

лабораторным работам, моделирования химических процессов и явлений, компьютеризация эксперимента, решение задач и проведение количественных расчетов, осуществление самоконтроля и стандартизированного контроля знаний. Важнейшим преимуществом является многосторонность рассмотрения изучаемых объектов, процессов или явлений. Это позволяет использовать компьютер для решения различного рода задач и проблем, то есть компьютер вводится в качестве инструментального средства в познавательный процесс при проблемном методе обучения.

УДК 519.24:616-005

КЛАССИФИКАЦИЯ ПАЦИЕНТОВ-ЖЕНЩИН ПО СТЕПЕНИ ЗАБОЛЕВАЕМОСТИ АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИЕЙ

Сенько Е.В., студ.

Гродненский государственный университет имени Я. Купалы,

г. Гродно, Республика Беларусь

Реферат. *Проведен анализ статистики заболеваемости артериальной гипертензией в мире, сформирована система медицинских показателей, характеризующих пациентов с артериальной гипертензией, осуществлена кластеризация пациентов-женщин по выбранным показателям. Построены классификационные функции.*

Ключевые слова: заболевание, артериальная гипертензия, моделирование, классификация, факторный анализ, кластерный анализ, дискриминация, функции.

Введение. Здоровье для каждого – неоценимое богатство человека, это условие счастья. Всемирная Организация Здравоохранения (ВОЗ) определяет здоровье, как состояние полного физического, духовного и социального благополучия, а не только отсутствие болезней и физических дефектов. Существует масса определений здоровья.

Часто встречающееся заболевание в современном мире – это артериальная гипертензия.

Артериальная гипертензия (АГ) – это заболевание сердечно-сосудистой системы, при котором давление крови в артериях системного (большого) круга кровообращения стабильно повышено.

Актуальность проблемы артериальных гипертензий для здравоохранения республики обусловлена распространенностью патологии и значительными экономическими потерями в связи с инвалидизацией и смертностью населения от осложнений артериальной гипертензии. Согласно эпидемиологическим исследованиям практически у каждого пятого жителя Республики Беларусь в возрасте старше 18 лет отмечается повышенное артериальное давление.

Европейские эксперты прогнозируют к 2025 году 29,0 % мужчин и 29,5 % женщин в мире будут иметь АГ. В большинстве стран мира имеются национальные программы по борьбе с АГ [2, с. 121].

Изолированное повышение артериального давления в 70 % случаев приводит к инсульту у женщин. Чтобы понимать в сравнение, у мужчин АГ без сопутствующих факторов риска заканчивается инсультом лишь в 30 % случаев. На сегодня осложнения АГ занимают первое место, если говорить о смертности в женской популяции.

Цель работы – исследовать степень заболеваемости группы пациентов-женщин АГ методами прикладной статистики. Объект исследования - это результаты пациентов-женщин, находящихся на клиническом обследовании, а предмет – артериальная гипертензия.

На начальном этапе настоящей работы была сформирована система показателей, характеризующих пациентов с АГ. После этого были применены методы прикладной статистики – факторный анализ: метод главных компонент, кластерный анализ: метод k-средних и дискриминантный анализ с пошаговым включением в статистическом пакете Statistica 8 [1, с. 22]. На следующем этапе была проведена классификация пациентов-женщин и определены функции классификации.

Для анализа пациентов с артериальной гипертензией были выбраны данные по 300

женщинам. В качестве системы показателей, характеризующих степень заболевания АГ, были взяты: вес, индекс массы тела (ИМТ), частота дыхания (ЧД), конечно-систолический размер левого желудочка (КСР), конечно-систолический объем левого желудочка (КСО), конечно-диастолический размер левого желудочка (КДР), конечно-диастолический объем левого желудочка (КДО) и артериальное давление: систолическое и диастолическое (АД = САД/ДАД).

В ходе факторного анализа были получены факторные рейтинги, а затем проделан кластерный анализ на основании полученных групп переменных. Таким образом, с помощью данных анализов были классифицированы переменные и категории респондентов по однородным кластерам.

Дискриминантный анализ позволил решить следующую задачу: какие переменные и как разделили входящий поток данных на группы [1, с. 48]. Дискриминантная функция, выведенная посредством дискриминантного анализа, – линейная комбинация независимых переменных, с помощью которой можно наилучшим образом различить категории зависимой переменной.

Полученные результаты. Для начала был применен факторный анализ, с помощью которого были построены рейтинги пациентов. Далее в ходе кластерного анализа пациенты-женщины были разбиты на 3 кластера. В таблице 1 отображены средние значения показателей в кластерах по классифицируемым признакам. Интерпретация средних значений показателей женщин представлена в таблице 2. Имеем: первая группа – женщины с нормальными показателями, вторая – пациенты-женщины, которые имеют заболевание АГ, а третью группу составили пациенты-женщины с незначительными нарушениями (нн) показателей.

Таблица 1 – Средние значения показателей в кластерах

Кластеры	Кол-во пациентов	вес	ИМТ, кг/м	ЧД	КСР, мм	КСО, мл	КДР, мм	КДО, мл	САД (мм. рт. ст.)	ДАД (мм. рт. ст.)
1	73	67,5	20,1	16,3	29,2	33,3	45,6	79,35	132,1	84,9
2	155	96,3	31,9	19,7	39,8	66,3	60,2	125,5	173,1	106,0
3	72	78,6	25,6	18,7	36,8	56,2	56,1	112,3	150,7	94,9

Источник: собственная разработка автора по расчетам в пакете Statistica.

Таблица 2 – Интерпретация средних значений показателей пациентов-женщин

Кластеры	Кол-во пациентов	вес	ИМТ, кг/м	ЧД	КСР, мм	КСО, мл	КДР, мм	КДО, мл	САД (мм. рт. ст.)	ДАД (мм. рт. ст.)
1	73	норма	норма	норма	норма	норма	норма	норма	норма	норма
2	155	болен	болен	болен	болен	болен	болен	болен	болен	болен
3	72	нн	нн	нн	нн	нн	нн	нн	нн	нн

Источник: собственная разработка автора по расчетам в пакете Statistica.

Далее был проведен дискриминантный анализ. В качестве группирующей переменной выступил кластер, а независимые переменные – это показатели, характеризующие АГ. Для анализа был выбран метод с пошаговым включением.

Значение лямбды Уилкса равно 0,042, значит, данная классификация является корректной. Статистика лямбда Уилкса служит для проверки качества дискриминации (чем ближе к 0, тем меньше вероятность ошибочного разделения).

При помощи стандартизированных коэффициентов можно непосредственно сравнивать относительный вклад каждой независимой переменной в различие исследуемых групп. Исходя из вычисленных данных, в первую дискриминантную функцию наибольший вклад вносят показатели КДР, КДО, а во вторую – САД, КДО и КДР.

Построенные дискриминантные функции являются статистически значимыми, так как вычисленные р-значения являются равными больше 0,05. С помощью функций классификации определяем принадлежность классифицируемых наблюдений к

определенному классу (табл. 3).

Наблюдение относится к тому классу, для которого функция классификации примет максимальное значение.

Таблица 3 – Функции классификации для пациентов-женщин

	G_1:1 - p=,24333	G_2:2 - p=,51667	G_3:3 - p=,24000
КДР, мм	22,4	65,4	56,8
КСР, мм	20,4	49,5	42,1
САД (мм. рт. ст.)	1,1	12,4	-0,8
КДО, мл	21,4	52,9	48,1
ИМТ, кг/м	5,4	19,9	11,8
КСО, мл	10,1	35,8	29,2
ЧД	4,9	17,5	14,5
вес, кг	3,1	9,6	5,7
ДАД (мм. рт. ст.)	-11,7	-19,4	-20,6
Constant	-13,0	-103,4	-67,4

Источник: собственная разработка автора по расчетам в пакете Statistica.

Затем для каждого конкретного наблюдения, которое должно быть классифицировано, вычисляются все функции классификации. Они получились следующие:

$$F1 = 22,4 * КДР + 20,4 * КСР + 1,1 * САД + 21,4 * КДО + 5,4 * ИМТ + 10,1 * КСО + 4,9 * ЧД + 3,1 * вес - 11,7 * ДАД - 13,0;$$

$$F2 = 65,4 * КДР + 49,5 * КСР + 12,4 * САД + 52,9 * КДО + 19,9 * ИМТ + 35,8 * КСО + 17,5 * ЧД + 9,6 * вес - 19,4 * ДАД - 103,4;$$

$$F3 = 56,8 * КДР + 42,1 * КСР - 0,8 * САД + 48,1 * КДО + 11,8 * ИМТ + 29,2 * КСО + 14 * ЧД + 35,7 * вес - 20,6 * ДАД - 67,4.$$

Например, F1 – это первый кластер, F2 – третий кластер, F3 – второй. Для того, чтобы определить принадлежность пациента, нужно подставить значения соответствующих медицинских показателей в формулы. Исходные значения показателей приведены в таблице 4.

Таблица 4 – Медицинские показатели пациента-женщины

КДР, мм	КСР, мм	САД, мм рт. ст.	КДО, мл	ИМТ	КСО, мл	ЧД	вес, кг	ДАД, мм рт. ст.м
53	32	133	69	17,43	25	15	59	80

Источник: собственная разработка автора.

Таким образом, значения функций получились равными:

$$F1 = 3116,922; F2 = 10764,857; F3 = 9108,274.$$

Так как F2 является максимальным значением, то пациент относится ко второму кластеру. А исходя из таблицы 2 у рассматриваемого пациента незначительные нарушения.

Заключение. Итак, в работе исследованы медицинские показатели пациентов-женщин с артериальной гипертензией. Построено три однородные группы пациентов-женщин, а также функции классификации. Полученные результаты позволяют определить принадлежность классифицируемых наблюдений к определенному классу. Применение методов прикладной статистики позволяет моделировать и анализировать различные ситуации, обеспечивать определенную точность и достоверность результатов.

Список использованных источников

1. Буреева, Н. Н. Многомерный статистический анализ с использованием ППП

- «STATISTICA» / Н. Н. Буреева. – Нижний Новгород, 2007. – 112 с.
2. Лисицын, Ю. П. Теории медицины XX века / Ю.П.Лисицын. – М.: Медицина, 1999. – 176 с.

УДК 004.02

РАБОТА С МАССИВАМИ. АНАЛИЗ МЕТОДОВ СОРТИРОВКИ

Стасеня Т.П., ст. преп., Мандрик О.Г., ст. преп., Яснев Д.А., студ.

Витебский государственный технологический университет,

г. Витебск, Республика Беларусь

Реферат. В статье рассматривается разработка программного проекта для демонстрации и анализа алгоритмов простых методов сортировки. Программа выполняет расчет количества шагов выбранных методов сортировки, что позволяет оценить эффективность использования данного метода.

Ключевые слова: программирование в среде Delphi, обработка массивов данных, методы сортировки.

Цель данной работы состояла в разработке программного проекта, который позволит визуально наблюдать этапы работы выбранных простых методов сортировки.

При работе с однотипными объектами часто возникает необходимость выполнить сортировку. Основное условие поставленной задачи: выбранный метод сортировки массивов должен экономно использовать доступную память. Это предполагает, что перестановки, приводящие элементы в порядок, должны выполняться на том же месте, т. е. методы, в которых элементы из массива А передаются в результирующий массив В, представляют меньший интерес.

Методов сортировки разработано много. С учетом экономии памяти будем выполнять подбор нужного метода среди многих возможных. Рассмотрим классификацию методов по их экономичности, т. е. по времени работы. Хорошей мерой эффективности может быть число необходимых сравнений и число пересылок (перестановок) элементов. Эти числа являются определяющими объём работ.

Простые и очевидные методы сортировки называют прямыми. Причины выбора прямых методов: прямые методы удобны для объяснения основных характеристик большинства сортировок; программы этих методов понятны и коротки; усложнённые методы требуют небольшого числа операций, но используют сложные операции; для достаточно малых наборов прямые методы оказываются быстрее, хотя при больших размерах массивов использовать простые методы не следует.

Простые методы сортировки можно разделить на три категории:

1. Сортировки с помощью включения (by insertio).
2. Сортировки с помощью выделения (by selection).
3. Сортировки с помощью обмена (by exchange).

Рассмотрим механизмы выбранных методов сортировки

Сортировка с помощью прямого включения

Элементы мысленно делятся на уже «готовую» последовательность $a_1 \dots a_{i-1}$ и исходную последовательность $a_i \dots a_n$. При каждом шаге, начиная с $i=2$ увеличивая i каждый раз на единицу, из исходной последовательности извлекается i -й элемент и перекладывается в готовую последовательность, при этом он вставляется в нужное место.

Минимальные работы в случае уже упорядоченной исходной последовательности элементов, наихудшие же вариант – когда исходные элементы расположены в обратном порядке.

Сортировка с помощью прямого выбора

Этот прием основан на следующих принципах: выбирается элемент с наименьшим значением; он меняется местами с первым элементом a_1 ; затем этот процесс повторяется с оставшимися $n-1$ элементами, $n-2$ элементами и т. д. до тех пор, пока не останется один, самый большой элемент.

Такой метод называют прямым выбором – в некотором смысле противоположным