Listy Biometryczne
Nr 34-36 (1972)

Mirosław Krzyśko, Andrzej Dyczkowski
Michał Karoński, Paweł Olejniczak (Poznań)

ZASTOSOWANIE METOD BIOMETRYCZNYCH DO PROBLEMÓW KLASYFIKACJI ZUPEŁNIE NIEMIAROWEJ CZYNNOŚCI SERCA

Przebadano trzydziestu chorych z zupełnie niemiarową czynnością serca spowodowaną migotaniem przedsionków, w tym z wadami mitralnymi - 17 chorych i ze zwyrodnieniem mięśnia sercowego - 13 chorych. Wszystkim wykonano elektrokardiogramy i pomierzono długości czasowe odcinków R-R zespołu QRS. W ten sposób każdy z pacjentów scharakteryzowany został k_i-elementowym ciągiem obserwacji $t_1, \ldots, t_{k_i}$ dla $i = 1, \ldots, 30$, stanowiącym próbę. U poszczególnych osób liczba obserwacji była różna i zdeterminowana czasem pomiaru wynoszącym przeciętnie 4 minuty. Na podstawie prób zbudowano szeregi rozdzielcze. Obserwacje grupowano w klasy o stałej długości 0,1 sekundy. Każdy z rozkładów scharakteryzowany został następującymi parametrami (por. [2]): wartością średnią $\bar{t}$, odchyleniem standardowym s, współczynnikiem asymetrii g_1 oraz współczynnikiem spłaszczenia g_2. Istotność współczynników g_1 i g_2 testowano (patrz [3]) przy użyciu statystyk $u_1(g_1) = g_1/D(g_1)$ i $u_2(g_2) = g_2/D(g_2)$, gdzie $D(g_1)$ i $D(g_2)$ oznaczają odchylenia standardowe tych współczynników. Wyniki testowania podano w tablicy 1.

Następnie przystąpiono do grupowania rozkładów. Niech n_{ij} będą licznościami w szeregach rozdzielczych zbudowanych dla dwóch pacjentów; $i = 1, 2$, oraz $j = 1, \ldots, r$, gdzie r jest liczbą klas w porównywanych szeregach.

Oznaczmy:

$$n_{.j} = \sum_{i=1}^{2} n_{ij} \quad \text{i} \quad n_{i.} = \sum_{j=1}^{r} n_{ij} .$$

Łatwo zauważyć, że

$$N = \sum_{i=1}^{2} \sum_{j=1}^{r} n_{ij} = \sum_{i=1}^{2} n_{i\cdot} = \sum_{j=1}^{r} n_{\cdot j} \, ,$$

jest liczbą wszystkich obserwacji w dwóch porównywanych szeregach rozdzielczych.

Za miarę odległości (niepodobieństwa) dwóch rozkładów odpowiadającym dwom chorym przyjęto $\sqrt{\chi^2}$, gdzie

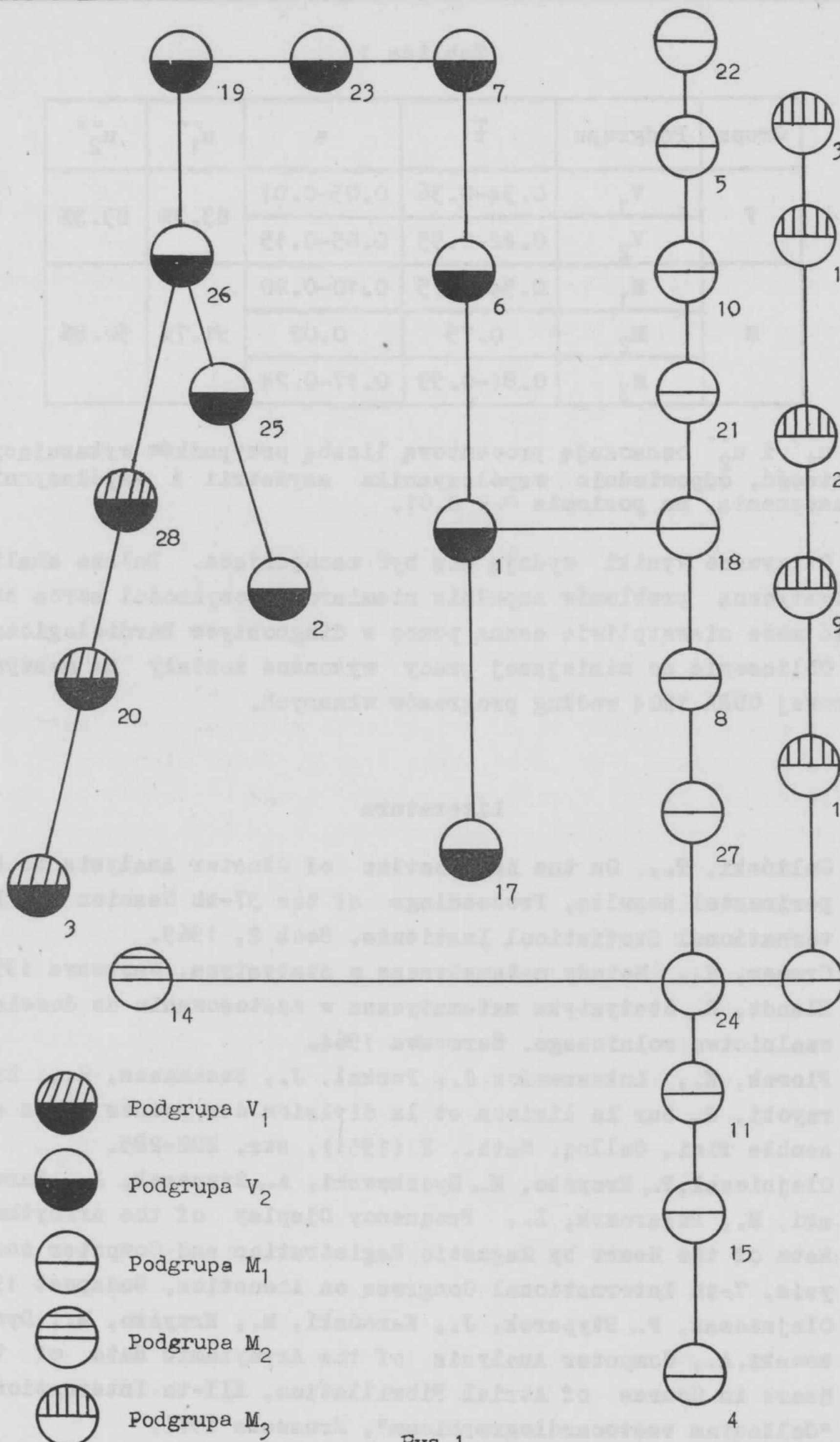
$$\chi^2 = N \sum_{i=1}^{2} \sum_{j=1}^{r} \left(\frac{n_{ij}^2}{n_{i\cdot} \, n_{\cdot j}} - 1 \right),$$

jest znaną statystyką K. Pearsona [2].

Zbudowano tablicę wzajemnych odległości między trzydziestoma obiektami (chorymi), która była punktem wyjścia do ich pogrupowania. Użyto niezależnie dwóch sposobów grupowania (skupiania) obiektów. Pierwszy z nich polega na podziale dendrytu (rys.1) na zwarte grupy [1], [4]. Jako kryterium podziału stosuje się minimum sumy kwadratów odchyleń dla zmienności wewnątrzgrupowej. Podział ten prowadzi się sukcesywnie aż do osiągnięcia minimum, a dla każdego podziału wyliczana jest statystyka typu F, która stanowi kryterium optymalnej liczby wydzielonych grup. Druga użyta metoda skupień traktuje początkowo wszystkie obiekty jako grupy jednoelementowe. W pierwszym kroku dokonuje się połączenia dwóch najbliższych grup i uważa się je jako jeden obiekt. Po utworzeniu nowego "łącznego" obiektu (grupy) buduje się nową tablicę odległości o wymiarze zredukowanym o jedność. W ten sposób postępuje się sukcesywnie aż do połączenia wszystkich obiektów w zadaną liczbę grup.

W wyniku grupowania otrzymano dwie zasadnicze grupy jednorodne oznaczone na rys. 1 przez V i M. Przy dalszym podziale otrzymano pięć podgrup oznaczonych przez V_1, V_2 i M_1, M_2, M_3 należących odpowiednio do grup V i M. Obie metody grupowania prowadzą do identycznych wyników. Poszczególne grupy scharakteryzowano w tablicy 1.

Porównując liczbę zaklasyfikowanych przypadków chorobowych w wyżej opisany sposób z rozpoznaniami klinicznymi, uzyskano przy wadach mitralnych 88,2% poprawnych zaklasyfikowań do grupy M (błędnie zaklasyfikowano obiekty nr 12 i 24) oraz w przypadku zwyrodnienia mięśnia sercowego 92,3% poprawnych zaklasyfikowań (błędnie zaliczony obiekt nr 3) do grupy V.


19
23
7
22
5
30
16
10
26
6
21
25
29
28
1
18
2
9
20
8
13
27
17
3
14
24
12
11
15
4
Podgrupa V_1
Podgrupa V_2
Podgrupa M_1
Podgrupa M_2
Podgrupa M_3

Rys.1.

Tablica 1

Grupa	Podgrupa	$\overline{t}$	s	u_1^{**}	u_2^{**}
V	V_1	0.34-0.36	0.05-0.07	83.3%	83.3%
	V_2	0.42-0.55	0.05-0.15		
M	M_1	0.54-0.75	0.10-0.20	91.7%	50.8%
	M_2	0.75	. 0.02		
	M_3	0.81-0.99	0.17-0.24		

u_1^{**} i u_2^{**} oznaczają procentową liczbę przypadków wykazujących istotność, odpowiednio współczynnika asymetrii i współczynnika spłaszczenia, na poziomie $\alpha = 0.01$.

Otrzymane wyniki wydają się być zachęcające. Dalsza analiza statystyczna problemów zupełnie niemiarowej czynności serca stanowić może niewątpliwie cenną pomoc w diagnostyce kardiologicznej.

Obliczenia do niniejszej pracy wykonane zostały na maszynie cyfrowej ODRA 1204 według programów własnych.

Literatura

[1] Caliński, T., On the Application of Cluster Analysis to Experimental Results, Proceedings of the 37-th Session of International Statistical Institute, Book 2, 1969.

[2] Cramer, H., Metody matematyczne w statystyce, Warszawa 1958.

[3] Elandt, R., Statystyka matematyczna w zastosowaniu do doświadczalnictwa rolniczego, Warszawa 1964.

[4] Florek, K., Łukaszewicz, J., Perkal, J., Steinhaus, H., Zubrzycki, S., Sur la liaison et la division des points d'un ensemble fini, Colloq. Math., 2 (1951), str. 282-285.

[5] Olejniczak, P., Krzyśko, M., Dyczkowski, A., Styperek, J., Karoński, M., Polarczyk, Z., Frequency Display of the Arrhythmic Rate of the Heart by Magnetic Registration and Computer Analysis, 7-th International Congress on Acoustics, Budapest 1971.

[6] Olejniczak, P., Styperek, J., Karoński, M., Krzyśko, M., Dyczkowski, A., Computer Analysis of the Arrhythmic Rate of the Heart in Course of Atrial Fibrillation, XII-th International "Colloqium vectocardiographicum", Brussels 1971.