodzie Bayesa jest nadanie zbiorowi parametrów struktury probabilistycznej. Krótko mówiąc, przyjmujemy, że parametr jest zmienną losową (wektorem losowym). Rozpatrzmy przykład. Niech eksperyment badawczy polega na obserwacji 100 klientów supermarketu, z których każdy zakupi bądź nie zakupi produkt z określonej grupy produktów. Przypisujemy naszemu eksperymentowi rodzinę rozkładów dwumianowych z parametrem $\theta \in (0, 1)$. Parametr ten ma jasno określoną interpretację. Odczytujemy go jako szanse na to, że każdy klient zakupuje produkt z częstością θ. W zakresie szkoły średniej w takim przypadku mówimy o prawdopodobieństwie sukcesu. Zawsze jednak mamy fragmentaryczną lub cząstkową wiedzę na temat parametru. Często proces badawczy zaczyna się od obserwacji wstępnych czy badań pilotażowych, które w konsekwencji wnoszą pewną wiedzę na temat parametru odpowiadającego szansie zakupu. Wiedza taka zostaje potraktowana jako informacja wstępna i skondensowana w postaci rozkładu probabilistycznego, jakiemu podlega cecha losowa. Bardziej formalnie rozkład ten jest rozkładem zmiennej losowej (cechy losowej) Θ nazywanym rozkładem *a priori* (por. [8]).

Zaczynamy więc analizę eksperymentu od dwóch kluczowych założeń. W pierwszym założeniu przyjmujemy określoną rodzinę rozkładów $\mathcal{P}$ indeksowaną parametrem. W drugim założeniu precyzujemy rozkład parametru. Zarówno przyjęcie rodziny rozkładów, jak i rozkładu parametru są subiektywne i wypływają z profesjonalności i merytorycznej wiedzy badacza na temat eksperymentu. W wyborach tych jednak nie można wszystkiego maksymalnie upraszczać. Nie należy również niepotrzebnie popisywać się komplikacją wybranych rozkładów. Podręczniki opisujące podejście bayesowskie podają szereg standardów pozwalających na skuteczne opisanie klasycznych eksperymentów badawczych.

Po wypełnieniu dwóch wyjściowych postulatów, dalsze postępowanie jest automatyczne. Wynikiem estymacji uzyskanym na bazie wyników pomiarów jest rozkład parametru nazwany rozkładem *a posteriori*, który otrzymujemy jako rozkład warunkowy parametru (jako zmiennej losowej Θ) pod warunkiem realizacji próby losowej prostej.

Podkreślmy, że procedura bayesowska polega na tym, że rozkład *a priori* („nasza wiedza") podlega transformacji w rozkład *a posteriori* („wiedza modyfikowana pomiarami"), w której rolę moderującą odgrywa wyjściowa rodzina rozkładów $\mathcal{P}$ (patrz rysunek 2). Ponieważ wynikiem estymacji Bayesa jest rozkład probabilistyczny, więc we wnioskach badawczych analizujemy charakterystyki tego wynikowego rozkładu. Najczęściej jest to wartość oczekiwana[9], jednak ponieważ wynikiem jest konkretny jednoznaczny rozkład probabilistyczny, więc prezentowana jest po prostu gęstość tego rozkładu (histogram). Dodatkowo podajemy kwantyle lub charakterystyki rozrzutu.

W sensie teoretycznym procedura jest do zaakceptowania i wydaje się bardzo efektywna. Problemem jest proces obliczeniowy. Jak wiadomo znalezienie rozkładu warunkowego jest na ogół bardzo skomplikowane. Właściwie poza prostymi standardowymi przykładami (które zazwyczaj także wymagają dużej wiedzy z zakresu analizy matematycznej) znalezienie rozkładu *a posteriori*, a czasami tylko jego charakterystyk jest niemal niemożliwe. Z pomocą w tych trudnościach przyspieszyli specjaliści z zakresu klasycznego rachunku prawdopodobieństwa i technik komputerowych. Powstała cała teoria nazwana MCMC (Markov Chain Monte Carlo) pozwalająca na wykorzystanie symulacji komputerowych do otrzymania aproksymacji wynikowego rozkładu *a posteriori* (patrz na przykład [13]).

Na koniec zwróćmy uwagę na naturę filozoficzną koncepcji bayesowskiej. Jest to egzemplifikacja rozumowania Euklidesa, nazwanego dedukcją (patrz na przykład [6]). W podejściu Euklidesa przyjmujemy założenia i na bazie reguł logicznych wysnuwany konsekwencje w postaci wniosków. W procesie estymacji bayesowskiej założenia obejmują wyróżnienie rodziny rozkładów i rozkładu *a priori*, wnioskiem jest rozkład *a posteriori*.

[9] W wielu monografiach to właśnie tę wartość oczekiwaną (warunkową) nazywa się estymatorem bayesowskim.

4. PODSUMOWANIE

Jak widać estymacja jest procesem badawczym głęboko osadzonym w filozofii nauki. Spór filozoficzny na temat metodologii nauki trwa setki lat i będzie trwał nadal. Tak samo w estymacji nie powiemy, która metoda jest lepsza, która gorsza. W procedurze doboru sposobu estymacji decyduje sumienność naukowa badacza, profesjonalność w zakresie meritum zagadnienia i zakres wynikowych wnioskowań. Niektórzy badacze kierowani specyfiką przedmiotu badań czy osobistym nastawieniem będą preferować jedno czy drugie podejście, ale nie powinno to przeważać w wyborze metodologii estymacji.

LITERATURA

[1] Ajdukiewicz K., [1955], *Klasyfikacja rozumowań*, Studia Logica, II:278-300.
[2] Barra J.R., [1982], *Matematyczne podstawy statystyki*, Państwowe Wydawnictwo Naukowe, Warszawa.
[3] Bickel P.J., Doksum K.A., [2001], *Mathematical Statistics. Basic Ideas and Selected Topics*, Vol. I, Prentice Hall, Upper Saddle River, New Jersey, 07458, wydanie II.
[4] Brillinger D.R., Tukey J.W., [2002], *His life and professional contributions*, The Annals of Statistics, 30(6):1535-1575.
[5] Gelman A., Carlin J.B., Stern H.S., Rubin D.B., [2000], *Bayesian Data Analysis*, CHAPMAN & HALL/CRC, wydanie II.
[6] Herbut J., (redaktor), [1997], *Leksykon filozofii klasycznej*, Katolicki Uniwersytet Lubelski, hasło DEDUKCJA – autor hasła Andrzej Bronk, s. 101-102.
[7] Krzykowski G., [2000], *Analiza wyników badań marketingowych. Metody symulacyjne*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.
[8] Krzykowski G., [2010], *Pomiar i statystyczna analiza cech binarnych*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.
[9] Lehmann E.L., [1991], *Teoria estymacji punktowej*, Państwowe Wydawnictwo Naukowe, Warszawa.
[10] Murphy J.J., [1999], *Analiza techniczna rynków finansowych*, WIG-PRESS.
[11] Petras I., Podlubny I., [2010], Least Squares or Least Circles? A comparsion of classical regression and orthogonal regression, Chance, 23(2):38-42.
[12] Reale G., *Historia filozofii starożytnej*, Katolicki Uniwersytet Lubelski, 2000-2002.
[13] Robert C.P., Casella G., [2004], *Monte Carlo Statistical Methods*, Springer-Verlag, wydanie II.
[14] Stigler S.M., [1981], *Gauss and invention of last squares*, The Annals of Statistics, 9(3):465-474.
[15] Tukey J.W., [1977], *Exploratory Data Analysis*, Addison-Wesley.
[16] Woleński J., [1985], *Filozoficzna szkoła lwowsko-warszawska*, Państwowe Wydawnictwo Naukowe, Warszawa.