ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1976

Seria: AUTOMATYKA z. 36 Nr kol. 481

Ernest Czogała

PRZEDMIOT BADAŃ STATYSTYCZNYCH W ŚWIETLE AKSJOMATYCZNEGO MODELU RACHUNKU PRAWDOPODOBIEŃSTWA

Streszczenie. W pracy przedstawiono podstawowe pojęcia dotyczące przedmiotu badań statystycznych w oparciu o aksjomatyczny model rachunku prawdopodobieństwa.Wskazano istniejące analogie między przedstawionymi pojęciami oraz niedostatki pojęć statystyki matematycznej na tle aksjomatycznego modelu rachunku prawdopodobieństwa.

1. Wstęp

Współczesna produkcja wymaga systematycznego kontrolowania jakości jej wyrobów. O właściwościach tych wyrobów wnioskuje się w oparciu o metody badań, których dostarcza samodzielna dyscyplina matematyki zwana statystyką matematyczną. Dyscyplina ta, ogólnie rzecz biorąc, zajmuje się metodami wnioskowania o pewnych zbiorach na podstawie badania podzbiorów tych zbiorów. Statystyka matematyczna posługuje się rachunkiem prawdopodobieństwa przy wnioskowaniu opartym na właściwych tylko jej zasadach,tym niemniej stanowi dyscyplinę odrębną od rachunku prawdopodobieństwa, mającą odrębny przedmiot badań. Tok postępowania w badaniach statystycznych można przedstawić schematycznie jak na rys. 1.

Rys. 1. Tok postępowania w badaniach statystycznych

W dalszych rozważaniach interesować nas będzie przedmiot badań statystycznych. Przedmiotem badań może być jeden element (jeden egzemplarz,jedna sztuka) bądź też zbiór elementów (partia egzemplarzy lub wyrobów).

Jeśli chodzi o zbiór elementów będący przedmiotem badań statystycznych to będziemy zakładać, że

1) wszystkie elementy są badane w taki sam sposób,
2) istnieje niezamierzony rozrzut wartości badanej właściwości, nazywanej cechą dla poszczególnych elementów w jednym zbiorze.

Całokształt badania statystycznego ustala tzw. program badań, w którym podaje się

a) opis badanego zbioru elementów,
b) przepis i sposób badania,
c) analizę wyników i sposób wnioskowania.

W zagadnieniach statystycznych, w których bada się prawidłowości pewnych zjawisk przez obserwowanie ich częstości (matematycznym odpowiednikiem pojęcia częstości w rachunku prawdopodobieństwa jest aksjomatyczne pojęcie prawdopodobieństwa) rachunek prawdopodobieństwa gra również dużą rolę. Z uwagi na fakt, że podstawowe pojęcia statystyki powstały znacznie wcześniej od aksjomatycznego modelu rachunku prawdopodobieństwa pochodzącego od KOŁMOGOROWA (1933 r.), konfrontacja pojęć obydwu dziedzin wydaje się celowa.

2. Podstawowe pojęcia dotyczące przedmiotu badań statystycznych w oparciu o aksjomatyczny model rachunku prawdopodobieństwa

W badaniach statystycznych przyjęto nazywać zbiór, którego elementy obserwujemy, populacją generalną lub zbiorowością generalną. Zbiór ten będziemy nazywać przedmiotem badań statystycznych. Odpowiednikiem tego zbioru w aksjomatycznym modelu rachunku prawdopodobieństwa jest przestrzeń zdarzeń elementarnych Ω. Jak wiadomo, przestrzeń zdarzeń elementarnych jest pojęciem pierwotnym w aksjomatycznym ujęciu prawdopodobieństwa. Zbadanie wszystkich elementów tego zbioru (badanie wyczerpujące) może być niewykonalne z następujących przyczyn:

1) zbyt liczna populacja generalna, tak że zbadanie wszystkich elementów byłoby zbyt kosztowne lub za długo by trwało,
2) cała populacja generalna może być niedostępna,
3) zbyt szybko wymagane wyniki badania,
4) badania elementów powodują ich zniszczenie.

W praktyce więc nie badamy całego zbioru Ω, a tylko jego część A, będącą podzbiorem Ω ($A \subset \Omega$). Podzbiór ten nazywamy <u>próbką</u> (próbą). Na podstawie wyników badania tej próbki wyciągamy wnioski co do reprezentowanego przez nią zbioru Ω. W celu lepszej reprezentatywności próbki, pobieramy poszczególne elementy w sposób losowy, nazywając ją <u>próbką losową.</u>

W aksjomatycznym modelu rachunku prawdopodobieństwa rozpatruje się pewną klasę (rodzinę) podzbiorów przestrzeni Ω zwaną σ-ciałem $\mathcal{B}$. Klasa $\mathcal{B}$ podzbiorów przestrzeni Ω, zwana również borelowskim σ- ciałem, jak wiadomo, czyni zadość następującym warunkom:

1) $\bigvee_{A \subset \Omega} A \in \mathcal{B}$ (niepustość)

2) $A \in \mathcal{B} \Rightarrow \bar{A} = \Omega \setminus A \in \mathcal{B}$ (komplementatywność)

3) $A_1, A_2, \ldots \in \mathcal{B} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$ (przeliczalna addytywność)

Odpowiednikiem abstrakcyjnego pojęcia prawdopodobieństwa w aksjomatycznym modelu rachunku prawdopodobieństwa jest w statystyce matematycznej częstość względna (von MISESa)

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

Należy tutaj zaznaczyć, że to określenie prawdopodobieństwa ze względu na jego intuicyjny charakter ma ograniczony zakres stosowalności i jest niezadowalające z punktu widzenia logicznej konstrukcji aksjomatycznego modelu rachunku prawdopodobieństwa, który to model postuluje następujące aksjomaty prawdopodobieństwa:

I $\bigwedge_{A \in \mathcal{B}} \; 0 \leqslant P(A) < 1$

II $P(\Omega) = 1$

III $A_1, \ldots, A_n, \ldots \in \mathcal{B}$, $A_i \cap A_j \underset{i \neq j}{=} \phi \Rightarrow P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Funkcja prawdopodobieństwa jest miarą unormowaną w przestrzeni zdarzeń elementarnych Ω, określoną na zbiorach σ- ciała $\mathcal{B}$. Przestrzeń zdarzeń elementarnych Ω, borelowską rodzinę jej podzbiorów $\mathcal{B}$ wraz z miarą unormowaną P, nazywamy przestrzenią probabilistyczną $(\Omega, \mathcal{B}, P)$.

Elementy populacji generalnej, a więc przedmiotu badań statystycznych posiadają różne właściwości zwane cechami. Rozróżniamy dwa zasadnicze rodzaje cech: skokowe (dyskretne) i ciągłe. W statystyce wprowadza się podział na cechy mierzalne (pozwalające się zmierzyć przyrządem pomiarowym) i niemierzalne. Cechą niemierzalną może być cecha skokowa o dwuelementowym zbiorze wartości.

Z punktu widzenia aksjomatycznego modelu rachunku prawdopodobieństwa określenie cechy odpowiada określeniu zmiennej losowej. Określenie to brzmi: zmienną losową $X(\omega)$ nazywamy funkcję rzeczywistą określoną na zbiorze Ω taką, że dla każdej liczby rzeczywistej $x \in R_X$ (R_X - oznaczamy zbiór liczb rzeczywistych, z którego zmienna losowa $X(\omega)$ może przybierać wartości) zbiór zdarzeń elementarnych, dla których $X(\omega) < x$ jest zdarzeniem losowym należącym do σ-ciała $\mathcal{B}$ czyli:

$$\bigwedge_{x \in R_X} \left\{ \omega : X(\omega) < x \right\} \in \mathcal{B} .$$

W statystyce dokonuje się podziału zbiorz R_X możliwych wartości danej cechy na rozłączne niepuste podzbiory zwane klasami. Dokonuje się w ten sposób tzw. klasyfikacji elementów zbioru R_X. Klasyfikacja elementów możliwych wartości stanowi podstawę klasyfikacji elementów rozważanego zbioru Ω na rozłączne podzbiory (klasy). Mówimy że klasyfikacja poszczególnych elementów w rozważanym zbiorze jest obrazem dokonanej klasyfikacji zbioru możliwych wyników. Liczba klas w zbiorze elementów i w zbiorze wyników jest taka sama i nie większa niż liczba możliwych różnych wyników. W oparciu o pojęcia częstości względnej można określić prawdopodobieństwa uzyskania klas odpowiednich elementów.

W aksjomatycznym modelu rachunku prawdopodobieństwa zmienna losowa $X(\omega)$ indukuje w naturalny sposób w przestrzeni swoich wartości R_X nowe σ-ciało $\mathcal{B}_X$ złożone z takich zbiorów $S \in R_X$, których przeciwzbiory (zdarzenia losowe) są elementami σ-ciała $\mathcal{B}$.

Na elementach tego nowego σ-ciała $\mathcal{B}_X$ zmienna losowa $X(\omega)$ indukuje również prawdopodobieństwo P_X, określone następująco:

$$\bigwedge_{S \in \mathcal{B}_X} P_X(S) = P(A).$$

W ten sposób zmienna losowa $X(\omega)$ indukuje w przestrzeni wartości nową przestrzeń probabilistyczną indukowaną $(R_X, \mathcal{B}_X, P_X)$. Można zauważyć, że istnieje pewna analogia między klasyfikacją w statystyce, a indukowaniem przestrzeni probabilistycznej w rachunku prawdopodobieństwa,

W statystyce często korzysta się z bardzo wygodnego modelu tzw. <u>probabilistycznej przestrzeni próbek</u> $\langle \Omega, \mathcal{B}, P \rangle$, przyjmując wprost że Ω jest prostą rzeczywistą, $\mathcal{B}$ jest σ-ciałem zbiorów borelowskich, natomiast $P_X(S) = P\{\omega : X(\omega) \in S\}$ dla każdego $S \in \mathcal{B}$. Zmienną losową przyjmujemy wtedy jako samą wartość zdarzenia elementarnego tj. $X(\omega) = \omega$, gdzie ω jest punktem przestrzeni R_X.

Ostatnim pojęciem, które podamy w niniejszej pracy jest pojęcie statystyki. Słowo "statystyka" należy do najbardziej wieloznacznych. Wiąże się to z ewolucją tego pojęcia. Często <u>statystyką</u> nazywamy zmienną losową będącą funkcją wielowymiarowej zmiennej losowej. Takie określenie statystyki jako funkcji wielowymiarowej zmiennej losowej jest jednoznaczne i staje się zrozumiałe na gruncie aksjomatycznego modelu rachunku prawdopodobieństwa.

3. <u>Uwagi końcowe</u>

Powszechnie zaakceptowanym modelem rachunku prawdopodobieństwa jest model aksjomatyczny KOŁMOGOROWA jako najbardziej zadowalający z teoretycznego punktu widzenia. Dokonując przeglądu podstawowych pojęć dotyczących przedmiotu badań statystycznych na tle aksjomatycznego modelu rachunku prawdopodobieństwa obserwuje się analogie w pojęciach obydwu dziedzin, Można jednak zauważyć, że niektóre pojęcia statystyki, jako że powstały wcześniej, nie są zbyt precyzyjne i zawierają wieloznaczności. Przykładem tego może być określenie prawdopodobieństwa oparte na częstości względnej termin "statystyka" jak również pojęcie klasy.Zbliżenie pojęć obydwu dziedzin pozwala uświadomić sobie pewne niedostatki tych pojęć statystyki matematycznej oraz może wskazać możliwości ich eliminacji na gruncie pojęć aksjomatycznego modelu rachunku prawdopodobieństwa. Warto o tym przypomnieć choćby z tego powodu, że rola statystyki matematycznej jako dyscypliny naukowej wciąż wzrasta obejmując coraz więcej dziedzin działalności praktycznej człowieka.

LITERATURA

[1] M. Fisz: Rachunek prawdopodobieństwa i statystyka matematyczna, PWN Warszawa 1972.

[2] M. Warmus: Wykłady z probabilistyki, Tom I, Tom II, PWN Warszawa 1971.

[3] Sz. Firkowicz: Statystyczne badanie wyrobów, WNT Warszawa 1970.

[4] Sz. Firkowicz: O metodach statystycznych w przemyśle, PAK 1974/10.

[5] K. Mańczak: Metody identyfikacji wielowymiarowych obiektów sterowania WNT Warszawa 1970.

ПРЕДМЕТ СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЙ
НА ФОНЕ АКСИОМАТИЧЕСКОЙ ТЕОРИИ ВЕРОЯТНОСТИ

Р е з ю м е

В работе даны основные понятия, касающиеся предмета статистических исследований на фоне аксиоматической теории вероятности. Указаны существующие аналогии между представленными понятиями, а также недостатки понятий математической статистики имея ввиду аксиоматическую теорию вероятности.

SUBJECT OF STATISTICAL INVESTIGATIONS BASED ON THE AXIOMATIC PROBABILITY THEORY

S u m m a r y

In the paper the essential ideas on a subject of statistical investigations based on the axiomatic probability theory have been presented. There have been indicated the analogies existing between the presented ideas and the defficiency of ideas of mathematical statistics against the background of the axiomatic probability theory.