

ALICJA TRITT

CECHY DIAGNOSTYCZNE RODZINY I CZYNNIKI UMIERALNOŚCI NIEMOWLĄT

I. WSTĘP

Celem niniejszego opracowania jest wybór zmiennych diagnostycznych oraz jednostek przestrzennych dla delimitacji regionów umieralności niemowląt. Zmienne nasze, potocznie nazywane czynnikami zgonów niemowląt, będą tylko czynnikami pośrednimi. Bezpośrednimi czynnikami umieralności niemowląt są choroby lub nieszczęśliwe wypadki, a to jest już sprawa medycyny. Ponieważ nie dla wszystkich interesujących nas zmiennych diagnostycznych znajdziemy dane w dostępnych nam źródłach w ujęciu jednostek przestrzennych, cały proces wyboru zmiennych podzielimy na dwa etapy: prezentację wszystkich preferowanych zmiennych diagnostycznych oraz wydzielenie dostępnych nam zmiennych.

Zastanowimy się także nad założeniami i warunkami wstępnymi zastosowania metod analizy czynnikowej i metod taksonometrycznych. Jako jednostki przestrzenne dla delimitacji regionów umieralności niemowląt wybieramy województwa Polski według podziału administracyjnego z 1 VI 1975 r., badanie przeprowadzimy na podstawie danych z tegoż roku.

Ponieważ ludność jest największym bogactwem kraju, cenne i przedmiotem wielkiej troski staje się życie i zdrowie ludzkie, a szczególnie życie niemowląt, zwłaszcza że prawdopodobieństwo zgonu niemowląt jest bardzo duże. Dopiero w późnej starości spotykamy równie duże prawdopodobieństwo zgonu, jak w niemowlęctwie. Umieralność niemowląt jest ważna również i z tego względu, że od jej wielkości zależy przeciętne trwanie życia ludzkiego. Zależność między przeciętnym trwaniem życia a umieralnością niemowląt jest widoczna. Im więcej umiera niemowląt, tym krótsze jest przeciętne trwanie życia i odwrotnie — w miarę jak cofa się umieralność niemowląt, przeciętne trwanie życia ulega wydłużeniu¹.

¹ E. Rosset, *Umieralność niemowląt a problem długowieczności*, *Studia Demograficzne* 1966, nr 10.

Współczynnik zgonu niemowląt uważa się nieraz za miarę poziomu kulturalnego i gospodarczego społeczeństwa. S. Szulc² w swoich badaniach nad współzależnością rozwoju społeczno-ekonomicznego społeczeństw a zjawiskami demograficznymi, podkreślał bardzo silny ujemny związek między kulturą społeczeństwa a umieralnością we wczesnym dzieciństwie — wraz ze wzrostem kultury obserwujemy bardzo duży spadek umieralności w tym okresie życia.

Polska do 1951 r. należała do krajów o bardzo wysokim poziomie umieralności niemowląt. Wpłynęły na to zapewne trudne warunki ekonomiczne ludności w okresie II wojny światowej oraz w pierwszych latach powojennych. Dziś Polska może poszczycić się ogromnymi sukcesami w walce z umieralnością niemowląt. Jak kształtowała się ona w Polsce w przeciągu ostatnich 25 lat podaje tab. 1.

Tabela 1

Umieralność niemowląt w Polsce w latach od 1950 do 1975 r. (dane co 5 lat)³

Rok	Liczba zgonów niemowląt	
	w tys.	w tym na 1000 urodzeń żywych
1950	82,4	111,2
1955	64,6	82,2
1960	37,5	54,8
1965	22,8	41,4
1970	18,1	33,4
1975	16,0	25,1

Źródeł tego sukcesu należy chyba szukać w tym, że niemowlęta przychodząc dziś na świat mają znacznie lepsze warunki środowiskowe niż dawniej, także przeciętny poziom zdrowia matek i ich uświadomienie w sprawach zdrowia są obecnie lepsze; opieka lekarska nad niemowlęciem stale się doskonali.

Mimo tak znacznego postępu w tej dziedzinie, Polska nie wygląda najlepiej w porównaniu z innymi krajami. Dla porównania podamy wielkości współczynników zgonów niemowląt w niektórych krajach w 1974 r.

Zdefiniujmy teraz pojęcie zgonów niemowląt. Zgodnie z Wielojęzycznym Słownikiem Demograficznym⁴ zjawisko śmierci dzieci, które nie dożyły swoich pierwszych urodzin, nazywamy umieralnością niemowląt. Dla przeprowadzenia głębszej analizy umieralności niemowląt konieczny

² S. Szulc, *Ludność*, Encyklopedia Nauk Politycznych, t. III, z. 1, Warszawa 1938, s. 591.

³ Źródło: Rocznik Statystyczny 1968, s. 48, 1976, s. 43, Warszawa.

⁴ *Wielojęzyczny Słownik Demograficzny*, Warszawa 1966, s. 55.

Tabela 2

Zgony niemowląt w 1974 r.

Kraje	Zgony niemowląt na 1000 urodzeń żywych	Kraje	Zgony niemowląt na 1000 urodzeń żywych
Szwecja	9,6	Stany Zjedn.	
Holandia	11,0	Ameryki P.	22,1
Japonia	11,3 ^a	Włochy	22,6
Dania	11,5 ^a	Austria	23,4
Francja	12,9	POLSKA	23,7
Hiszpania	13,8	Grecja	24,0
Kanada	15,5 ^a	Bułgaria	25,5
NRD	15,9	ZSRR	27,8
Belgia	16,2	Węgry	34,0
W. Brytania	16,3	Rumunia	35,0
Australia	18,5 ^a	Portugalia	38,4
Czechosłowacja	20,4	Jugosławia	40,0
RFN	21,1	Meksyk	51,4 ^a

^a - rok 1973.

Źródło: Rocznik Statystyczny 1976, Warszawa 1976, s. 552.

jest podział umieralności ogólnej na zgony niemowląt z przyczyn endogenicznych i egzogenicznych. W zasadzie do pierwszej grupy zalicza się zgony wywołane słabą konstytucją noworodka i warunkami porodu (przyczyny endogeniczne = wrodzone, wewnętrzne); do drugiej grupy zalicza się zgony, których przyczyną są kontakty dziecka z otoczeniem, w szczególności zgony wywołane chorobami zakaźnymi i wypadkami (przyczyny egzogeniczne = zewnętrzne)⁵. Do przyczyn umieralności endogenicznej zalicza się: 1) wady rozwojowe wrodzone, 2) inne choroby wczesnego dzieciństwa i wcześniactwo. Do umieralności egzogenicznej zalicza się natomiast: 1) choroby zakaźne i pasożytnicze, 2) zapalenie płuc i inne choroby narządów oddechowych, 3) zapalenie żołądkowo-jelitowe i jelita grubego, 4) uszkodzenia porodowe, poporodowe, zamartwica i niedodma noworodków, 5) inne⁶.

J. Bourgeois-Pichat⁷ w swojej metodzie oblicza współczynniki zgonów endogenicznych i współczynniki zgonów egzogenicznych, przyjmując jako kryterium podziału — wiek niemowlęcia w momencie zgonu. Przyjmuje on, że zgony od 0 do 1 miesiąca życia, to zgony endogeniczne, a zgony od 1 miesiąca do 1 roku to zgony egzogeniczne.

⁵ *Wielojęzyczny ...*, s. 59.⁶ Por. też S. Żyromski, *Umieralność niemowląt w województwie olsztyńskim w latach 1950 - 1969*, *Studia Demograficzne* 27/71, s. 87 - 98.⁷ J. Bourgeois-Pichat, *Population* 1951, nr 2 i 3.

O poziomie zgonów w pierwszym miesiącu życia decydują właściwie dwa czynniki: 1) natury biologicznej (wczesniactwo, niedorozwój i wszelkiego rodzaju anomalie), 2) brak właściwej pomocy przy porodzie i złe warunki, w jakich znajduje się kobieta ciężarna⁸. Czynniki biologiczne są bardzo trwałe i wolno ulegają zmianom. Natomiast odpowiednia pomoc przy porodzie może być stosunkowo łatwo wprowadzona i szybko daje dodatnie wyniki. Podłoże zgonów z przyczyn endogenicznych jest wielorakie — przede wszystkim są to wady wrodzone powstałe wskutek wad w komórkach zarodkowych ojca lub matki, wskutek uszkodzeń zarodka przez czynniki szkodliwe (alkohol, niektóre leki używane bez nadzoru lekarza w pierwszych miesiącach ciąży, niewłaściwe odżywianie się matki), uszkodzenia płodu w wyniku chorób matki⁹.

II. PREFEROWANE ZMIENNE DIAGNOSTYCZNE

Zgonów niemowląt, jak zresztą wszystkich pozostałych zjawisk demograficznych nie można zrozumieć i należy ocenić w oderwaniu od ich ekonomicznego i społecznego podłoża. Dlatego zastanowimy się teraz nad czynnikami zgonów niemowląt. Będą to oczywiście czynniki pośrednie, bo bezpośrednimi zajmuje się medycyna. Słowo „czynnik” jest tutaj użyte w potocznym znaczeniu, bowiem chodzi nam nie o ładunki czynnikowe wyodrębnione drogą analizy czynnikowej, a o zmienne diagnostyczne wpływające w pośredni sposób na poziom umieralności niemowląt.

J. Adam i H. Enkle¹⁰ na podstawie ankiety przeprowadzonej dla jednego powiatu przeprowadzili analizę czynnikową umieralności niemowląt. Wyodrębnili oni następujące zmienne:

- 1) płeć: 0 — męska, 1 — żeńska,
- 2) miejsce urodzenia: 0 — zakład (szpital), 1 — dom,
- 3) stan niemowlęcia w momencie urodzenia: 0 — dojrzały, 1 — wcześniak,
- 4) waga niemowlęcia,
- 5) wiek niemowlęcia,
- 6) pochodzenie: 0 — małżeńskie, 1 — niemalżeńskie,
- 7) środowisko społeczne: 0 — normalne, 1 — nienormalne,
- 8) uczęszczanie matki w okresie ciąży do poradni,
- 9) liczba mieszkańców w miejscowości zamieszkania niemowlęcia.

⁸ M. Kacprzak i S. Adamowiczowa, *Zapobieganie zgonom niemowląt*, Studia Demograficzne 1963, nr 2, s. 33 - 44.

⁹ Por. J. Bogdanowicz, Z. Lajembart, I. Sunderland, *Choroby wieku dziecięcego*, Warszawa 1963, s. 55.

¹⁰ *Zur Anwendung statistischer Methoden*, Leipzig 1963.

Nasze preferowane zmienne ujmujemy w pięć grup czynników hipotetycznych:

I. Czynniki demograficzno-biologiczne

- 1) płeć niemowlęcia,
- 2) stan niemowlęcia w momencie urodzenia (dojrzały, wcześniak),
- 3) waga noworodka,
- 4) pochodzenie noworodka ¹¹ (dziecko ślubne, nieślubne),
- 5) wiek matki w chwili urodzenia dziecka,
- 6) kolejny numer urodzenia dziecka,
- 7) liczba porodów martwych matki niemowlęcia,
- 8) liczba poronień samoistnych matki niemowlęcia,
- 9) liczba poronień sztucznych matki niemowlęcia,
- 10) rodność kobiet,
- 11) płodność kobiet.

II. Czynniki ekonomiczno-społeczne

- 1) roczne dochody rodziny,
- 2) zawód wykonywany matki niemowlęcia,
- 3) zawód wykonywany ojca niemowlęcia,
- 4) wielkość zajmowanego mieszkania w izbach,
- 5) wielkość zajmowanego mieszkania w m²,
- 6) liczba utrzymywanych z rolnictwa na 1000 ogółu ludności.

III. Czynniki kulturowe

- 1) wykształcenie matki,
- 2) wykształcenie ojca,
- 3) liczba telewizorów na 1000 mieszkańców,
- 4) księgozbiór w bibliotekach publicznych w woluminach,
- 5) szkoły ogólnokształcące na 1000 ludności.

IV. Czynniki sanitarny i ochrony zdrowia

- 1) liczba lekarzy na 10 000 ludności,
- 2) liczba położnych,
- 3) liczba łóżek szpitalnych na 1000 ludności,
- 4) punkty położnicze,
- 5) wyposażenie mieszkania w wodociąg,
- 6) wyposażenie mieszkania w urządzenia kanalizacyjne,
- 7) zużycie wody, energii elektrycznej i gazu,
- 8) wyposażenie mieszkania w centralne ogrzewanie.

V. Czynniki środowiskowe

- 1) zanieczyszczenie wód,
- 2) emisja pyłów i gazów,
- 3) ścieki przemysłowe.

¹¹ E. Vielrose stwierdza, że zgony niemowląt są wyraźnie częstsze wśród dzieci nieślubnych, niż wśród dzieci ślubnych (E. Vielrose, *Elementy ruchu naturalnego ludności*, Warszawa 1961).

III. DOSTĘPNE ZMIENNE DIAGNOSTYCZNE

Celem naszego opracowania, obok badania współzależności korelacyjnej wyróżnionych zmiennych i wyodrębnienia czynników metodą analizy czynnikowej, jest także delimitacja regionów umieralności niemowląt na podstawie uzyskanych czynników. Naszą podstawową jednostką przestrzenną jest województwo. Musimy więc z preferowanych zmiennych odrzucić te, które nie znajdują pokrycia w dostępnych źródłach w ujęciu wojewódzkim. W ten sposób liczbę naszych zmiennych redukujemy do 17:

I. Czynniki demograficzno-biologiczne

- 1) rodność kobiet,
- 2) płodność kobiet.

II. Czynniki ekonomiczno-społeczne

1) wkłady oszczędnościowe w PKO i bankach spółdzielczych w złotych na 1 mieszkańca (z oficjalnych źródeł nie uzyskamy dochodów ludności, dlatego tę wielkość przyjęliśmy jako miarę poziomu zamożności ludności),

2) liczba utrzymywanych z rolnictwa na 1000 ogółu ludności (ta wielkość zastąpi nam podział na miasto i wieś oraz w pewnym stopniu uwzględni zmienną: zawód rodziców przy ludności rolniczej i nierolniczej),

3) zagęszczenie izb — liczba osób na 1 izbę (jest to wskaźnik sytuacji mieszkaniowej ludności).

III. Czynniki kulturowe

- 1) szkoły ogólnokształcące na 1000 ludności,
- 2) sprzedaż prasy krajowej (gazet) na 1 mieszkańca w egzemplarzach,
- 3) abonenci telewizyjni na 1000 ludności,
- 4) księgozbiór w bibliotekach publicznych w woluminach na 1000 ludności,

IV. Czynniki sanitarny i ochrony zdrowia

- 1) liczba lekarzy na 10 000 ludności,
- 2) liczba położnych na 10 000 ludności,
- 3) liczba łóżek szpitalnych na 10 000 ludności,
- 4) punkty położnicze,
- 5) zużycie wody, energii elektrycznej i gazu na 1 mieszkańca.

V. Czynniki środowiskowe

- 1) zanieczyszczenie wód — ścieki przemysłowe i komunalne w mln m³,
- 2) emisja pyłów zanieczyszczających powietrze atmosferyczne w tys. ton,
- 3) emisja gazów zanieczyszczających powietrze atmosferyczne w tys. ton.

Analizując siedemnaście wybranych zmiennych diagnostycznych wi-

dzimy, że odbiegają one dość znacznie od pierwotnego zespołu zmiennych. Przede wszystkim pominięto szereg ważnych cech demograficzno-społecznych. Różnice te wynikają jednak z przyczyny obiektywnej — braku źródeł do opracowania ogólnokrajowego wszystkich zamierzonych cech, charakteryzujących umieralność niemowląt¹².

Dlatego też dla porównania przeprowadzimy to samo badanie dla pięciu województw: poznańskiego, pilskiego, kaliskiego, leszczyńskiego i konińskiego (obszar byłego województwa poznańskiego). Dla uzyskania pełnego zespołu interesujących nas zmiennych diagnostycznych przeprowadzimy na terenie tych województw ankietę: kwestionariusz w sprawie okoliczności zgonu niemowlęcia (załącznik 1).

Do badania przyjmujemy następujące zmienne:

1) płeć — ponieważ jest to cecha jakościowa, dokonujemy kwantyfikacji, przyjmując: płeć męska — 0, żeńska — 1,

2) stan noworodka — tę zmienną również musimy kwantyfikować; mamy dwa warianty cechy: dojrzały — 0, wcześniak — 1,

3) waga noworodka — miarą tej zmiennej jest ciężar dziecka w chwili urodzenia wyrażony w gramach,

4) wiek niemowlęcia w momencie zgonu podawany w miesiącach,

5) pochodzenie dziecka — uwzględniamy tylko dwa warianty cechy: ślubne — 0, nieślubne — 1,

6) kolejność urodzenia dziecka, tzn. którym z kolei jest zmarłe dziecko,

7) liczba porodów martwych matki niemowlęcia,

8) liczba poronień samoistnych,

9) liczba poronień sztucznych,

10) wiek matki w chwili urodzenia dziecka podany w ukończonych latach,

11) wykształcenie matki,

12) wykształcenie ojca — dwie ostatnie zmienne są mierzone w latach nauki szkolnej; w przypadku, gdy respondent nie poda liczby ukończonych klas (czy lat studiów), a tylko określi stopień wykształcenia, będziemy przyjmować, jak podano w zestawieniu na s. 108.

13) wykonywany zawód matki — zmienna ta należy do jakościowych, musimy więc poszczególnym wariantom tej cechy przyporządkować liczby, tworząc następujące grupy:

matka pracuje w rolnictwie — 0,

matka pracuje poza rolnictwem — 1,

matka jest pracownikiem umysłowym — 0,

matka jest pracownikiem fizycznym — 1,

¹² Departament Badań Demograficznych i Społecznych GUS przeprowadził w końcu 1975 r. badanie ankietowe dotyczące umieralności niemowląt. Badaniem objęto 30-procentową populację niemowląt zmarłych w ciągu 1975 r. Ankieta jest bardzo obszerna, jednak wyniki badania nie są jeszcze opublikowane.

Wykształcenie	Lata nauki szkolnej	Wykształcenie	Lata nauki szkolnej
Niepełne podstawowe	4	Niepełne średnie	9
Podstawowe		Średnie ogólne	11
ukończone	7	Średnie zawodowe	12
Zawodowe	9	Wyższe	15
		Doktorat	18

14) wielkość mieszkania w izbach — otrzymujemy dodając do liczby pokoi samodzielną kuchnię,

15) wielkość mieszkania w m²,

16) wyposażenie mieszkania w wodociąg,

17) wyposażenie mieszkania w kanalizację,

18) wyposażenie mieszkania w centralne ogrzewanie,

19) wyposażenie mieszkania w wc; cztery ostatnie zmienne (16 - 19) jakościowe mają po dwa warianty: mieszkanie jest wyposażone w dane urządzenie (przyporządkowujemy 0) i mieszkanie nie jest wyposażone w dane urządzenie (przyporządkowujemy 1),

20) liczba mieszkańców w miejscu zamieszkania rodziców niemowlęcia, wielkość tę podajemy w tysiącach,

21) odległość w kilometrach od miejsca zamieszkania do stałego ośrodka zdrowia,

22) odległość w kilometrach od miejsca zamieszkania do szpitala.

Jak już pisaliśmy nasze badanie przeprowadzimy w dwóch wariantach. Najpierw obliczymy współzależność korelacyjną między umieralnością niemowląt a 17 zmiennymi ją objaśniającymi oraz między wszystkimi zmiennymi objaśniającymi. Liczba jednostek obserwacji będzie wynosiła 49, ponieważ jako jednostkę obserwacji przyjęliśmy jedno województwo Polski według nowego podziału administracyjnego z 1 VI 1975 r. Następnie przeprowadzimy analizę korelacyjną dla zbiorowości utworzonej z rodzin zamieszkujących 5 wybranych województw, w których wystąpił przypadek zgonu niemowlęcia. Zmiennych objaśniających umieralność niemowląt mamy 22, a liczba obserwacji będzie tutaj zależała od liczby zebranych ankiet.

Dla obu tych zbiorowości wyodrębnimy ładunki czynnikowe drogą analizy czynnikowej. Jeżeli nasza wstępna hipoteza określania czynników poziomu umieralności niemowląt okaże się słuszna, otrzymamy pięć czynników: demograficzno-biologiczny, ekonomiczno-społeczny, kulturowy, sanitarny i ochrony zdrowia, środowiskowy. Trudno nam stawiać hipotezy co do kolejności wymienionych czynników; ale po dokonaniu konkretnych obliczeń będziemy w stanie powiedzieć, jaki czynnik wywiera najistotniejszy wpływ na poziom umieralności niemowląt.

Na podstawie wyodrębnionych czynników dla pierwszej zbiorowości (49 województw) przeprowadzimy delimitację regionów umieralności niemowląt. Dokonujemy rejonizacji umieralności niemowląt, ponieważ istnieje przestrzenne zróżnicowanie tej cechy (różne współczynniki zgonów niemowląt w poszczególnych województwach). E. Vielrose¹³ badał, jak kształtuje się współczynnik zgonów endogenicznych w poszczególnych byłych województwach. Współczynniki zgonów zmniejszają się w każdym województwie z roku na rok, jednak kolejność województw pozostaje mniej więcej taka sama. Wydzielił on województwa, które charakteryzowały się niskimi współczynnikami zgonów niemowląt oraz województwa o wysokich współczynnikach. E. Vielrose interpretuje to w ten sposób, że niektóre województwa mają lepsze warunki bytowania ludzi i to odbija się na mniejszej częstości zgonów endogenicznych. Na tej podstawie uznaje obszar byłych województw: krakowskiego, łódzkiego, kieleckiego, katowickiego i szczecińskiego za szczególnie niezdrowy, natomiast były województwa: olsztyńskie, gdańskie, koszalińskie i opolskie za bardziej zdrowe od pozostałych. Niezdrowe byłyby zatem województwa najbardziej uprzemysłowione, a zdrowe — województwa położone nad morzem i ponadto woj. opolskie. Jaki wpływ wywiera uprzemysłowienie na poziom umieralności niemowląt? Z jednej strony uprzemysłowienie wiąże się z podnoszeniem poziomu życia, z drugiej strony zanieczyszczenia powietrza i wody ujemnie wpływają na zdrowie ludzkie.

Taksonomia pozwoli nam na podzielenie całego kraju na kilka regionów umieralności niemowląt. W danych rejonach otrzymamy województwa, które będą do siebie podobne pod względem wyodrębnionych czynników umieralności niemowląt, natomiast rejony będą się różniły od siebie stopniem podobieństwa pod względem wyodrębnionych czynników.

Przejdziemy teraz do omówienia założeń oraz wstępnych warunków zastosowania metod analizy czynnikowej i metod taksonometrycznych do delimitacji regionów umieralności niemowląt.

IV. ZAŁOŻENIA METODOLOGICZNE ANALIZY CZYNNIKOWEJ

Analiza czynnikowa jest metodą matematyczno-statystyczną. Za twórcę tej metody uważa się Charlesa Spearmana¹⁴, który w 1904 r. opublikował swą pierwszą pracę na temat inteligencji ogólnej obiektywnie określonej i mierzonej (*General intelligence objectively determined and measured*, Am. J. Psychol. 1904, s. 201 - 293). Do rozwoju metody analizy

¹³ E. Vielrose, *Kilka uwag o umieralności niemowląt w Polsce*, Studia Demograficzne 1967, nr 14, Warszawa 1967, s. 35 - 44.

¹⁴ Koncepcję Spearmana dokładniej omawia J. Okoń, *Analiza czynnikowa w psychologii*, Warszawa 1964, s. 167 - 177.

czynnikowej silnie przyczynił się L. L. Thurstone. Jego przełomowe koncepcje pojawiły się w latach trzydziestych okresu międzywojennego.

Analiza czynnikowa jest metodą analizowania zmienności (współ-zmienności) rozmaitych zmiennych, metodą, która prowadzi do wykrycia określonej liczby podstawowych czynników, leżących u podłoża stwierdzonej korelacji.

Charakterystykę analizy czynnikowej można ująć sumarycznie w kilku punktach:

1) analiza czynnikowa opiera się na obserwacjach zmienności w jej naturalnym przebiegu,

2) nie wybiera zespołu zmiennych, które mają być badane, ale umożliwia wydobycie z danego zespołu zmiennych podstawowych czynników,

3) nie wymaga wstępnych hipotez, może natomiast stanowić podstawę dla hipotez, czy też może potwierdzać hipotezy postawione na podstawie danych uzyskanych innymi metodami,

4) nie wymaga wstępnych założeń, które zmienne są zależne, a które niezależne, ani też nie przesądza związku przyczynowego.

Słabą stroną analizy czynnikowej jest to, że brak jest jednoznacznego postępowania przy określaniu tzw. ładunków czynnikowych. Istnieje wiele metod wyodrębniania ładunków czynnikowych: metoda centroidalna i blisko z nią spokrewnione metody: przekątna, grup, grupowania i metoda grup wielokrotnych; metoda największej wiarygodności Lawleya; metoda dwuczynnikowa (spokrewniona z koncepcją Spearmana); metoda czynników dwubiegunowych oraz metoda osi głównych (Hotellinga).

W naszym badaniu zastosujemy metodę centroidalną, zwaną też metodą środka ciężkości¹⁵.

Niech będzie danych n cech, w których interweniuje m czynników. Oznaczamy dowolną cechę przez $x_i (i=1, 2, \dots, n)$, a dowolny czynnik przez $f_j (j=1, 2, \dots, m)$. Zakładamy, że spełniona jest nierówność $m < n$.

Dowolną cechę można przedstawić w postaci równania liniowego:

$$x_i = \sum_{j=1}^m c_{ij} \cdot f_j + b_i \cdot s_i$$

gdzie: $i = 1, 2, 3, \dots, n$

$j = 1, 2, 3, \dots, m$

$n > m$.

Czynnik f_j jest wspólny co najmniej dwu zmiennym podczas gdy s_i jest czynnikiem specyficznym dla każdej cechy. Stała c_j i b_i są wagami czynników, zwanymi także w literaturze — ładunkami czynników.

Założmy, że czynnik $s_i = 0$. Wówczas znalezienie ładunków c_j czyn-

¹⁵ Pochodzenie nazwy tej metody wyjaśnia na przykładzie geometrycznego ujęcia tego problemu J. Okoń, op. cit., s. 91.

ników wspólnych f_j , będzie się sprowadzało do rozwiązania układu równań:

$$x_i = \sum_{j=1}^m c_{ij} \cdot f_j \quad i=1, 2, 3, \dots, n.$$

W interpretacji matematycznej uzyskane z rozwiązania układu równań ładunki czynników wspólnych są współczynnikami kierunkowymi funkcji liniowej każdej cechy.

Teraz przejdziemy do szczegółowego omówienia wszystkich etapów wyodrębniania czynników w metodzie centroidalnej.

Podstawą do przeprowadzenia analizy czynnikowej jest macierz korelacji rzędu zerowego. Wpisując na przekątnej zamiast jedynek zasoby zmienności wspólnej (h^2), otrzymujemy macierz zredukowaną R. Istnieje wiele metod wstępnego określania zasobów zmienności wspólnej. Najprostszą, a zarazem najszerzej stosowaną metodą — jest tzw. metoda najwyższej korelacji. Polega ona na tym, że do kratki leżącej na przekątnej wpisujemy najwyższy współczynnik korelacji z danej kolumny i bez względu na pierwotny znak algebraiczny dajemy mu znak dodatni. Obok metody najwyższej korelacji mamy pokrewną metodę Burta¹⁶. Ścisłą, lecz bardziej pracowitą metodą jest tzw. metoda małego centroidu. Polega ona na tym, że a) dla każdej zmiennej formujemy macierz korelacji 4×4 , obejmującą samą badaną zmienną i trzy inne, z którymi koreluje ona najwyżej. W kratce na przekątnej wpisujemy najwyższe korelacje z każdej kolumny. Zasób zmienności wspólnej określamy według wzoru: $h^2 = (\Sigma r_i)^2 : \Sigma T$, gdzie Σr_i = suma wartości pierwszej kolumny, ΣT = suma wartości całej macierzy 4×4 ; robimy to kolejno dla każdej zmiennej; b) wpisujemy obliczone oceny wartości zasobu zmienności wspólnej do kretek na przekątnej naszej zasadniczej macierzy korelacji.

W celu obliczenia ładunków pierwszego czynnika sumujemy wszystkie kolumny macierzy R łącznie z wartościami na przekątnej (wiersz Σr). Dodajemy wszystkie sumy poszczególnych kolumn, otrzymując wartość T. Sumy poszczególnych kolumn dzielimy przez $\sqrt{T}$ i uzyskujemy ładunki pierwszego czynnika. Możemy to ująć we wzór:

$$C_{1a} = \frac{\Sigma r_a}{\sqrt{T}}$$

Następnie ustalamy macierz pierwszych pozostałości i przystępujemy do odwracania znaków algebraicznych w tej macierzy. Macierz pozostałości obliczamy w ten sposób, że od pierwotnego współczynnika korelacji odejmujemy iloczyn jego ładunków w tych zmiennych. Korzystamy tutaj z twierdzenia: korelacja dwóch zmiennych wywołana jakimś czynni-

¹⁶ Patrz J. Okoń, op. cit., s. 113.

kiem wspólnym (dla tych zmiennych), jest równa iloczynowi ładunków tego czynnika w obu zmiennych. Po odwróceniu znaków obliczamy ładunki drugiego czynnika w sposób podany przy wyodrębnianiu pierwszego czynnika. Przy obliczaniu ładunków dalszych czynników k -tego stopnia postępujemy analogicznie jak przy wyodrębnianiu drugiego czynnika. Zaczynamy zawsze od ustalenia macierzy $k-1$ pozostałości.

Istotnym zagadnieniem jest, kiedy należy przerwać wyodrębnianie dalszych czynników, to znaczy, kiedy możemy mieć pewność, że uzyskaliśmy wystarczającą ilość „wymiarów”, tłumaczących całkowicie współczynniki zawarte w macierzy korelacji. Istnieją różne metody sprawdzania czy zawartość czynnikowa macierzy została całkowicie wyczerpana. Thurstone podaje wzór na minimalną liczbę zmiennych (n), konieczną dla jednoznacznego określenia m czynników:

$$n = \frac{2m+1 + \sqrt{8m+1}}{2}.$$

Rozwiązując powyższy wzór dla m wyznaczamy największą ilość czynników, które mogą być jednoznacznie określone przez n zmiennych:

$$m = \frac{2n+1 - \sqrt{8n+1}}{2}.$$

Według T. Czyż¹⁷, proces wyodrębniania czynników z praktycznego punktu widzenia należy skończyć, gdy wartości własne poszczególnych czynników spadają poniżej jedności. Wartości własne poszczególnych czynników otrzymamy sumując kwadraty ładunków czynnika w każdej ze zmiennych.

W wyniku kolejnych operacji stosowanych przy metodzie centroidalnej otrzymamy układ m czynników, charakteryzujący n początkowych zmiennych. Porównując obliczone i przyjęte zasoby zmienności wspólnej dowiadujemy się czy przyjęte wartości h^2 są dostatecznie ściśle. Jeżeli rozbieżności są zbyt duże, powtarzamy cały proces wyodrębniania czynników przyjmując nowe zasoby zmienności wspólnej (h^2) — zasoby obliczone z pierwszej macierzy czynników.

Na podstawie macierzy czynników możemy wykreślić układ wektorów reprezentujących nasze zmienne. Mamy m czynników, a więc wykres byłby m -wymiarowy. Ponieważ jednak rysunek taki byłby bardzo skomplikowany, a przez to nieczytelny, rozkładamy niejako m -wymiarowy układ wektorów na z wykresów dwuwymiarowych. W miarę wzrostu liczby czynników, liczba potrzebnych wykresów rośnie według wzoru:

$$z = \frac{m(m-1)}{2}, \quad \text{gdzie:} \quad \begin{array}{l} z - \text{liczba wykresów,} \\ m - \text{liczba czynników.} \end{array}$$

¹⁷ T. Czyż, *Zastosowanie metody analizy czynnikowej do badania ekonomicznej struktury regionalnej Polski*, Wrocław, Warszawa, Gdańsk 1971.

Zbiór ładunków czynnikowych n zmiennych, który uzyskamy w wyniku wyodrębniania m kolejnych czynników metodą eentroidalną, nie jest jedynym możliwym ich zbiorem. Mamy tu do czynienia z dwoma zasadniczymi elementami: 1) stałym, niezmiennym układem (konfiguracją) wektorów reprezentujących nasze zmienne; 2) zmiennym elementem układem odniesienia, które nałożyliśmy na konfigurację wektorów.

Obracając układ odniesienia można znaleźć nieskończenie wiele rozwiązań. Stwierdzono, że istnieje jakaś jedna określona pozycja układu odniesienia, która daje zbiór ładunków czynnikowych o szczególnym znaczeniu, a wszelkie pozostałe pozycje są jej matematycznym przekształceniem.

Ustalenie najwłaściwszej pozycji układu odniesienia (jednej z nieskończenie wielu) jest jednym z najtrudniejszych i najsubtelniejszych problemów analizy czynnikowej.

Obrót osi współrzędnych nazywamy rotacją. Walter John i Hans Vahle podają w swojej pracy¹⁸ trzy metody rotacji: 1) rotacja graficzna, 2) metoda Warimax, 3) transformacja specjalna.

Według L. L. Thurstone'a jest to dążenie do tzw. „prostej struktury”. W interpretacji geometrycznej dążenie do „prostej struktury” przejawia się jako skupienie się wektorów reprezentujących zmienne wokół osi czynnikowych. Z dużym uproszczeniem można powiedzieć, że dążąc do „prostej struktury” kierujemy się dwiema następującymi zasadami:

- 1) usiłujemy uzyskać największą możliwą liczbę dużych ładunków czynnikowych,
- 2) usiłujemy uzyskać największą liczbę ładunków zerowych (względnie bliskich zeru) możliwą w danym układzie zmiennych.

Pierwszym krokiem, po otrzymaniu macierzy czynników po wszystkich rotacjach, jest rozpoznanie natury uzyskanych czynników. Badamy szczegółowo końcową macierz czynników, by wykryć, które zmienne są wysoko naładowane danym czynnikiem, a które go nie zawierają. I na tej podstawie dokonujemy interpretacji otrzymanych czynników.

Przy metodzie centroidalnej wyodrębniania czynników konieczna jest rotacja układu odniesienia, natomiast przy metodzie osi głównych Hotellinga uzyskujemy jednoznaczne wyniki. Dlatego też metoda ta, spośród wszystkich metod wyodrębniania ładunków czynnikowych, zyskała największe uznanie matematyków. Twórcą tej metody jest Hotelling, natomiast Kalley wprowadził pewne ulepszenia. Metody tej nie można jednak stosować bez maszyny elektronowej.

Procedura obliczeń przy metodzie Hotellinga rozpada się na dwa zasadnicze etapy: 1) wstępny proces tzw. iteracji, tj. wielokrotnego powta-

¹⁸ W. John i H. Vahle, *Die Faktoranalyse und ihre Anwendung*, Berlin, s. 123-144.

rzania tej samej operacji obliczeniowej dla uzyskania dostatecznie ścisłych wartości, koniecznych dla określenia ładunków czynnikowych; 2) drugi etap obejmuje samo obliczanie tych ładunków.

V. ZAŁOŻENIA METODOLOGICZNE METOD TAKSONOMICZNYCH

Rejonizacja polega na wyodrębnianiu na danym terytorium obszarów składowych. Obszary składowe są charakterystyczne z punktu widzenia przyjętych właściwości. Przy rejonizacji przyjmuje się, że istnieje przestrzenne zróżnicowanie zjawisk, które daje się wyodrębnić i klasyfikować.

Istnieje wiele pojęć rejonu (ponad 100). W jednej ze swych definicji rejonu S. Berezowski¹⁹ pisze: rejon — jest to obszar, którego poszczególne części mają możliwie wiele cech wspólnych i który wykazuje możliwie wiele różnic w stosunku do obszarów otaczających.

Dla dokonania rejonizacji posługujemy się metodami taksonomicznymi. Istota metod taksonomicznych polega na podziale niejednolitego z punktu widzenia określonych i wymiernych cech typologicznych zbioru jednostek na zespoły wewnętrzne bardziej jednolite. Inaczej mówiąc, polega ona na klasyfikacji jednostek przestrzennych badanej zbiorowości ze względu na kilka celowo dobranych cech. Według J. Fiericha²⁰ klasyfikacja jednostek metodą taksonomiczną ma sens przy co najmniej dwóch cechach, ale nie więcej niż dziesięciu. Zbyt duża liczba cech poważnie komplikuje procedurę obliczeniową, a także może zacierać pierwotny obraz wyodrębnionych zespołów jednostek.

Metodę taksonomiczną można stosować operując bezpośrednio wartościami cech diagnostycznych albo też zastosować analizę czynnikową i wówczas większą liczbę cech diagnostycznych zastępujemy mniejszą liczbą czynników wspólnych. Wykorzystując analizę czynnikową zagadnienie taksonomiczne ulega w znacznej mierze uproszczeniu.

Wyróżniamy dwie zasadnicze grupy metod taksonomicznych.

1) Metody wzorcowe, polegające na klasyfikacji badanych jednostek pod względem podobieństwa pewnych cech według z góry określonych wzorów. Do tych metod zaliczamy np. metodę dyskryminacji Fishera, metodę korelacji ortogonalnej Reeda.

2) Metody bezwzorcowe, przy których wyodrębnianie grup klasyfikacyjnych następuje w drodze ściśle określonego postępowania, w wyniku czego łączy się ze sobą elementy najbardziej podobne. Metody bezwzorcowe mogą opierać się na typie jako miejscu największego skupienia cech (do tych metod zaliczamy między innymi metodę dendrytową Steinhausa oraz metodę różnic przeciętnych Czekanowskiego) lub też mogą opierać się na typie jako miejscu nadwyżek (tutaj występuje między

¹⁹ S. Berezowski, *Geografia gospodarcza Polski*, Warszawa 1959, s. 402.

²⁰ J. Fierich, *Próba zastosowania metod taksonomicznych do rejonizacji systemów rolniczych w woj. krakowskim*, *Myśl Gospodarcza*, 1957, nr 1.

innymi metoda przekrojów korelacyjnych Stołyhwy i metoda stochastycznej korelacji wielorakiej, zastosowana przez Wankego).

Inicjatorem metod taksonomicznych był wybitny uczony J. Czekanowski. Wprowadził on dwa warianty tych metod do badań w dziedzinie antropologii: metodę przeciętnych różnic i metodę podobieństw²¹. Metody te były poddawane wielu modyfikacjom i znalazły zastosowanie w innych dziedzinach nauki. Szczególnie popularna jest tzw. metoda przeciętnych różnic. Praca J. Fiericha²² zawiera udaną próbę wykorzystania tej metody w badaniach ekonomicznych (do rejonizacji systemów rolniczych). Metoda przeciętnych różnic Czekanowskiego przetrwała wszystkie późniejsze krytyki i uległa daleko idącym zmianom, które ją udoskonaliły. Jest ona coraz częściej wykorzystywana w pracach badawczych.

Do głównych zalet metody przeciętnych różnic zaliczamy to, że²³:

a) obiektywizuje ona proces poznawczy, nie krępując merytorycznego podejścia poszczególnego badacza, który dobiera sobie zespół cech uznany przez niego za najodpowiedniejszy;

b) zezwala na porównywanie poszczególnych opracowań;

c) jest prosta (oparta na podstawowych działaniach algebraicznych i technice graficznej) i zrozumiała;

d) wyniki uzyskuje się stosunkowo szybko i tanim kosztem,

e) stawia w typowaniu rejonów na pierwszym miejscu związki i podobieństwa pierwszego rzędu, ale nie zaciera związków i podobieństw drugorzędnych; z diagramu odczytać można zarówno syntezę badań, jak i wszystkie szczegółowe powiązania występujące między jednostkami obszarowymi²⁴.

Jako wady metody przeciętnych różnic w jej pierwotnej postaci, podaje się, że:

a) jest to metoda mechaniczna (automatycznie grupuje elementy w zespoły),

b) operuje wielkościami niewspółmiernymi i różnorodnymi, a traktuje je jako jednoznaczne (jednowartościowe),

c) nie uwzględnia różnic wartości poszczególnych cech i ich współzależności,

d) nie można jej stosować do badania cech jakościowych,

e) jest subiektywna w ujmowaniu poszczególnych elementów w diagramie²⁵.

²¹ J. Czekanowski, *Zarys metod statystycznych w zastosowaniu do antropologii*, Warszawa 1913, s. 167 - 173.

²² J. Fierich, op .cit.

²³ Por. J. Steczkowski, *Zasady i metody rejonizacji produkcji rolniczej*, Warszawa 1966.

²⁴ Por. Z. Gługiewicz, *Rejony podaży kontraktowanej trzody chlewnej w woj. poznańskim*, *Ruch Prawniczy, Ekonomiczny i Socjologiczny*, 1960, z. 2.

²⁵ Zdaniem J. Steczkowskiego, op. cit., ostatni zarzut jest sprzeczny z zarzutem sformułowanym w punkcie a.

Zasadniczy proces delimitacji obszarów podobnych przy użyciu metod analizy czynnikowej oraz taksonomicznej metody przeciętnych różnic Czekanowskiego przebiega w następujących etapach:

- 1) dobór cech typologicznych oraz uzyskanie macierzy (F) czynników wspólnych w wyniku przeprowadzenia analizy czynnikowej,
- 2) standaryzacja cech,
- 3) obliczenie współrzędnych wyjściowych punktów (województw) w układzie wyodrębnionych czynników,
- 4) obliczenie numerycznych różnic pomiędzy jednostkami ze względu na dobrane cechy,
- 5) określenie skali podobieństw (odległości) dla jednostek badanej zbiorowości,
- 6) wykreślenie diagramu Czekanowskiego,
- 7) uporządkowanie diagramu.

Ad 1 i 2) O sposobie uzyskania macierzy czynników (F) już pisaliśmy. Po otrzymaniu macierzy F przechodzimy do standaryzacji wartości poszczególnych zmiennych wyjściowych, które użyto dla obliczenia współczynników korelacji²⁶. Standaryzacji (normalizacji) cech dokonujemy za pomocą wzoru²⁷:

$$t_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \quad \begin{matrix} i = 1, 2, 3, \dots, n \\ j = 1, 2, 3, \dots, m \end{matrix}$$

gdzie: x_{ji} — zaobserwowane wartości cechy x_j dla każdego i

$\bar{x}_j$ — średnia wartość cechy x_j

S_j — odchylenie standardowe cechy x_j , przy czym:

$$\bar{x}_j = \frac{1}{N} \sum_{i=1}^N x_{ji}$$

$$S_j^2 = \frac{1}{N} \sum_{i=1}^N (x_{ij} - \bar{x}_j)^2$$

Wartości t_{ij} są wartościami poszczególnych cech diagnostycznych znormalizowanych. Istnieje jeszcze szereg innych sposobów standaryzacji, omawia je obszernie J. Fierich²⁸.

Znormalizowane wartości charakteryzują się tym, że ich średnia jest

²⁶ Można także obliczyć współczynniki korelacji dopiero po normalizacji danych wyjściowych i zastosować prosty wzór: $r_{jk} = \frac{1}{N} \left(\sum_{i=1}^N z_{ij} \cdot z_{ik} \right)$. Wartości współczynników korelacji będą oczywiście takie same, jak wyliczone wzorem Pearsona.

²⁷ J. Paradysz, A. Zeliaś, *Wykorzystanie analizy czynnikowej do badania poziomu warunków bytu ludności w układach regionalnych*, Problemy Ekonomiczne, 1971, 1, s. 31 - 48.

²⁸ A. Zeliaś, *Ekonometryczne metody budowy prognoz (na przykładzie produkcji roślinnej)*, Kraków 1970, s. 26.

równa 0, a odchylenie standardowe równe jedności. Dzięki normalizacji zmiennych można wszystkie badane cechy porównywać i sumować. Macierz znormalizowanych zmiennych oznaczamy symbolem **T**. Stanowi ona ostateczną postać macierzy obserwacji i jest przedmiotem dalszej analizy.

Ad 3) Współrzędne wyjściowe województw w układzie otrzymanych czynników otrzymamy przez przemnożenie macierzy znormalizowanych zmiennych (**T**) przez macierz czynników (**F**). Obliczoną macierz oznaczymy symbolem **Y**. W wyniku tego działania wyjściowy zbiór K punktów ($k_i = t_{i1}, t_{i2}, \dots, t_{im}; i=1, 2, \dots, n$) z przestrzeni m -wymiarowej przeprowadzamy na zbiór **Y** w przestrzeni r -wymiarowej czynników wspólnych, którego elementami są następujące punkty:

$$Y_i = k_i F \quad i=1, 2, \dots, n^{29},$$

Jak widać z powyższego równania — podobszary (województwa) przedstawione zostaną jako punkty w przestrzeni r -wymiarowej.

Ad 4) W rezultacie obliczenia sumarycznych różnic pomiędzy województwami ze względu na dobrane cechy otrzymujemy sumaryczną macierz różnic odległości między każdym punktem a wszystkimi innymi punktami. Zasadnicza idea metody różnic przeciętnych polega na tym, że im bliżej siebie znajdują się poszczególne powiaty w przestrzeni r -wymiarowej, tym bardziej są podobne podobszary reprezentowane przez te punkty, a odległości między nimi są dokładnymi wskaźnikami wielocechowego podobieństwa tych obszarów.

Istnieje wiele sposobów obliczania różnic między elementami. Omawia je dokładnie J. Steczkowski w swojej pracy³⁰.

W naszych obliczeniach posłużymy się sposobem Czekanowskiego³¹:

$$w_{ik} = \frac{a_{ij} - a_{kj}}{m} = \frac{R_{ik}}{m},$$

gdzie: m — liczba cech diagnostycznych³²,

n — liczba elementów (województw),

a_{ij} , a_{kj} — porównywany i -ty oraz k -ty element pod względem j -tej cechy.

Ad 5) Aby wykorzystać macierz odległości **D** dla grupowania podobszarów o maksymalnej jednorodności, należy odpowiednio zakwalifikować sumaryczne różnice między poszczególnymi elementami do określonych przedziałów klasowych, czyli ustalić skalę podobieństw.

²⁹ Por. J. Steczkowski, op. cit.

³⁰ Ibidem.

³¹ Por. J. Czekanowski, op. cit.

³² Przy obliczaniu różnic — można podzielić sumę różnic przez liczbę cech diagnostycznych (kolumny) — ale nie jest to konieczne, choć najczęściej stosowane (stąd nazwa).

J. Paradysz i A. Zeliaś w swojej pracy³³ podkreślają, że ustalenie poprawnej skali podobieństw jest problemem o niezwykle znaczeniu dla ostatecznego wyodrębnienia jednorodnych rejonów. Trudności, jakie napotyka się w tym zakresie rozwiązuje się najczęściej w praktyce na drodze kolejnych przybliżeń.

H. Baranek³⁴ w celu określenia prawidłowej skali podobieństw proponuje budowę szeregu rozdzielczego (i jego graficzną prezentację) na podstawie uporządkowanych w relacji rosnącej wartości sumarycznych różnic między każdą parą punktów. Wielobok liczebności określony w przedziale zmienności sumarycznych różnic ma w tym przedziale kilka maksimów i kilka minimów. Jest to oznaką, że należy dany zbiór punktów zróżnicować, tworząc kilka odrębnych zbiorowości możliwie do siebie podobnych. Trzeba więc odnaleźć proste oddzielające te zbiory. Szukamy tych prostych w ten sposób, aby uzyskać szeregi rozdzielcze o jednym maksimum. Tak postępując, otrzymamy z klas podobieństw.

J. Steczkowski³⁵ uważa, że poprawna skala podobieństw odpowiadać powinna równości:

$$\frac{n^z}{z} \approx f_1 \approx f_z,$$

gdzie: f_1 — liczebność podobieństw w klasie pierwszej,

f_z — liczebność podobieństw w klasie ostatniej,

z — liczba klas,

n^z — liczba sumarycznych różnic.

Ad 6) Opierając się na ustalonych klasach podobieństw, budujemy — z macierzy różnic odległości **D** — nieuporządkowany diagram Czekanowskiego.

Ad 7) Z kolei naszym zadaniem jest możliwie najlepsze uporządkowanie diagramu w ten sposób, aby otrzymać największe skupienie podobieństw pierwszego stopnia wzdłuż głównej przekątnej. Wykorzystamy do tego metodę taksonomii wrocławskiej³⁶. Metoda ta służy do porządkowania zbioru elementów i jego podziału na podzbiory wewnętrznie bardziej podobne. Porządkowanie zbioru przeprowadza się na podstawie

³³ J. Paradysz, A. Zeliaś, *Wykorzystanie analizy...*, s. 41.

³⁴ H. Baranek, *O pewnej procedurze wyboru klas podobieństwa w taksonomicznej metodzie różnic*, w: *Studia statystyczne i demograficzne. Praca z zastosowań statystyki w demografii, socjologii*, Pr. zbior. pod red. K. Zająca, Wrocław, Warszawa, Kraków, Gdańsk.

³⁵ J. Steczkowski, op. cit.

³⁶ Omawiają tę metodę między innymi K. Florek, J. Łukaszewicz, J. Perkal, H. Steinhaus, S. Zubrzycki w: *Taksonomia wrocławska*, Przegląd Antropologiczny, t. XVIII, Poznań 1951. O metodzie dendrytowej pisze także B. Podolec, *Ekonomiczne metody ustalania regionów konsumpcyjnych w Polsce* (praca doktorska), Kraków 1973.

tablicy sumarycznych różnic odległości. Wybiera się z niej odległości najmniejsze, które świadczą o tym, że dwa elementy stoją (pod względem dobranych cech) najbliżej siebie. Przez łączenie linią najbliższych stojących elementów (województw) tego zbioru otrzymuje się dendryt. Jest on linią łamaną, która rozgałęzia się, lecz nie może zawierać łamanych zamkniętych, a równocześnie taka, że każde dwa punkty danego zbioru są przez nie połączone.

Najlepsze uporządkowanie zbioru będzie wtedy, gdy otrzymamy dendryt o najmniejszej długości (a więc najmniejszą sumę odległości pomiędzy poszczególnymi parami punktów sąsiadujących ze sobą w danym uporządkowaniu).

Na podstawie zróżnicowanej skali podobieństw oraz znajomości dendrytu sporządzamy możliwie najlepiej liniowo uporządkowany diagram Czekanowskiego. Na podstawie tego diagramu wyróżniamy jednolite regiony z punktu widzenia wyodrębnionych czynników.

*

W wyniku dokonanych rozważań doszliśmy do tego, że wyodrębnienie czynników umieralności niemowląt drogą analizy czynnikowej przeprowadzimy w dwóch wariantach. Na podstawie publikowanych źródeł wybraliśmy 17 zmiennych diagnostycznych określających poziom umieralności niemowląt i na podstawie tych 17 zmiennych zbadamy współzależność korelacyjną i przeprowadzimy analizę czynnikową. Ponieważ wśród zmiennych tych brakuje istotnych zmiennych demograficzno-biologicznych i społeczno-ekonomicznych analizę czynnikową dokonamy również na podstawie współzależności korelacyjnej między dwudziestoma dwoma zmiennymi uzyskanymi drogą ankietową („Kwestionariusz w sprawie okoliczności zgonu niemowlęcia”). Interesujące będzie, w jakiej mierze czynniki uzyskane z obu tych wariantów, będą się pokrywały. Przypuszczamy jednak, że najistotniejszy będzie czynnik demograficzny, ponieważ pośrednie przyczyny zgonów niemowląt tkwią przede wszystkim w warunkach rodzinnych, w jakich znajduje się dziecko. Dlatego chcielibyśmy badać umieralność niemowląt w ścisłym powiązaniu z rodziną. Jeżeli uzyskamy potrzebne informacje (22 zmienne z drugiego wariantu) w przekroju wojewódzkim dla całego kraju, zrezygnujemy z przeprowadzania badań w pierwszym wariantcie. Otrzymane wyniki byłyby na pewno doskonalsze.

KWESTIONARIUSZ W SPRAWIE OKOLICZNOŚCI ZGONU NIEMOWLĘCIA

1. Płeć: męska, żeńska (właściwe podkreślić)
2. Stan noworodka: dojrzały, wcześniak (właściwe podkreślić)
3. Waga noworodka

4. Wiek niemowlęcia w chwili zgonu
5. Pochodzenie: małżeńskie, niemalżeńskie (właściwe podkreślić)
6. Kolejność urodzenia dziecka
7. Liczba porodów martwych, w których latach
8. Liczba poronień, w których latach
9. Wiek matki w chwili urodzenia dziecka
10. Wykształcenie ojca, matki
11. Nazwa zakład pracy, miejscowość, zawód i stanowisko
 - a) ojca
 - b) matki
12. Mieszkanie:

liczba pokoi powierzchnia w m² samodzielna

kuchnia (tak, nie) powierzchnia
13. Urządzenia sanitarne w mieszkaniu (pisać: tak, nie) wodociąg, kanalizacja, centralne ogrzewanie, wc
14. Nazwa miejsca zamieszkania
miasto, wieś (właściwe podkreślić), liczba mieszkańców
15. Podać nazwę miejscowości, w której znajduje się najbliższy:
 - a) stały ośrodek zdrowia, odległość w km
 - b) ośrodek czynny w niektóre dni, odległość w km
 - c) szpital, odległość w km

FACTOR ANALYSIS AND DELIMITATION OF REGIONS BY FACTORS OF INFANT MORTALITY

S u m m a r y

In the first part of the article the Author has chosen 33 preliminary diagnostic variables differentiating the level of infant mortality. The variables belong to five groups of hypothetical factors of infant mortality, namely: 1) demo-biological factor, 2) socio-economical factor, 3) cultural factor, 4) health protecting factor, 5) environmental factor. Afterwards 16 variables without necessary source information have been excluded from the total. In order to improve the information on the 17 variables left in the set the Author has elaborated a survey questionnaires concerning the infant deaths in the recent years.

In the second part of the article the Author has discussed the methods of factor analysis and taxonometrics. He has found the centroidal method (L. L. Thurstone's method) as the appropriate method of factor analysis and the average difference method (Czekanowski's method) as the best taxonomic method for region delimitation. At last, the new voivodships created on the 1 st of June 1975 have been accepted as the basic spatial units for region delimitation.