

УДК 519.237.8:612 + 378.037.1-057.875:519.237.8

## **КЛАСТЕРНЫЙ АНАЛИЗ МОРФОФУНКЦИОНАЛЬНЫХ ПОКАЗАТЕЛЕЙ ЗДОРОВЬЯ СПОРТСМЕНОК**

*А.Н. Пермина*

В работе представлены результаты изучения некоторых морфофункциональных показателей здоровья спортсменок. Реализована статистическая модель классификации спортсменок на основе полученных экспериментальных данных

Ключевые слова: кластеризация, иерархический кластерный анализ, метод Уорда.

В настоящей работе исследовалось влияние основных антропометрических и физиологических показателей студентов, ранжированных в соответствии с особенностями состояния здоровья и двигательной активности на три группы (т.н. группы здоровья), на их физическую подготовленность. Изучение различных аспектов подобного влияния представляет определенный интерес с точки зрения оптимизации практики физического воспитания и повышения уровня здоровья студентов.

Для исследования были отобраны студенты университета, которые, в соответствии с результатами предварительного обследования, были разбиты на три специальные группы для прохождения физической подготовки.

В первую группу (16 чел.) отбирали практически здоровых индивидов, во вторую (15 чел.) – индивидов с незначительными отклонениями от нормы, в третью (15 чел.) – со значительными отклонениями в здоровье. Возраст испытуемых колебался в диапазоне от 18 до 21 года.

Каждая из групп занималась физической подготовкой по специально разработанным программам. Физическая нагрузка первой группы предполагала повышенную двигательную активность (два занятия с использованием фитнес-технологий на уроках физического воспитания, и три раза в неделю занятия различными видами спорта по интересам) с общим объемом 8–10 часов в неделю. Вторая группа занималась по обычной программе физической подготовки (2 раза в неделю) и, дополнительно – туризмом выходного дня. Суммарно – 8 часов в неделю. Участники третьей группы занимались по индивидуальным программам лечебной физической культуры два раза в неделю с ежедневными самостоятельными занятиями в те дни, когда уроков лечебной физической культуры не было.

Анализ двигательной активности испытуемых проводился на основе исследования следующих показателей: возраст (Age), рост (H) и вес (W) обследуемых, уровни систолического и диастолического давления (SAD, DAD), реакции частоты сердцебиений на ортовоздействие (Ort), индекс

Скибинской (Ind\_Skib), 6 смешанных упоров, подтягиваний (6\_s\_u) плюс пресс (Pod\_Press), показатели ручной динамометрии – правой (F\_right) и левой (F\_left).

Отметим, что по показателям частоты сердцебиений можно судить о нарушениях регуляции сердечно сосудистой системы. Индекс ручной динамометрии отражает специфику спортивных занятий студентов, индекс Скибинской – функциональные резервы дыхательной и сердечно-сосудистой систем человека.

В качестве тестового показателя физической подготовленности были использованы результаты испытуемых в беге на 100 м (Sprint\_100) и кроссе на 500 и 1000 м (Cross).

Целью исследования было, в первую очередь, установление того, насколько обоснованным является отнесение того или иного испытуемого к той или иной группе физической подготовки.

Одним из способов решения этой задачи может служить разбиение всего пула испытуемых на однородные кластеры в соответствии с похожестью или непохожестью характеризующих их признаков. Такое разбиение дает возможность проводить структурный и смысловой анализ полученных кластеров, и в дальнейшем успешно прогнозировать спортивную результативность и выявлять индикаторы повышения успешности спортивной деятельности.

Для выполнения кластеризации был выбран метод Уорда (Ward's minimum variance method, 1963 [1], [2]), как метод, расширяющий пространство, т.е. склонный к образованию мелких, более «отчетливых» кластеров.

Первоначально, обозначим набор показателей как  $\{\xi_1, \xi_2, \dots, \xi_{14}\}$ , группу индивидов, как  $\{u_1, u_2, \dots, u_{46}\}$ . Таким образом, все полученные в результате эксперимента данные можно представить в виде матрицы-таблицы:

Таблица 1

Полученные в результате эксперимента данные

Показатель Индивид	$\xi_1$	$\xi_2$	...	$\xi_n$
$u_1$	$x_{11}$	$x_{12}$	...	$x_{1n}$
$u_2$	$x_{21}$	$x_{22}$	...	$x_{2n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$u_m$	$x_{m1}$	$x_{m2}$	...	$x_{mn}$

Таким образом,  $x_{ij}$  – это значение  $j$ -ого показателя для  $i$ -ого индивида.

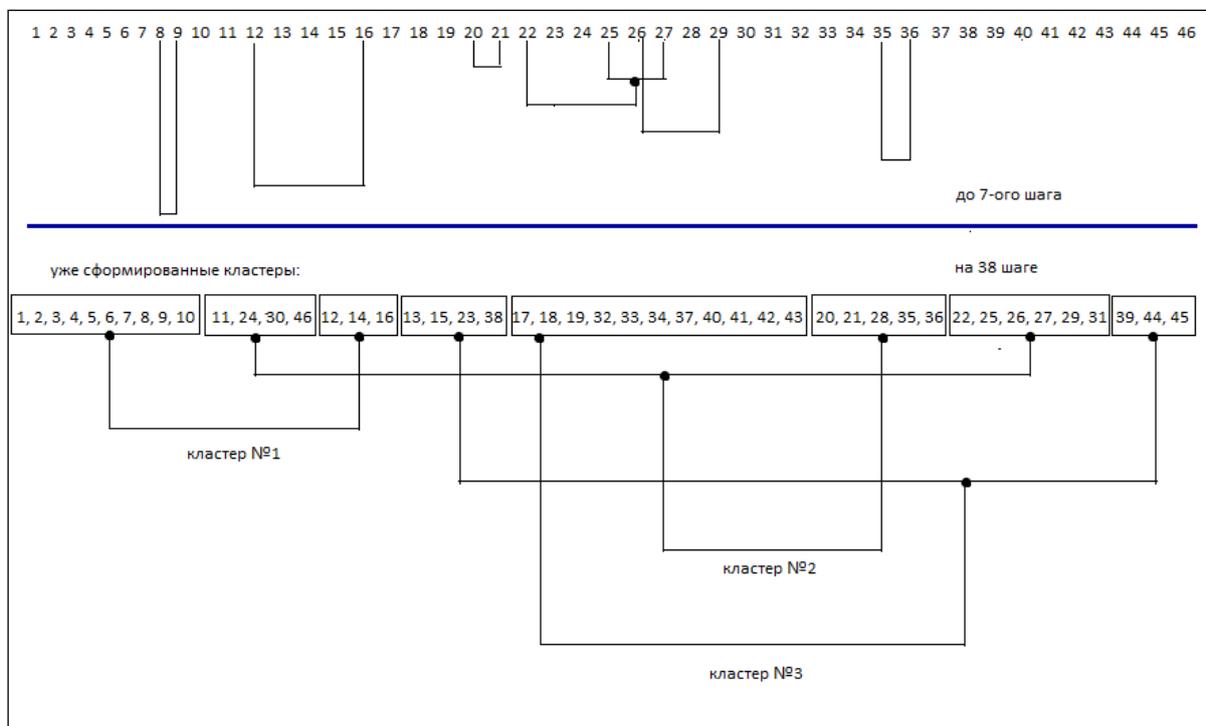
На первом шаге кластеризации каждый индивид объявляется отдельным одноэлементным кластером. На каждом последующем шаге объединяются такие два кластера, которые приводят к минимальному увеличению целевой

функции, т.е. внутригрупповой суммы квадратов (т.н.  $SS_E$  – the error sum of squares [3]), в наших обозначениях:

$$SS_E = \sum_{i=1}^n \left( \sum_{j=1}^m x_{ij}^2 - \frac{1}{m} \sum_{j=1}^m (x_{ij})^2 \right).$$

В конечном итоге все индивиды объединяются в один общий кластер. С целью определения стадии кластеризации, на которой произошло образование наиболее однородных, схожих по значениям показателей, кластеров, была построена и изучена дендрограмма, отображающая динамику процесса объединения кластеров.

Схематично дендрограмма объединения кластеров представлена на рисунке:



Дендрограмма процесса кластеризации

На рис. 1 видно, что на первом шаге в один кластер объединились индивиды № 20 и № 21, на втором шаге № 25 и № 27 и так далее. После синей черты показана ситуация уже на 38 шаге, когда образовались 8 кластеров. Далее процесс кластеризации развивается до 43 шага, на котором четко выделяются три кластера, каждый из которых включает в себя однородных, по включенным в анализ показателям, испытуемых.

Результаты кластерного анализа приведены в таблице 2.

Таблица 2

Кластерный анализ

Кластер Группа	I кластер	II кластер	III кластер	Количество	% верной классификации
Группа 1	13	2	1	16	81,25 %
Группа 2	0	11	4	15	73,8 %
Группа 3	4	2	9	15	60,0 %
Итого	17	15	14	46	71,7 %

Проведенный сравнительный анализ принадлежности к группе здоровья продемонстрировал, что идентификация принадлежности к первой группе оказалась верной в 81,25 % случаев, ко второй – в 73,8 %, к третьей – не столь успешная – в 60,0 %. Общее количество правильно классифицированных индивидов – более 71 %.

Для сравнения результатов классификации, параллельно с кластеризацией был проведен дискриминантный анализ экспериментальных данных. В рассматриваемой ситуации он оказался более эффективным, как показывает таблица 3.

Таблица 3

Дискриминантный анализ

Кластер Группа	I кластер	II кластер	III кластер	Количество	% верной классификации
Группа 1	16	0	0	16	100,00 %
Группа 2	0	15	0	15	100,0 %
Группа 3	2	1	12	15	80,0 %
Итого	18	16	12	46	93,48 %

Таблицы 2 и 3 свидетельствуют, что результаты предварительного медицинского обследования, на основании которого комплектовались группы здоровья, хорошо согласуются с полученными классификационными результатами.

Библиографический список

1. Ким, Дж.-О. Факторный, дискриминантный и кластерный / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др. – М.: Финансы и статистика, 1989. – С. 141–176.
2. Gabor J. Szekely. Hierarchical clustering via joint between-within distances: Extending Ward's minimum variance method / G.J. Szekely, M.L. Rizzo // Journal of classification. – 2005. – № 22. – Pp. 151–183.
3. Joe H. Ward, Jr. Hierarchical grouping to optimize an objective function / J.H. Ward Jr. // Journal of the American Statistical Association. – 1963 – № 58. – Pp. 235–240.

[К содержанию](#)