ZESZYTY NAUKOWE WYŻSZEJ SZKOŁY PEDAGOGICZNEJ W BYDGOSZCZY
Problemy Matematyczne 1980 z. 2

Lech T. Kubik
Uniwersytet Warszawski

O NAUCZANIU RACHUNKU PRAWDOPODOBIENSTWA NA KIERUNKU NAUCZYCIELSKIM STUDIÓW MATEMATYCZNYCH

Ucząc jakiegokolwiek przedmiotu należy zdawać sobie sprawę z tego, jaki jest cel nauczania. Otóż wydaje się, że celem nauczania rachunku prawdopodobieństwa na każdym poziomie i na każdej specjalności powinno być:

1/ przedstawienie rachunku prawdopodobieństwa jako ścisłej teorii matematycznej,

2/ wskazanie na związek tej teorii z doświadczeniem i na możliwości jej zastosowań.

Opinię tę wyraża m. in. A. Rényi w artykule [12]. W zależności od potrzeby przy wyborze treści i metod nauczania można uznać za ważniejszy cel 1/ lub 2/, ale nigdy nie powinno się ignorować żadnego z nich. Na kierunku nauczycielskim trzeba uwzględnić jeszcze trzeci cel nauczania rachunku prawdopodobieństwa:

3/ Przygotowanie studenta do nauczania tego przedmiotu w szkole.

W tym referacie chciałbym przedstawić, jak moim zdaniem należy dążyć do osiągnięcia powyższych celów. Część uwag, jakie tu podam, będzie powtórzeniem uwag zawartych w moich artykułach opublikowanych w latach 1973-1976 w czasopiśmie dla nauczycieli "Matematyka" oraz w mojej książce [8], a część - będą to uwagi nowe.

Otóż punktem wyjścia wykładu powinny być doświadczenia losowe odznaczające się stabilnością częstości. Tego rodzaju doświadczenia powinny towarzyszyć całemu wykładowi rachunku prawdopodobieństwa. Przestrzeń probabilistyczna $/\Omega, M, P/$ powinna być traktowana jako model matematyczny pewnego doświadczenia losowego. Każde pojęcie w modelu powinno być inspirowane przez rozważania dotyczące doświadczeń losowych. Rozważania dotyczące doświadczeń i ich modeli powinny się przeplatać, ale powinny być starannie od siebie oddzielone. W przypadku przeciwnym następuje pomieszanie elementów teoretycznych i empirycznych, co może prowadzić do zniechęcenia do rachunku prawdopodobieństwa, a w konsekwencji do zniechęcenia do matematyki, lub do stwierdzenia, że rachu- w której wszystko jest mniej ścisłe i mniej pewne. Niestety, wiele książek z rachunku prawdopodobieństwa może myślącego czytelnika zniechęcić do tego przedmiotu. W książkach tych oprócz pomieszania elementów teoretycznych i empirycznych często zdarzają się rozbieżności między definicjami pojęć, a późniejszym stosowaniem tych pojęć. Konkretne przykłady takich sytuacji będą kilkakrotnie wskazane w tym referacie.

Rachunek prawdopodobieństwa stanowi, oczywiście, równie ścisły dział matematyki, jak każdy inny. Jest to jak wiadomo teoria aksjomatyczna, w której dowodzi się twierdzeń na drodze dedukcji. Twierdzenia te są prawdziwe w każdym modelu probabilistycznym. Nie dotyczą one jednak samych doświadczeń losowych. Jeżeli chcemy rozwiązać jakieś zagadnienie dotyczące konkretnego doświadczenia losowego, to musimy wykonać kolejno trzy następujące etapy: 1^o zbudować model probabilistyczny tego doświadczenia, 2^o rozwiązać w modelu zagadnienie matematyczne będące odpowiednikiem postawionego zagadnienia praktycznego, 3^o zinterpretować otrzymany wynik w odniesieniu do rozpatrywanego doświadczenia losowego.
W podanych wyżej etapach 1^o i 3^o stosujemy rozumowanie indukcyjne /korzystamy z praktyki i intuicji/, a w etapie 2^o stosujemy rozumowanie dedukcyjne /korzystamy z twierdzeń matematycznych/. Model probabilistyczny gra w zastosowaniach tylko pomocniczą rolę, ale etap 1^o konstrukcji odpowiedniego modelu doświadczenia losowego jest często najistotniejszy

i najtrudniejszy. Należy przy tym pamiętać, że można skonstruować różne sensowne modele tego samego doświadczenia, i że ostatecznym kryterium dobroci modelu jest praktyka. Sprawy te są bardzo ważne dla człowieka stykającego się z doświadczeniami losowymi, a więc w szczególności dla nauczyciela. Niestety sprawom tym poświęca się w większości książek bardzo mało uwagi.

Autorzy niektórych książek traktują aspekt praktyczny rachunku prawdopodobieństwa marginesowo, koncentrując się na rozpatrywaniu samej teorii matematycznej. W szczególności np. A.A. Borowkow w swojej książce [2] na str. 16 po wprowadzeniu pojęcia prawdopodobieństwa pisze wręcz: "Zauważmy tu, że konkretne wartości liczbowe funkcji P nie będą nas interesować - jest to jedynie sprawa praktycznej wartości tego czy innego modelu".

Autorzy większości książek uwzględniających aspekt praktyczny rachunku prawdopodobieństwa przy rozwiązywaniu zagadnień praktycznych postępują tak, jak gdyby każde doświadczenie losowe miało jeden, oczywisty model, często zresztą nie precyzując go wyraźnie. Tak postępuje m. in. M. Fisz [4].

Może się wydawać, że w zakresie elementarnego rachunku prawdopodobieństwa sprawa ta jest oczywista: w zadaniach dotyczących rzutów symetrycznymi monetami, kostkami, czy wyciągania kul z urn należy po prostu uznać wszystkie możliwe wyniki za jednakowo prawdopodobne, przyjmując tym samym model klasyczny doświadczenia. Otóż jest to sąd całkowicie błędny. Konstrukcja modelu nawet prostego doświadczenia losowego może wielu ludziom sprawiać kłopoty.

Dla przykładu rozpatrzmy doświadczenie losowe polegające na rzucie parą identycznych, symetrycznych kostek sześciennych. Jeżeli na jednej z kostek wypadło i oczek, a na drugiej j oczek, to jest rzeczą sensowną jako wynik tego doświadczenia uznać parę liczb /i, j/. Ponieważ kostki są identyczne, więc jako zbiór zdarzeń elementarnych przyjmiemy zbiór par nieuporządkowanych /i, j/ /i, j = 1, 2, ..., 6/. Par takich jest 21. Zgodnie z sugerowaną wyżej uwagą, że w zadaniach dotyczących rzutów kostkami symetrycznymi wszystkie możliwe wyniki należy uznać za jednakowo prawdopodobne, należałoby więc przyjąć

$$(1) \qquad P(i,j) = \frac{1}{21} \quad (i \leq j).$$

Nauczyciel powinien wiedzieć, że model rzutu parą kostek sześciennych określony wzorem (1) jest zły. Powinien jednak także wiedzieć, że nie ma na to dowodu matematycznego. Można oczywiście przytoczyć różne argumenty przeciwko modelowi (1). Gdyby jednak kogoś te argumenty nie przekonały, to pozostałaby tylko jedna droga: konfrontacja modelu z rzeczywistością - a więc wielokrotne rzuty parą kostek, obserwacja częstości zdarzeń (i,j) i porównanie tych częstości z prawdopodobieństwami (1).

Można sugerować, że rzut parą kostek jest równoważny dwukrotnemu rzutowi jedną kostką, co prowadzi w naturalny sposób do następującego modelu tego doświadczenia:

$$(2) \qquad P(i,j) = \frac{1}{36} \quad (i,j = 1,2,\ldots,6).$$

Nie dla każdego ta równoważność musi jednak być oczywista. Warto może przypomnieć, że d'Alembert nie uważał tego rodzaju doświadczeń za równoważne /patrz np. L.J. Majstrow [10], str. 151/.

Oczywiście błędność modelu (1) rzutu parą kostek symetrycznych nie wynika z niewłaściwego wyboru zdarzeń elementarnych, ale z niewłaściwego przyporządkowania im prawdopodobieństw. Określając prawdopodobieństwa zgodnie ze wzorem

$$(1') \qquad P(i,j) = \begin{cases} \frac{1}{36} & \text{dla } i = j, \\ \frac{1}{18} & \text{dla } i < j, \end{cases}$$

dostajemy dobry model rzutu parą kostek.

Niestety niewiele jest książek, które uwzględniają trudności, jakie może mieć czytelnik rozwiązując zadania dotyczące rzeczywistych doświadczeń losowych. Jedną z nich jest książka J. Neymana [11]. Ponieważ autor ogranicza swoje rozważania do schematu klasycznego, więc sprawa konstrukcji modelu doświadczenia losowego sprowadza się do wyboru zbioru zdarzeń elementarnych. Autor wielokrotnie zwraca uwagę na ważność tej sprawy. M. in. na str. 20 pisze: "Powodem wielu błędów w obliczaniu prawdopodobieństw jest zapomnienie o zbiorze zdarzeń elementarnych", a na str. 55: "Nie jest możliwe przejaskrawienie znaczenia precyzyjnej definicji zbioru zdarzeń elementarnych. Musi ona być wiadoma przed przystąpieniem do rozwiązywania jakiegokolwiek problemu rachunku prawdopodobieństwa." Uogólniając to na przypadek schematów niekoniecznie klasycznych, można powiedzieć, że nie jest możliwe przejaskrawienie znaczenia konstrukcji modelu doświadczenia losowego. Błędy i paradoksy w rachunku prawdopodobieństwa są wynikiem nieesprecyzowania, na czym polega doświadczenie losowe, braku konstrukcji modelu doświadczenia lub konstrukcji złego modelu.

Wszystko to należy wpoić przyszłemu nauczycielowi, aby nie uczył później rachunku prawdopodobieństwa w sposób bezmyślny i dogmatyczny i nie dyskwalifikował ucznia np. za to, że zbudował inny model probabilistyczny niż oczekiwał tego nauczyciel. Należy uświadomić studentom trudności nauczania rachunku prawdopodobieństwa, wskazać na istniejące pułapki, w które mogą wpaść oni lub ich uczniowie, przygotować ich do rzeczowego ustosunkowania się do tzw. "głupich" uwag ucznia w rodzaju: "wypadnięcie orła lub reszki nie jest zdarzeniem pewnym, bo moneta może zatrzymać się na kancie."

Rozpatrzymy teraz kolejno kilka dalszych spraw.

Zaczniemy od pojęcia niezależności. Jak wiadomo, zdarzenia $A_1, A_2, \ldots, A_n$ nazywa się niezależnymi, jeżeli dla dowolnych $1 \leq i_1 < i_2 < \ldots < i_r \leq n$, $r = 2,3,\ldots,n$

$$(3) \qquad P\left(\bigcap_{k=1}^{r} A_{i_k}\right) = \prod_{k=1}^{r} P\left(A_{i_k}\right).$$

Przy rozwiązywaniu zadań wielu autorów sprawę jednak odwraca i z niezależności zdarzeń /jakoby oczywistej/ wnioskuje o spełnieniu wzoru (3). Tak postępują np. T. Gerstenkorn i T. Sródka [5] w przykładzie 4.10.4, uznając za oczywistą niezależność wyników kolejnych rzutów kostką.

Poruszona sprawa jest bardzo ważna i wymaga rzetelnego wyjaśnienia. Otóż trzeba odróżnić pojęcie niezależności zdarzeń od pojęcia niezależności doświadczeń. Niezależność zdarzeń jest to pojęcie matematyczne, określone wzorem (3); niezależność zdarzeń można stwierdzić tylko wtedy, gdy rozkład prawdopodobieństwa P jest już odpowiednio wybrany. Bardzo ważnym pojęciem jest niezależność doświadczeń. O niezależności dwóch /lub większej liczby/ doświadczeń mówimy, gdy wynik jednego z nich nie ma żadnego wpływu na wynik drugiego. Niezależność doświadczeń nie jest pojęciem matematycznym. O niezależności doświadczeń wnioskujemy w oparciu o praktykę i intuicję. Jeżeli modelem probabilistycznym pierwszego doświadczenia jest trójka (Ω_1, M_1, P_1), a modelem drugiego doświadczenia trójka (Ω_2, M_2, P_2) i doświadczenia te są niezależne /w sensie potocznym/, to jest rzeczą sensowną jako model doświadczenia łącznego, polegającego na łącznej /kolejnej lub równoczesnej/ realizacji obu tych doświadczeń przyjąć model produktowy (Ω, M, P), gdzie $\Omega = \Omega_1 \times \Omega_2$, $M = M_1 \times M_2$, $P = P_1 \times P_2$. Łatwo udowodnić, że w tym modelu każda para zdarzeń postaci $A_1 \times \Omega_2$ $(A_1 \in M_1)$ i $\Omega_1 \times A_2$ $(A_2 \in M_2)$ stanowi parę zdarzeń niezależnych w sensie definicji (3), tzn. że

$$P\Big((A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)\Big) = P(A_1 \times \Omega_2)\ P(\Omega_1 \times A_2).$$

Przyszły nauczyciel powinien wiedzieć, że niezależność kolejnych rzutów kostką czy monetą jest faktem empirycznym, a nie matematycznym i że jego uczeń może mieć wątpliwości co do tej niezależności. Warto może przytoczyć dotyczące tej sprawy słowa W. Fellera ([3], str. 132): "Zwykły człowiek, a także filozof Marbe wierzą, że po serii siedemnastu orłów reszka staje się bardziej prawdopodobna. Argumentacja ta nie ma nic wspólnego z niedoskonałością fizycznych monet; obdarza ona pamięcią przyrodę lub, stosując przyjętą przez nas terminologię, przeczy ona niezależności statystycznej kolejnych prób. Teoria Marbego nie może być obalona przez logikę, ale została odrzucona z powodu braku poparcia empirycznego." Przytoczymy jeszcze słowa A.N. Kołmogorowa ([7], str. 19): "... jednym z ważniejszych zadań filozofii nauk przyrodniczych /.../ jest wyjaśnienie i uściślenie tych przesłanek, przy których można jakiekolwiek dane rzeczywiste zjawiska traktować jako niezależne. Zagadnienie to wykracza jednak za ramy tej książki."

Zagadnienie to wykracza także za ramy kursu rachunku prawdopodobieństwa na kierunku nauczycielskim. Warto jednak, żeby nauczyciel wiedział o istnieniu tego zagadnienia i nie usiłować przekonywać swojego ucznia, że niezależność wyników dwóch rzutów kostką jest faktem dającym się dowieść matematycznie.

Przejdźmy do kolejnej sprawy. Prawdopodobieństwo warunkowe definiuje się - jak wiadomo - wzorem

$$(4) \qquad P(A|B) \overset{\text{def}}{=} \frac{P(A \cap B)}{P(B)} \qquad (P(B) \neq 0),$$

gdzie P jest prawdopodobieństwem określonym uprzednio na podzbiorach rozpatrywanej przestrzeni Ω. Po wprowadzeniu tej definicji autorzy rozwiązując zagadnienia praktyczne z reguły nie przejmują się jednak wzorem (4) i bez słowa wyjaśnienia liczą prawdopodobieństwo warunkowe po prostu traktując warunek jako odrębny zbiór zdarzeń elementarnych, nie mówiąc zresztą tego wyraźnie. W szczególności postępują tak przy stosowaniu twierdzenia o prawdopodobieństwie całkowitym i twierdzenia Bayesa. Tę niezgodność obliczania prawdopodobieństwa warunkowego z definicją można znaleźć m. in. w książkach M. Fisza [4], B.W. Gniedenki [6] i A. Rényiego [13].

Na prostym, typowym przykładzie pokażę krytykowany przeze mnie sposób rozwiązywania zadań, a następnie wskażę na inny sposób, zastosowany w moim podręczniku [8].

Przykład 1. Mamy dwie urny z kulami. W pierwszej urnie są 2 białe i 8 czarnych kul, w drugiej 6 białych i 4 czarne kule. Z losowo wybranej urny wyciągamy w sposób losowy kulę. Jakie jest prawdopodobieństwo, że będzie to kula biała?

Rozwiązanie 1. Jest powszechnie przyjęte rozwiązywanie tego typu zadań w następujący sposób. Wprowadźmy oznaczenia:

A_1 - wybór pierwszej urny
A_2 - wybór drugiej urny
B - wybór kuli białej.

Zgodnie z twierdzeniem o prawdopodobieństwie całkowitym

$$(5) \qquad P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2).$$

Z treści zadania wynika, że

$$(6) \quad P(A_1) = P(A_2) = \tfrac{1}{2}, \quad P(B|A_1) = \tfrac{2}{10}, \quad P(B|A_2) = \tfrac{6}{10},$$

a więc

$$P(B) = \tfrac{1}{2} \cdot \tfrac{2}{10} + \tfrac{1}{2} \cdot \tfrac{6}{10} = \tfrac{4}{10}.$$

W podanym rozwiązaniu 1 nie skonstruowano modelu probabilistycznego rozpatrywanego doświadczenia losowego, zdarzenia określa się w sposób opisowy, a nie jako podzbiory odpowiedniej przestrzeni Ω, a prawdopodobieństw warunkowych nie oblicza się zgodnie z definicją (4).

Rozwiązanie 2. Całe opisane doświadczenie losowe można rozbić na dwa etapy: 1/ wybór urny, 2/ wybór kuli z wybranej urny. Jako model doświadczenia losowego z pierwszego etapu jest sensownie przyjąć /o ile urna pierwsza i druga są wybierane jednakowo często/ parę [1] (Ω_1, P_1), gdzie $\Omega_1 = \{1,2\}$,

[1] W przypadku dyskretnej przestrzeni zdarzeń elementarnych Ω w trójce (Ω, M, P) opuszczam σ- ciało M jako oczywiste i piszę krócej (Ω, P).

$$(7) \qquad P_1(1) = P_1(2) = \tfrac{1}{2}.$$

W drugim etapie wykonujemy jedno z dwóch doświadczeń: wybieramy kulę z pierwszej urny lub z drugiej urny. Sensownymi modelami tych doświadczeń są odpowiednio pary $(\Omega_2, P_{2|1})$ i $(\Omega_2, P_{2|2})$, gdzie $\Omega_2 = \{b, c\}$,

$$(8) \qquad P_{2|1}(b) = \tfrac{2}{10}, \qquad P_{2|1}(c) = \tfrac{8}{10},$$

$$(9) \qquad P_{2|2}(b) = \frac{6}{10}, \qquad P_{2|2}(c) = \frac{4}{10}.$$

Łatwo teraz z tych trzech modeli utworzyć sensowny model (Ω, P) całego dwuetapowego doświadczenia, przyjmując $\Omega = \Omega_1 \times \Omega_2$ oraz

$$(10) \quad \begin{aligned} P(1,b) &\overset{\text{def}}{=} P_1(1)P_{2|1}(b) = \frac{1}{2}\cdot\frac{2}{10} = \frac{1}{10},\\ P(1,c) &\overset{\text{def}}{=} P_1(1)P_{2|1}(c) = \frac{1}{2}\cdot\frac{8}{10} = \frac{4}{10},\\ P(2,b) &\overset{\text{def}}{=} P_1(2)P_{2|2}(b) = \frac{1}{2}\cdot\frac{6}{10} = \frac{3}{10},\\ P(2,c) &\overset{\text{def}}{=} P_1(2)P_{2|2}(c) = \frac{1}{2}\cdot\frac{4}{10} = \frac{2}{10}. \end{aligned}$$

To, że powyższy model jest sensownym modelem rozpatrywanego dwuetapowego doświadczenia, można uzasadnić dwojako.

1° Na drodze rozważań dotyczących oczekiwanych częstości poszczególnych wyników przy wielokrotnym powtarzaniu dwuetapowego doświadczenia.

2° Wskazując na pewne własności rozkładu (10), dotyczące zdarzeń A_1, A_2 i B opisanych w rozwiązaniu 1. Otóż zdarzenia te mają obecnie postać:

$$A_1 = \{1\} \times \Omega_2, \qquad A_2 = \{2\} \times \Omega_2, \qquad B = \Omega_1 \times \{b\}.$$

Korzystając z definicji (10) rozkładu P na $\Omega_1 \times \Omega_2$ dostajemy bez trudu

$$P(A_1) = P(A_2) = \frac{1}{2},$$

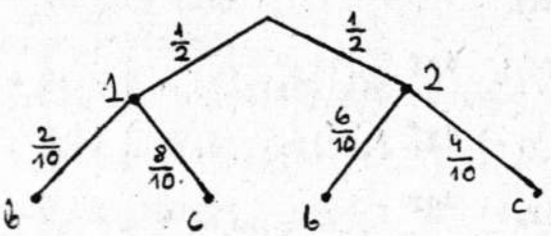
$$P(B|A_1) = \frac{P(B \cap A_1)}{P(A_1)} = \frac{2}{10}, \qquad P(B|A_2) = \frac{6}{10}.$$

Wzory te są identyczne ze wzorami (6) przyjętymi w rozwiązaniu 1 jako intuicyjnie oczywiste.

Mając zbudowany model (Ω, P) całego dwuetapowego doświadczenia rozwiązujemy zadanie natychmiast:

$$P(B) = P(\Omega_1 \times \{b\}) = P(1,b) + P(2,b) = \frac{1}{2}\cdot\frac{2}{10} + \frac{1}{2}\cdot\frac{6}{10} = \frac{4}{10}.$$

Dwuetapowe /i ogólniej wieloetapowe/ doświadczenie losowe można przedstawić poglądowo przy pomocy tzw. drzew. Na rysunku 1 przedstawione jest drzewo ilustrujące rozpatrywane przez nas doświadczenie dwuetapowe. Pierwszy poziom drzewa ilustruje doświadczenie z pierwszego etapu, a drugi poziom - dwa doświadczenia z drugiego etapu.



Rys. 1

Przeprowadzone tu rozważania można bez trudu uogólnić na dowolne doświadczenie dwuetapowe, trójetapowe, n-etapowe, czy wreszcie o nieograniczonej liczbie etapów. Przypadkiem szczególnym tego ostatniego doświadczenia jest łańcuch Markowa. Szczegóły można znaleźć w mojej książce [8].

Wróćmy na chwilę do przykładu 1. W szkole zadania tego typu rozwiązuje się sposobem przedstawionym w rozwiązaniu 1. Nauczyciel powinien jednak zdawać sobie sprawę z tego, że jest to jedynie rozwiązanie skrótowe i że we wzorze (5) w istocie występują prawdopodobieństwa liczone w czterech różnych modelach, mimo oznaczania ich tym samym symbolem P. Nauczyciel powinien także wiedzieć, że stosowanie zbytnich skrótów może prowadzić do błędnych rozwiązań. Dla przestrogi można np. podać następujące rozwiązanie zadania stanowiącego niewielką modyfikację zadania rozpatrywanego przez J. Neymana [11], str. 54.

Przykładem 2. Stwierdzono, że obrona przeciwlotnicza obiektu A zestrzeliwuje 50 % pojawiających się nad nim nieprzyjacielskich samolotów, a obrona przeciwlotnicza obiektu B - 60 % samolotów. Obliczyć prawdopodobieństwo, że samolot nieprzyjacielski wysłany nad oba obiekty zostanie zestrzelony.

Rozwiązanie. Oznaczmy zestrzelenie samolotu nad obiektem A przez Z_A, a zestrzelenie samolotu nad obiektem B przez Z_B. Zgodnie z treścią zadania należy przyjąć, że

$$P(Z_A) = 0,5 , \qquad P(Z_B) = 0,6 .$$

Nas interesuje zestrzelenie samolotu, które może nastąpić bądź nad A, bądź nad B, a więc zdarzenie $Z_A \cup Z_B$. Zdarzenia Z_A i Z_B wyłączają się /samolot może zostać tylko raz zestrzelony/, więc

$$P(Z_A \cup Z_B) = P(Z_A) + P(Z_B) = 0,5 + 0,6 = 1,1.$$

Wynik jest oczywiście błędny. Zadanie to jest szczegółowo omówione w mojej książce [8]. Podana tam jest konstrukcja modelu probabilistycznego rozpatrywanego doświadczenia losowego.

Przejdźmy do pewnych spraw związanych z pojęciem zmiennej losowej. Przypomnijmy definicję.

Definicja. Niech (Ω, M, P) będzie przestrzenią probabilistyczną. Zmienną losową nazywamy każdą funkcję $X: \Omega \to R^1$, mierzalną względem σ- ciała M. Rozkładem zmiennej losowej X nazywamy rozkład P_X określony na klasie B zbiorów borelowskich prostej R^1 następującym wzorem:

$$P_X(A) \overset{def}{=} P\left(X^{-1}(A)\right) = P\left(\{\omega : \omega \in \Omega , X(\omega) \in A\}\right)$$

dla każdego $A \in B$.

Zmienna losowa X prowadzi więc od przestrzeni probabilistycznej (Ω, M, P) do przestrzeni probabilistycznej (R^1, B, P_X). Należy zwrócić uwagę, że powszechnie stosowany zapis $P(X \in A)$ stanowi jedynie inną postać bardziej poprawnego zapisu $P_X(A)$.

Autorzy wielu książek rozwiązując zagadnienia praktyczne przez zmienną losową rozumieją wielkość liczbową będącą wynikiem doświadczenia losowego, nie troszcząc się o zgodność tego terminu z wprowadzoną uprzednio definicją. Nauczycielowi powinno się to wyjaśnić. Spróbujmy wyjaśnić, jak np. należy rozumieć następujące zdanie:

Długość świecenia żarówki I gatunku można traktować jako zmienną losową X o rozkładzie wykładniczym o wartości średniej 1000 godzin.

Otóż sens tego zdania jest w istocie następujący: Dobrym modelem doświadczenia losowego polegającego na świeceniu żarówki gatunku I aż do jej przepalenia jest trójka (R^1, B, P), gdzie P jest rozkładem wykładniczym o wartości średniej 1000, tzn. rozkładem określonym wzorem

$$P(A) = \int_A f(x)dx \qquad (A \in B), \tag{11}$$

gdzie

$$f(x) = \begin{cases} 0 & \text{dla } x < 0, \\ ae^{-ax} & \text{dla } x \geqslant 0, \end{cases} \qquad (a = 1/1000). \tag{12}$$

Wprowadzenie jałowego przekształcenia $X(x) = x$ dla $x \in R^1$ pozwala traktować losową długość świecenia żarówki jako zmienną losową w sensie podanej poprzednio definicji. Rozkładem P_X tej zmiennej losowej X jest oczywiście rozkład wykładniczy o gęstości (12), tzn. rozkład dany wzorem

$$P_X(A) = P(X \in A) = \int_A f(x)dx. \tag{13}$$

Wprowadzanie tu zmiennej losowej X jest jak widzimy niepotrzebne. Traktowanie długości świecenia żarówki jako zmiennej losowej stanowi jedynie powszechnie stosowany żargon.

Niezależność zmiennych losowych X i Y określonych na tej samej przestrzeni zdarzeń elementarnych Ω definiuje się wzorem

$$P(X < x,\ Y < y) = P(X < x)P(Y < y). \tag{14}$$

W praktyce wielu autorów przez niezależne zmienne losowe X i Y rozumie wielkości liczbowe będące wynikami niezależnie od siebie przeprowadzanych doświadczeń losowych, nie troszcząc się o zgodność tego z wprowadzoną uprzednio definicją (14). Studentowi powinno się to wyjaśnić. Wyjaśnijmy, jak np. należy rozumieć następujące zdanie:

Długości świecenia dwóch żarówek I gatunku można traktować jako niezależne zmienne losowe X i Y o rozkładach wykładniczych o wartości średniej 1000 godzin.

Otóż sens tego zdania jest w istocie następujący: Jako model długości świecenia pierwszej żarówki przyjmujemy $\left(R^1_x, B_x, P_1\right)$, gdzie P_1 jest rozkładem wykładniczym o gęstości $f_1(x) = f(x)$ danej wzorem (12). Jako model długości świecenia drugiej żarówki przyjmujemy $\left(R^1_y, B_y, P_2\right)$, gdzie P_2 jest rozkładem wykładniczym o gęstości $f_2(y) = f(y)$ danej wzorem (12). Oba doświadczenia cząstkowe uznajemy za niezależne od siebie /w sensie potocznym/. Dlatego jako model doświadczenia łącznego, polegającego na świeceniu obu żarówek, przyjmujemy model produktowy $\left(R^2, B^2, P\right)$, gdzie $P = P_1 \times P_2$, tzn. gdzie P jest rozkładem określonym wzorem

$$P(A) = \iint_A f_1(x)f_2(y)\,dxdy \qquad \left(A \in B^2\right). \tag{15}$$

Wprowadzenie w R^2 dwóch zmiennych losowych:

$$X(x,y) = x, \qquad Y(x,y) = y$$

pozwala traktować długości świecenia dwóch żarówek jako zmienne losowe. Łatwo przy tym stwierdzamy, że te zmienne losowe są niezależne w sensie definicji (14) i że rozkład P_X zmiennej losowej X jest identyczny z rozkładem P_1, rozkład P_Y zmiennej losowej Y identyczny z rozkładem P_2, a ich rozkład łączny $P_{X,Y}$ identyczny z rozkładem P danym wzorem (15).

Wprowadzenie tu zmiennych losowych X i Y jest właściwie niepotrzebne. Traktowanie długości świecenia pierwszej i drugiej żarówki jako niezależnych zmiennych losowych stanowi jedynie powszechnie stosowany żargon. Ogólnie, jest powszechnie przyjęte traktowanie liczbowych wyników niezależnych doświadczeń losowych jako niezależnych zmiennych losowych. Sprawa zgodności tego z uprzednio przyjętą definicją niezależności zmiennych losowych jest - jak widzieliśmy - prosta, wymaga jednak rzetelnego wyjaśnienia.

Warto zaznaczyć, że wielu autorów rozpatruje rozkłady na prostej /lub ogólniej: w R^n/ wyłącznie jako rozkłady zmiennych losowych /lub ogólniej: wektorów losowych/, co moim zdaniem nie tylko jest dydaktycznie niewłaściwe, ale także może być mylące. Jest to szczególnie rozpowszechnione w książkach nawiązujących do zastosowań. Do pozytywnych wyjątków w tym względzie należy książka S. Zubrzyckiego [16].

Ostatnia sprawa, którą chciałbym tu poruszyć, dotyczy wnioskowania statystycznego. Zacznę od zacytowania kilka uwag R. Bartoszyńskiego, zawartych w jego recenzji [1]:

"Druga uwaga dydaktyczna natury ogólnej dotyczy pojęcia populacji i próbki. Problem nie jest może zbyt skomplikowany, niemniej przeto kryje w sobie różne możliwości nieporozumienia, a nawet pułapki, przed którymi warto ostrzec czytelników, szczególnie jeśli są to studenci, i to stykający się po raz pierwszy z matematycznymi podstawami statystyki.

W największym uproszczeniu /zaniedbując celowo np. zagadnienia wyboru schematu eksperymentu, losowań zależnych, itp./, problem statystyczny da się scharakteryzować następująco. Obserwujemy realizację n-wymiarowej zmiennej losowej $(X_1,\ldots,X_n)$, gdzie wszystkie X_i są niezależne i mają ten sam nieznany rozkład F. Konkretne zadania można formułować w różny sposób, w każdym jednak przypadku chodzi o to, aby na podstawie obserwacji $X_1,\ldots,X_n$ orzekać coś o rozkładzie F /testować hipotezę, że $F = F_0$, szacować parametry F, itp./. Twierdzenia statystyki /a właściwie teorii prawdopodobieństwa/ orzekają o własnościach prawdopodobieństw pewnych zdarzeń wyrażonych poprzez $X_1,\ldots,X_n$ dla różnych rozkładów F; stwarza to z kolei możliwość budowy postępowań dla wnioskowania o nieznanym rozkładzie F.

Praktyk, który chce stosować metody statystyki matematycznej dla rozwiązania określonego zagadnienia, musi wobec tego zinterpretować je w taki sposób, aby wyniki obserwacji można było uznać za realizację n-wymiarowej zmiennej losowej $(X_1,\ldots,X_n)$, a podstawowe pytanie empiryczne stało się pytaniem o rozkład F.

W wielu sytuacjach taką interpretację uzyskuje się w sposób naturalny, wprowadzając pojęcia populacji i próbki. Pojęcia te wywodzą się zapewne z czasów, gdy zastosowania statystyki dotyczyły głównie zagadnień demografii. Wydaje się jednak, że tradycyjne ujmowanie każdego problemu statystycznego w terminach populacji i próbki /zamiast przedstawienia go bezpośrednio w terminach zmiennych losowych/ nie jest najszczęśliwszym dydaktycznie rozwiązaniem i może w efekcie przynieść więcej szkody niż pożytku; w wielu sytuacjach interpretacja w terminach populacja - próbka jest na tyle nienaturalna, że powoduje trudności w zrozumieniu istoty zagadnienia."

Do tych uwag chciałbym dodać następujące. Należy uwypuklić, że wspomniane wyżej zagadnienia statystyczne dotyczą w gruncie rzeczy pewnego doświadczenia losowego, którego losowy wynik traktuje się jako zmienną losową X. Wynik n-krotnej realizacji tego doświadczenia traktuje się jako układ $(X_1,\ldots,X_n)$ n niezależnych zmiennych losowych o rozkładzie identycznym jak X. Rozkład zmiennej losowej X to po prostu model rozpatrywanego doświadczenia losowego. Testowanie hipotezy dotyczącej rozkładu zmiennej losowej X to konfrontacja hipotetycznego modelu z rzeczywistością /z n-krotną realizacją doświadczenia losowego/. Szacowanie parametru rozkładu zmiennej losowej X to wybór odpowiedniej wartości parametru θ, od którego zależy sugerowany przez nas rozkład zmiennej losowej X; inaczej mówiąc - to wybór odpowiedniego modelu doświadczenia spośród jednoparametrowej rodziny modeli. Warto przy tym zauważyć, że nie zawsze zagadnienie powszechnie nazywane szacowaniem parametru θ istotnie polega na szacowaniu prawdziwej, nieznanej wartości θ. Np. trudno byłoby sprecyzować, co należy rozumieć przez prawdziwą wartość parametru a rozkładu wykładniczego uznanego za właściwy model długości świecenia żarówki danego typu. Chodzi tu po prostu o to, aby uznawszy - na podstawie praktyki - rozkład wykładniczy za właściwy typ rozkładu, wybrać jeden z nich, odpowiadający wybranej jakoś na podstawie danych doświadczalnych wartości a. Podobnie trudno byłoby sprecyzować, co należy rozumieć przez prawdziwą wartość parametru p rozkładu dwupunktowego stanowiącego model rzutu monetą.

Naszkicowane wyżej podejście do wnioskowania statystycznego zastosowałem w mojej książce [9].

Sprawy poruszone przeze mnie w tym referacie są proste, ale nie można ich lekceważyć. Trzeba je choć raz przedstawić jasno, z pełną ścisłością. Potem można stosować pewne uproszczenia. Pedantyzm i zbytni formalizm na dalszą metę mogą zaciemniać istotę zagadnienia, ale niedomówienia mogą zrobić dużo więcej szkody.

W moim referacie skupiłem się w zasadzie na sprawach dotyczących nauczania rachunku prawdopodobieństwa na kierunku nauczycielskim. Wydaje mi się jednak, że większość podanych tu uwag odnosi się także do każdego kursu rachunku prawdopodobieństwa nawiązującego do zastosowań. Twierdzę, że dla zastosowań, a także dla celów nauczania rachunku prawdopodobieństwa w szkole nie jest niezbędna znajomość zaawansowanej, abstrakcyjnej teorii. Chociaż, rzecz prosta, można się spotkać z różnymi opiniami na temat stosowalności tego czy innego działu teorii prawdopodobieństwa. Np. W. Feller [3] na str. 137 pisze, że "słabe prawo wielkich liczb ma bardzo ograniczone zastosowanie i powinno być zastąpione przez dokładniejsze i użyteczniejsze mocne prawo wielkich liczb." Natomiast W.N. Tutubalin [14] na str. 164 wygłasza opinię przeciwną: "... jest wątpliwe, aby twierdzenia typu mocnego prawa wielkich liczb mogły z aplikacyjnego punktu widzenia dać więcej niż zwykłe prawo wielkich liczb." Podobną opinię wygłasza B.L. van der Waerden [15] w paragrafie 24. W każdym razie jest dla mnie oczywiste, że dla celów nauczania w szkole oraz dla celów aplikacyjnych ważniejsze niż pedantyczne dowody twierdzeń w ramach dowolnego modelu probabilistycznego jest wyrobienie u słuchaczy intuicji i umiejętności właściwego doboru modelu do danej sytuacji praktycznej. Przy zbyt abstrakcyjnym podejściu od samego początku wykładu rachunku prawdopodobieństwa traci się z pola widzenia doświadczenia losowe. Wydaje się, że rachunek prawdopodobieństwa nie powinien być wykładany w oparciu o ogólną teorię miary, jako jej zastosowanie, ale odwrotnie: ogólna teoria miary powinna być wykładana jako uogólnienie znanych już studentowi szczególnych miar, takich jak miara Lebesgue'a czy prawdopodobieństwo. Elementarny rachunek prawdopodobieństwa może i powinien być wykładany przed ogólną teorią miary. I tak, o ile mi wiadomo, jest na kierunku zastosowań matematyki Uniwersytetu Wrocławskiego, gdzie na 3-cim i 4-tym semestrze wykładany jest Wstęp do rachunku prawdopodobieństwa, a na 4-tym i 5-tym semestrze odrębny wykład z Rachunku prawdopodobieństwa. Podobnie można by zrobić na kierunku nauczycielskim, choć sądzę, że można by się obyć bez tego drugiego wykładu, włączając elementy abstrakcyjnego rachunku prawdopodobieństwa do wykładu ze Wstępu do rachunku prawdopodobieństwa, albo nawet do wykładu z ogólnej teorii miary. Nie wyobrażam sobie natomiast, aby nauczyciel mógł dobrze uczyć rachunku prawdopodobieństwa bez wysłuchania wykładu ze Wstępu do rachunku prawdopodobieństwa, szczegółowo omawiającego poruszone tu przeze mnie sprawy.

LITERATURA

[1] R. Bartoszyński, rec. Wiktor Oktaba i Edward Niedokos, Matematyka i podstawy statystyki matematycznej, Wiadomości Matematyczne XVI /1973/, str. 151-154.

[2] A.A. Borewkow, Rachunek prawdopodobieństwa. Warszawa 1975.

[3] W. Feller, Wstęp do rachunku prawdopodobieństwa. Warszawa 1966.

[4] M. Fisz, Rachunek prawdopodobieństwa i statystyka matematyczna. Warszawa 1967.

[5] T. Gerstenkorn i T. Śródka, Kombinatoryka i rachunek prawdopodobieństwa. Warszawa 1973.

[6] B.W. Gniedenko, Kurs teorii wierojatnostiej. Moskwa 1954.

[7] A.N. Kołmogorow, Osnownyje poniatia teorii wierojatnostiej. Moskwa 1974.

[8] L.T. Kubik, Rachunek prawdopodobieństwa. Podręcznik dla kierunków nauczycielskich studiów matematycznych. Warszawa 1976.

[9] L.T. Kubik, Wprowadzenie do rachunku prawdopodobieństwa

i jego zastosowań. /W przygotowaniu/.

10 L.J. Majstrow, Teoria wierojatnostiej. Istoriczeskij oczierk. Moskwa 1967.

11 J. Neyman, Zasady rachunku prawdopodobieństwa i statystyki matematycznej. Warszawa 1969.

12 A. Rényi, Remarks on the teaching of probability, artykuł w wydawnictwie: L. Råde /ed./, The teaching of probability and statistics, Proceedings of the first CSMP International Conference cosponsored by SOUTHERN Illinois University and Central Midwestern Regional Educational Laboratory, 1970.

13 A. Rényi, Probability theory. Budapest 1970.

14 W.N. Tutubalin, Teoria wierojatnostiej. Moskwa 1972.

15 B.L. van der Waerden, Mathematische Statistik. Berlin-Göttingen-Heidelberg 1957.

16 S. Zubrzycki, Wykłady z rachunku prawdopodobieństwa i statystyki matematycznej. Warszawa 1966.

THE EDUCATION IN THE PROBABILITY THEORY FOR STUDENTS AND TEACHERS A MATHEMATICS

Summary

Some critical remarks concerning many books on probability theory are given. For instance, the use of the notion of conditional probability and that of independence of events is in those books not consistent with definitions of these notions. To avoid these defects it is necessary to make a distinction between random experiment and its probability model. In author's opinion the teaching of probability theory should be based on the construction of probability models.

ОБ ОБУЧЕНИИ ТЕОРИЙ ВЕРОЯТНОСТЕЙ СТУДЕНТОВ И УЧИТЕЛЕЙ МАТЕМАТИКИ

Содержание

В работе приведено несколько критических замечаний о книгах по теории вероятностей. В частности использование понятия условной вероятности и независимости событий не согласовано в этих книгах с определениями этих понятий. Чтобы избежать этих недостатков, нужно различать случайное испытание и его вероятностную модель. По мнению автора построение вероятностных моделей - это основа обучения теории вероятностей.