ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 227, 2009

*Jan Kordos**

Z DOŚWIADCZEŃ NAUCZANIA STATYSTYKI MATEMATYCZNEJ I WNIOSKOWANIA STATYSTYCZNEGO NA UCZELNIACH EKONOMICZNYCH

1. WSTĘP

W ciągu ponad 50-letniej mojej pracy naukowo-badawczej i dydaktycznej w statystyce w kraju i zagranicą wiele czasu poświęcałem dydaktyce, tj. szkoleniom statystycznym pracowników statystyki oficjalnej oraz wykładom i ćwiczeniom z różnych dziedzin statystyki na kilku uczelniach. Brałem udział w intensywnych szkoleniach z metod statystycznych prowadzonych przez Komisję Matematyczną GUS w latach sześćdziesiątych i siedemdziesiątych ubiegłego wieku. W Szkole Głównej Handlowej (dawnej SGPiS) prowadziłem wykłady z metody reprezentacyjnej w latach 1971–1974. W czasie pracy w Etiopii w latach 1974–1980, oprócz szkolenia personelu statystycznego, prowadziłem wykłady z metod statystycznych na Uniwersytecie w Addis Abebie. Po powrocie do kraju w końcu 1980 r., oprócz pracy w GUS, podjąłem współpracę z Instytutem Gospodarstwa Społecznego SGH, którą kontynuowałem do 1992 r. W czasie pełnienia funkcji wiceprezesa GUS w latach 1992–1996 przerwałem współpracę, którą podjąłem na początku 1996 r. w Instytucie Statystyki i Demografii SGH. Prowadziłem tam wykłady do lipca 2007 r. z projektowania badań społecznych, statystyki społecznej i metod ilościowych. Wykłady i ćwiczenia ze statystyki matematycznej lub wnioskowania statystycznego oraz ze statystyki z elementami wnioskowania statystycznego prowadziłem w ostatnich latach na dwóch uczelniach ekonomicznych w ciągu ponad 20 semestrów.

Warto wspomnieć, że problemy nauczania statystyki i szkoleń statystycznych były przedmiotem kilku spotkań zorganizowanych przez Międzynarodowe Stowarzyszenie Edukacji Statystycznej (*International Association for Statistical Education* – w skrócie IASE). Chodziło tam o edukację społeczeństwa, szkolenie

* Prof. dr hab., Wyższa Szkoła Menedżerska w Warszawie.

statystyków i użytkowników statystyki oraz nauczanie statystyki na różnych poziomach edukacji. IASE wydaje periodyk oraz różne opracowania poświęcone edukacji statystycznej. W mojej pracy dydaktycznej korzystałem także z doświadczeń IASE (Kordos, 1999, s. 45–47).

Znaczne doświadczenie w szkoleniu statystyków w urzędach statystycznych zdobył Eurostat prowadząc znany program TES („*Training of European Statisticians*" – Szkolenie Statystyków Europejskich). Eurostat dąży do permanentnego szkolenia statystyków europejskich, a pewne prace z tego zakresu prowadzono również w Polsce (Kordos, 2000, s. 50–62; Kordos i Platek, 1999).

Postanowiłem jednak w tej prezentacji skoncentrować się tylko na kilku zagadnieniach ze statystyki matematycznej i wnioskowania statystycznego (tak nazywano te wykłady, chociaż syllabusy były podobne), które sprawiały mi pewne kłopoty w dydaktyce, a także przedstawić próby ich rozwiązania. Wspomnę również o pewnych nieścisłościach interpretacyjnych oraz terminologicznych napotkanych w czasie mojej pracy.

2. DLACZEGO NALEŻY LOSOWAĆ PRÓBĘ, ABY MOŻNA BYŁO UOGÓLNIĆ UZYSKANE WYNIKI NA CAŁĄ POPULACJĘ?

W pierwszej kolejności przedstawię trudności, na jakie napotkałem przy wyjaśnieniu studentom konieczności losowania próby z populacji, gdy chcemy uogólnić uzyskane wyniki z próby na zbiorowość generalną. Dlaczego próba musi być losowana, a nie wybrana w inny sposób? Z początku wydawało się to dość proste, bo przecież wynika to z *prawa wielkich liczb*. Po wprowadzeniu podstawowych pojęć i definicji z elementów rachunku prawdopodobieństwa przedstawiałem *prawo wielkich liczb* oraz *centralne twierdzenie graniczne*.

Zakładałem, że zrozumienie istoty *statystyki matematycznej* wymaga zapoznania się z niektórymi osiągnięciami rachunku prawdopodobieństwa. Część studentów nie miała przed tym wykładów ze statystyki, więc przypominałem podstawy statystyki opisowej, aby później przejść do podstawowych pojęć jak *zdarzenie losowe, prawdopodobieństwo zajścia zdarzenia losowego* (trzy definicje prawdopodobieństwa), *zmienna losowa, rozkład zmiennej losowej, dystrybuanta rozkładu*. Szczególnie dużo czasu poświęcałem *rozkładowi normalnemu*, ze względu na jego podstawowe znaczenie w statystyce, a także *rozkładowi t-Studenta, rozkładowi dwumianowemu* oraz *rozkładowi chi-kwadrat*.

Przedstawiałem *twierdzenia graniczne*, których szczególnym przypadkiem są twierdzenia znane najczęściej pod nazwą *prawa wielkich liczb*. W zasadzie na studiach ekonomicznych nie prowadziłem dowodów twierdzeń. Z twierdzeń granicznych wyróżniałem przede wszystkim twierdzenie zwane *centralnym twierdzeniem granicznym*, które jest ściśle związane z rozkładem normalnym.

Uważałem, że twierdzenia te wyjaśniają sens losowania próby z populacji generalnej oraz uogólnienia wyników uzyskanych z próby na całą populację.

Przekonałem się jednak, że takie formalne przedstawienie przedmiotu nie jest dla studentów zbyt jasne. Podawanie warunku, iż liczebność próby dąży do nieskończoności ($n \to \infty$), a w rzeczywistości mamy zawsze próby skończone. Dlatego postanowiłem dokonać pewnego rodzaju analizy słabego prawa wielkich liczb, aby krok po kroku pokazać studentom, w jaki sposób przechodzi się od teorii do praktyki.

W praktyce, jak wiadomo, mamy populację składającą się z N jednostek, w której badamy cechę X, przyjmującą dla jednostki i wartość x_i ($i = 1, 2, \ldots, N$). Dla tej populacji chcemy oszacować nieznaną wielkość średnią, tj. μ, na podstawie próby liczącej n jednostek. Ale jak tego dokonać, aby uzyskane wyniki uogólnić na całą populację?

Analizuję sposób przekazania studentom przejścia z teorii do praktyki, na przykładzie słabego prawa wielkich liczb, z którego wynika konieczność wykorzystania tablicy liczb losowych do wybrania próby umożliwiającej uogólnienie uzyskanych wyników na całą populacją ze średnim błędem szacunku ε.

Podam tu *słabe prawo wielkich liczb*, nieco zmodyfikowane, według S. Zubrzyckiego (Zubrzycki, 1966, s. 139):

Jeśli $X_1, X_2, \ldots, X_n$ *jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie prawdopodobieństwa z wartością oczekiwaną* μ *i odchyleniem standardowym* σ, *a* $\overline{X}_n = \frac{\sum_{i=1} X_i}{n}$ *oznacza średnią pierwszych n zmiennych, to dla każdego* $\varepsilon > 0$ *i* $\eta > 0$ *zachodzi nierówność*

$$P(|\overline{x}_n - \mu| < \varepsilon) > 1 - \eta$$

gdy n jest dostatecznie duże.

W pierwszej kolejności próbuję wyjaśnić studentom, dlaczego próba powinna być wybrana przy pomocy tablic liczb losowych. Pierwsza część słabego prawa wielkich liczb głosi:

„*Jeśli* $X_1, X_2, \ldots, X_n$ *jest ciągiem <u>niezależnych</u> zmiennych losowych o jednakowym rozkładzie prawdopodobieństwa z wartością oczekiwaną* μ *i odchyleniem standardowym* σ*"*. Podkreśliłem słowo „niezależnych" zmiennych losowych, gdyż w praktyce, aby spełnić ten warunek teoretyczny, statystycy zbudowali „*tablice liczb losowych"*. Przy pomocy tablic liczb losowych wybieramy najpierw x_1 i traktujemy jako realizację zmiennej losowej X_1, a następnie x_2 – jako realizację zmiennej losowej X_2, itd. Próba $x_1, x_2, \ldots, x_n$ jest traktowana jako pojedyncza realizacja n-wymiarowej zmiennej losowej (X_1, X_2, ..., X_n). Tak więc, jednostki badania wybierane są do próby przy pomocy *tablic liczb losowych*, aby spełnić warunek niezależności.

Druga część twierdzenia słabego prawa wielkich liczb brzmi: „*wtedy dla każdego ε>0 i 0< η<1 zachodzi* $P\{|\overline{x}_n - \mu| < \varepsilon\} > 1-\eta$, *gdy n jest dostatecznie duże*". W praktyce mamy $P\{|\overline{x}_n - \mu| < \varepsilon\} = 1-\alpha$, gdzie ε oznacza precyzję oceny, a 1 – α poziom ufności. Stwierdzam następnie, że zagadnienia te będziemy omawiali szerzej przy prezentacji metod estymacji. Wydaje mi się, że Neyman wykorzystał tę nierówność $P\{|\overline{x}_n - \mu| < \varepsilon\} > 1-\eta$ i zastąpił ją $P\{|\overline{x}_n - \mu| < \varepsilon\} = 1-\alpha$. Ale na ten temat ze studentami nie dyskutuję.

3. ZASADA PRAKTYCZNEJ PEWNOŚCI

Dlaczego w praktyce przyjmujemy oceny z badań reprezentacyjnych oraz odrzucamy hipotezę zerową? Jakie są podstawy teoretyczne takiego postępowania? Podstawą stosowania prawa wielkich liczb w praktyce statystycznej jest pewna reguła postępowania praktycznego, którą określamy mianem ***zasady praktycznej pewności***. Zasadę tę spotkałem tylko w podręczniku W. Sadowskiego (Sadowski, 1965).

Zasada ta opiera się na empirycznym stwierdzeniu, że zdarzenia losowe, których prawdopodobieństwo jest bardzo małe, zachodzą bardzo rzadko. Jeśli prawdopodobieństwo pewnego zjawiska losowego jest tak małe, że zachodzi ono niezmiernie rzadko, to w praktyce można nie liczyć się z możliwością zajścia takiego zjawiska. Wówczas powiadamy, że istnieje praktyczna pewność, że zjawisko nie zajdzie. Na podstawie zasady praktycznej pewności oczekujemy, że błąd szacunku podlegającemu prawu wielkich liczb jest niewielki, jeśli liczba obserwacji jest dostatecznie duża. Hipotezę zerową odrzucamy, gdy prawdopodobieństwo popełnienia błędu pierwszego rodzaju jest odpowiednio małe.

Zasada praktycznej pewności, jako podstawa stosowania w praktyce prawa wielkich liczb, dopuszcza wyjątki, wprawdzie niezmiernie rzadkie, ale możliwe. Mogą więc zajść zdarzenia losowe, co do których istnieje praktyczna pewność, że nie zachodzą.

Podkreślałem, że teoria rozkładów granicznych stanowi jeden z większych i ważniejszych działów rachunku prawdopodobieństwa, który rozwinął się szczególnie szybko na przestrzeni ostatnich pięćdziesięciu lat, jakkolwiek niektóre twierdzenia z tego zakresu zostały udowodnione jeszcze w XVIII i XIX wieku. Jako pionierów prac badawczych w tej dziedzinie wymieniałem A. de Moivre'a, C.F. Gaussa i A. Laplace'a, odkrywców rozkładu normalnego, który w twierdzeniach granicznych odgrywa podstawową rolę. Dalsze prace szeregu wybitnych matematyków i statystyków, a przede wszystkim pracę Levy'ego, Lindeberga, Lapunowa, Fellera i Chinczyna, doprowadziły do pełnego rozwiązania tego zagadnienia. Formułowano i udowodniono kilka ważnych twierdzeń, podając warunki, które muszą być spełnione, aby zmienna losowa miała rozkład normalny lub zbieżny do rozkładu normalnego. Podawałem tylko *twierdzenie Lindeberga-Lévy'ego* dotyczące centralnego twierdzenia granicznego.

4. ESTYMACJA PUNKTOWA CZY PRZEDZIAŁOWA?

Zgodnie z programem, przedstawiam zarówno estymację punktową jak i przedziałową. Dla studentów, a także dla użytkowników danych, bardziej zrozumiała jest estymacja punktowa ze średnim błędem standardowym. Estymacja przedziałowa jest często tak interpretowana, jak gdyby parametr był zmienną losową. Podawane są interpretację, że „*parametr leży w przedziale*", a przecież zbudowany przedział jest realizacją zmiennej losowej i można go tylko zinterpretować, iż „*pokrywa wartość parametru*" przy określonym poziomie ufności. Ciekawa interpretacja przedziału ufności jest podana w książce S. M. Kota i in. (2007) z powołaniem się na prof. Z. Hellwiga z chusteczką. Studenci rozumieją taką interpretację.

Neyman (1933) wprowadził termin „*półprzedział ufności*", a chodzi tu o wielkość $d = t_{1-\frac{\alpha}{2}} s(\bar{x})$ lub względny półprzedział ufności, tj. $\Delta = t_{1-\frac{\alpha}{2}} v(\bar{x})$, gdzie $v(\bar{x}) = \frac{s(\bar{x})}{x}$. Wydaje mi się, że w dydaktyce termin ten jest przydatny i wprowadziłem go na wykładach. W niektórych podręcznikach i w publikacjach statystycznych spotkałem się z interpretacją, że ocena ma niewielką wartość, gdy względny błąd standardowy jest większy niż 5%, a gdy jego wartość jest większa niż 10%, to ocena całkowicie jest dyskwalifikowana. Takie kategoryczne stwierdzenie jest nieuzasadnione, gdyż zależy to od celu badania, znaczenia poszczególnych szacowanych parametrów, kosztów oraz ważności domen, których oceny dotyczą. Na ten temat warto podyskutować, gdyż występują nieporozumienia.

5. RÓŻNICE TERMINOLOGICZNE

Zauważyłem, że w różnych podręcznikach statystycznych nie jest wykorzystywana taka sama terminologia. W niektórych podręcznikach $1 - \alpha$ nazywany jest współczynnikiem ufności, a w innych poziomem ufności. Neyman (1933) współczynnikiem ufności nazwał α. Na wykładach podaję, że $1 - \alpha$ powinno się nazywać *poziomem ufności*, jak to jest podawane w „*Tablicach statystycznych*" R. Zielińskiego (Zieliński, 1982), ale wspominam o różnicach terminologicznych.

Na przykład, używa się terminu „*próba*" i „*próbka*", ale chodzi tu o to samo pojęcie. Taki przykładów można znaleźć więcej.

6. EFEKTYWNOŚĆ NAUCZANIA

Po każdym wykładzie dokonywałem pewnego rodzaju analizy procesu dydaktycznego. Na podstawie takiej analizy można było ustalić, od jakich czynników zależy efektywność nauczania. Zwykle prowadziłem wykłady i ćwiczenia na zaocznych studiach magisterskich. Najlepsze wyniki można było uzyskać w tych grupach, gdy wykładowca prowadził zarówno wykład jak i ćwiczenia. Najsłabsze wyniki uzyskiwałem, gdy prowadzący ćwiczenia zmieniali się w każdym semestrze. Efektywność nauczania zależała również od tego, czy wykładowca dał studentom dodatkowe materiały przygotowane do wykładów. Niejednokrotnie dawałem studentom tzw. slajdy rozszerzone, które wykorzystywałem na wykładach, ale zawierające uzupełniające informacje, które nie byłem w stanie omówić szerzej na wykładzie. Dodatni wpływ miały także dodatkowe konsultacje poza wykładami, aby wyjaśnić trudniejsze problemy. Zwykle przychodzili na nie studenci „dociekliwi", ale zauważyłem, iż przekazywali swoją wiedzę słabszym kolegom.

7. UWAGI KOŃCOWE

Podałem tu tylko wybrane uwagi odnośnie problemów napotkanych w procesie nauczania statystyki matematycznej lub wnioskowania statystycznego na studiach ekonomicznych. Podobne problemy występowały także na innych wykładach z metod statystycznych. Zdaję sobie sprawę, że każdy dydaktyk ma indywidualne podejście do poruszanych zagadnień. Zastanawia mnie, czy prawidłowo przekazuję wiedzę studentom i w jaki sposób ulepszyć proces dydaktyczny w następnym semestrze. Interesuje mnie jak tego dokonują statystycy w innych krajach.

Zagadnieniami oceny nauczania statystyki na wyższych uczelniach oraz specjalistycznego szkolenia z zakresu statystyki w urzędach statystycznych na świecie zajmują się różne instytucje. Przy Międzynarodowym Instytucie Statystycznym działa od 1988 r. międzynarodowe stowarzyszenie zajmujące się edukacją statystyczną – *International Association for Statistical Education*, o którym wspomniałem na początku. Stowarzyszenie to dużo czasu poświęca problemom nauczania statystyki na różnych poziomach i metodom jego oceny. Wydaje się celowe, aby nasi statystycy zajmujący się nauczaniem statystyki oraz zawodowym szkoleniem statystyków zapoznawali się z pracami tego stowarzyszenia międzynarodowego. Warto zwrócić uwagę na niektóre zagadnienia związane z nauczaniem statystyki na wyższych uczelniach.

LITERATURA

Fisz M. (1969), *Rachunek prawdopodobieństwa i statystyka matematyczna*, WNT, Warszawa.

Kordos J. (1999), *Prace badawcze w zakresie edukacji statystycznej*, Kwartalnik Statystyczny, nr 4.

Kordos J. (2000), *Instytut TES – Szkolenie statystyków europejskich*, Kwartalnik Statystyczny, nr 6.

Kordos J. (2000), *Permanentne szkolenie statystyków,* Wiadomości Statystyczne, nr 5.

Kordos J., Platek R. (1999), *Zarys programu szkolenia statystycznego pracowników Głównego Urzędu Statystycznego*, GUS, Warszawa, grudzień (maszynopis).

Kot S., Jakubowski J., Sokołowski A. (2007), *Statystyka, Podręcznik dla studiów ekonomicznych.* Difin, Warszawa.

Luszniewicz A. (1987), *Metody wnioskowania statystycznego*, PWE, Warszawa.

Neyman J. (1933), *Zarys teorii i praktyki badania struktury ludności metodą reprezentacyjną*, Instytut Gospodarstwa Społecznego, Warszawa.

Pawłowski Z. (1980), *Statystyka matematyczna*, PWN, Warszawa.

Sadowski W. (1965), *Statystyka matematyczna*, PWE, Warszawa.

Zieliński R. (1982), *Tablice statystyczne*. PWN, Warszawa.

Zubrzycki S. (1966), *Wykłady z rachunku prawdopodobieństwa i statystyki matematycznej*, PWN, Warszawa.

Jan Kordos

PRACTICAL EXPERIENCE OF TEACHING MATHEMATICAL STATISTICS AND STATISTICAL INFERENCE AT ECONOMIC UNIVERSITIES

The author begins by mentioning his lectures on statistical methods he has given in Poland and abroad during the last fifty years, focusing on lectures and exercises in mathematical statistics and statistical inference presented at economic universities. He presents difficulties he has encountered while teaching certain topics, particularly *law of large numbers* and *central limit theorem.* He is analyzing the way of showing students how to get from theory to practice on an example of weak law of large numbers which implies using a table of random numbers for sample selection to get estimates for the whole population with the standard errors. Next he presents advantages of taking into account the *principle of practical certainty* to explain while we reject null hypothesis and accept in practice results of sample surveys knowing their precision. He also discusses some terminology differences he has encountered in some Polish handbooks on mathematical statistics. Some concluding remarks are given at the end.