

СТАТЬИ

УДК 378:796

ОСОБЕННОСТИ ПРИМЕНЕНИЯ ПАРАМЕТРИЧЕСКИХ И НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ ПРИ АНАЛИЗЕ ДАННЫХ В ФИЗИЧЕСКОЙ КУЛЬТУРЕ И СПОРТЕ**Абдрахманова И.В., Лущик И.В.***ФГБОУ ВО «Волгоградская государственная академия физической культуры», Волгоград, e-mail: academy@vgafk.ru*

В статье описаны трудности, возникающие у молодых исследователей при решении задачи проверки статистических гипотез на выборках малого объема. Данное положение обусловлено специфическими особенностями сбора данных в сфере физической культуры и спорта. Большинство рассмотренных авторами трудов, представленных в открытых информационных ресурсах, включает сопоставление экспериментальной и контрольной групп по уровню подготовленности, определение эффективности новаций в методическом обеспечении тренировочного процесса посредством сравнения показателей спортсменов в разные моменты времени. В условиях работы с небольшим массивом данных основной проблемой является оптимальный выбор наиболее мощного непараметрического критерия. Авторами выделены основные ошибки, допускаемые магистрантами и обучающимися на бакалавриате при определении статистического инструментария. Представлены результаты анализа причин, обуславливающих некорректность интерпретации математических моделей, построенных на основе применения параметрических критериев. Приведены примеры неоднозначной трактовки итогов использования непараметрических критериев, представленные в современной научно-методической литературе. Описаны процедуры оптимизации отбора статистического критерия для решения локальных задач статистического анализа, позволяющие существенно снизить вероятность получения неадекватных или противоречивых результатов обработки данных малого объема в области физической культуры и спорта.

Ключевые слова: статистический анализ, физическая культура, непараметрические критерии, проверка статистических гипотез

FEATURES OF APPLICATION OF PARAMETRIC AND NONPARAMETRIC CRITERIA IN DATA ANALYSIS IN PHYSICAL CULTURE AND SPORT**Abdrakhmanova I.V., Luschik I.V.***Volgograd State Physical Education Academy, Volgograd, e-mail: academy@vgafk.ru*

The article describes the difficulties that young researchers face when solving the problem of testing statistical hypotheses on small samples. This provision is due to the specific features of data collection in the field of physical culture and sports. Most of the works reviewed by the authors, presented in open information resources, include a comparison of the experimental and control groups according to the level of preparedness, determining the effectiveness of innovations in the methodological support of the training process by comparing the performance of athletes at different points in time. When working with a small amount of data, the main problem is the optimal choice of the most powerful nonparametric criterion. The authors have identified the main mistakes made by undergraduates and undergraduate students in the determination of statistical tools. The results of the analysis of the reasons for the incorrect interpretation of mathematical models built on the basis of the use of parametric criteria are presented. Examples of ambiguous interpretation of the results of the use of non-parametric criteria, presented in modern scientific and methodological literature, are given. The procedures for optimizing the selection of a statistical criterion for solving local problems of statistical analysis are described, which can significantly reduce the likelihood of obtaining inadequate or inconsistent results of processing small data in the field of physical culture and sports.

Keywords: statistical analysis, physical education, nonparametric criteria, statistical hypothesis testing

Высокая публикационная активность современных молодых ученых определяется высоким уровнем развития информационной среды. Возможность ознакомления с широким спектром исследований, представленных студентами и магистрантами в научных изданиях, позволяет сделать следующие выводы о составе испытуемых, исследуемых признаках и наиболее популярных критериях:

– выборки формируются из спортсменов различных возрастных групп с идентичной специализацией, или производится дифференциация по половому признаку

в случае принадлежности к одной возрастной группе;

– целью исследования является оценка эффективности экспериментальных авторских методик на основе анализа изменений уровня сформированности физических качеств или выявление зависимости между различными показателями скоростных, силовых, скоростно-силовых, антропологических или морфологических характеристик;

– широко используются параметрический t-критерий Стьюдента для связанных или несвязанных выборок, непараметрические критерии Манна – Уитни и Краскела – Уоллиса.

Популярность параметрического t-критерия Стьюдента обусловлена многоаспектностью его применения. Среди решаемых посредством данного критерия задач доминируют следующие:

- построение интервальной оценки среднего значения и дисперсии исследуемого признака;
- выявление выбросов на основе сопоставления зарегистрированных значений и построенного интервала;
- оценка степени однородности выборочной совокупности.

Математическое моделирование реализуется на основе учета взаимосвязи следующих характеристик: «ширины доверительного интервала, соответствующей ему доверительной вероятности и объема выборки» [1, с. 43].

Выбор t-критерия Стьюдента должен иметь в качестве обоснования для использования подтверждение нормального закона распределения выборки, что часто игнорируется молодыми исследователями.

Критерий Манна – Уитни используется для сопоставления средних значений двух выборок, при условии, что между исходными данными отсутствует корреляционная связь. Допускается наличие в одной из выборок трех значений. Критерий Краскала – Уоллиса есть многомерное обобщение критерия Вилкоксона – Манна – Уитни, представляющее собой непараметрический аналог однофакторного дисперсионного анализа. При отборе непараметрических критериев учитываются количество и объем сопоставляемых выборок. Студенты и магистранты испытывают сложности при отборе или ошибаются в выборе вследствие игнорирования такого показателя, как мощность критерия.

С целью решения представленных выше проблем были определены условия, снижающие вероятность неверного выбора или некорректного применения статистического критерия для анализа данных малого объема.

Цель исследования – выявить причины некорректности выводов, полученных на основе использования параметрического t-критерия Стьюдента для проверки статистических гипотез на выборках малого объема; описать алгоритм отбора непараметрического критерия, минимизирующего риск получения противоречивых результатов анализа малого массива данных.

Материалы и методы исследования

В процессе исследования произведено изучение соответствующих материалов, представленных в открытых информационных источниках.

Результаты исследования и их обсуждение

Указанные преимущества t-критерия Стьюдента не могут быть в полной мере реализованы в области физической культуры и спорта, так как его использование ограничено следующими требованиями:

- распределение выборочных данных должно подчиняться нормальному закону;
- дисперсии сопоставляемых выборок должны быть равны [2, с. 129].

В современной специальной научной литературе детально рассматривается проблема отбора мощного статистического критерия для проверки гипотезы о нормальном распределении выборочных данных малого объема. При этом выделяются простые гипотезы (о соответствии распределения данных какому-либо закону с заданными априори параметрами) и сложные (о соответствии распределения данных какому-либо закону с произвольными параметрами, определенными согласно эмпирическим значениям). Выделяют общие критерии согласия, разработанные для проверки простых гипотез, и частные.

В этой связи отмечается высокая популярность аппроксимации эмпирических данных распределением из семейства Пирсона, при которой возможно построение кривой, описывающей плотности их распределения. Данная процедура позволяет определить характеристики распределения, например его симметричность, поведение на хвостах и прочее. Негативной при этом является сложность определения части параметров (например, квантилей) по построенной аналитической модели распределения плотности вероятности.

Развитие информационных технологий позволяет модернизировать процедуру расчета и интерпретации классических критериев, разработанных в XX в. Разработка онлайн-калькуляторов, отображающих количественные результаты статистического анализа, словесное резюме и соответствующую графическую интерпретацию, сменилась созданием нейронов, эквивалентных критериям согласия [3].

Сравнительный анализ итогов использования данных продуктов, а также результатов разностороннего сравнения критериев, описанных в специальной научной литературе, позволил сделать следующие заключения:

- критерий Пирсона, являясь оптимальным для применения в полиномиальных схемах и наиболее популярным критерием согласия при проверке простых гипотез, не имеет достаточной мощности при работе

с малыми массивами данных (менее 20 измерений) и высокую вероятность ошибки второго рода;

– критерий Смирнова – Крамера – фон Мизеса, разработанный почти на 30 лет позже критерия Пирсона, превышает его по показателю вероятности ошибок первого и второго рода почти на 7%, при этом данный критерий предполагает конкретизацию значений параметров теоретической функции распределения, оценивание параметров по выборке, в свою очередь, приводит к существенному росту вероятности ошибки второго рода;

– критерий Андерсона – Дарлинга, являясь непараметрическим, используется для проверки простых гипотез о согласии эмпирического распределения с любым теоретическим распределением и является чувствительным к поведению функции распределения вероятностей на хвостах;

– критерий, разработанный Дэвидом, Хартли и Пирсоном, предполагает учет варьированности данных на основе анализа отношения размаха вариации и нормированного стандартного отклонения;

– критерий Шапиро – Уилка, созданный специально для анализа малых выборок и определяющий ее принадлежность к нормально распределенной генеральной совокупности, базируется на построении оптимальной линейной несмещенной оценки дисперсии при помощи метода максимального правдоподобия, дает приемлемые результаты при условии, что объем выборки превышает 20 измерений;

– критерий Хегази – Грина предусматривает использование аппроксимаций и является более мощным, чем критерий Шапиро – Уилка, но не позволяет автоматизировать аппроксимации и при исследовании малых выборок дает существенное смещение;

– критерий Васичека имеет своей основой следующее положение: энтропия нормального распределения превышает энтропию любого другого распределения с той же дисперсией (при этом для использования данного критерия нет необходимости в построении таблицы коэффициентов); он является одним из лучших с точки зрения математического моделирования, имеет удовлетворительную асимптотическую эффективность и более чувствителен к выбросам по сравнению с критерием Шапиро – Уилка;

– критерий Гири имеет высокую мощность (особенно при сравнении с критерием Дэвида – Хартли – Пирсона) и удобный алгоритм расчета и интерпретации с точки зрения его автоматизации, но его статистики могут быть признанными асимптоти-

чески нормальными только при условии, что объем выборки не менее 40 (при меньших объемах они нестабильны).

Спектр критериев согласия достаточно широк: в течение последних 120 лет разработано около 230 различных статистик, некоторая доля которых может быть применена посредством использования возможностей нейросетевых технологий.

Непараметрические статистики, используемые для анализа данных малого объема, могут быть объединены в следующие группы: ранговые и неранговые непараметрические статистики. Молодые исследователи относительно редко в своих работах обращаются к непараметрическим ранговым статистикам, основанным на эмпирических функциях распределения. На их основе разработаны следующие критерии:

– Колмогорова,

– Смирнова,

– Реньи,

– Крамера – Мизеса и др.

Критерий согласия Колмогорова предназначен для проверки гипотез о совпадении законов распределения двух эмпирических вариационных рядов. При этом отсутствует требование группировки дискретных данных или построения интервалов для непрерывных случайных величин. Данный критерий является чувствительным, так как позволяет отклонить нулевую гипотезу на более низком уровне значимости (по сравнению с медианным критерием). Критерий симметричности Смирнова используется для проверки гипотезы о симметричности распределения эмпирических данных относительно некоторого центра (в частности, среднего выборочного). Критерии Реньи и Крамера – Мизеса применяются для проверки гипотез об однородности двух независимых выборок.

Анализ научных трудов в данной области позволил выделить следующие особенности применения наиболее широко используемых критериев Колмогорова и Смирнова:

1. Критерий Смирнова применим для ранжированных данных (в частности, измеренных по порядковой шкале). Таблицы критических значений составлены для выборок малого объема.

2. При работе с выборками большого объема используются таблицы критических значений предельного распределения Колмогорова, предполагающие анализ данных, измеренных по интервальной шкале. В этом случае объем каждой из сопоставляемых выборок должен включать не менее 150 измерений [4].

Организация исследования, обеспечивающая репрезентативность малых выбо-

рок и признание статистических данных легитимными, предполагает обработку количественных признаков различными способами, среди которых научный интерес представляют следующие:

1. Линеаризующие преобразования.

2. Расчет «размера эффекта» (Effect Size – ES) [5].

Метод линеаризации функций, в математическом аспекте, предполагает построение прямой пропорциональной зависимости между значениями факторного и результирующего признака. Магистранты в большинстве случаев используют парные или множественные линейные регрессионные модели. При рассмотрении реальных процессов линейные модели часто оказываются несостоятельными, так как изучение малых массивов данных сопровождается регистрацией некоторых значений с низкой корреляционной связью и, как следствие, построением моделей с недостаточным интерполированием. Адекватную интерпретацию можно получить, используя построение полиномиальной регрессии. Однако в этом случае следует учитывать специфичность данных для правильного подбора показателей степеней, иначе возникает опасность перенасыщения модели. Линеаризация позволяет учитывать все зарегистрированные значения. Применимость метода к нелинейным моделям практически не имеет ограничений.

Современным аналогом критерия достоверности при проведении анализа эффективности тренировочного процесса, развития спортивно-технического мастерства спортсменов, результатов применения экспериментальных методик в спорте является ES – «величина эффективности». Методологически данный параметр противопоставляется пороговому значению уровня значимости ($p = 0,05$) для оценки достоверности коэффициента корреляции, и при его расчете используются следующие значения:

- объем выборки;
- коэффициент корреляции;
- t-критерий Стьюдента.

Для корреляционного анализа интерпретация расчетного значения ES (варьируется от 0 до 2) выглядит следующим образом:

- менее 0,1 – эффект отсутствует;
- от 0,1 до 0,3 – эффект низкий;
- от 0,3 до 0,5 – эффект средний;
- более 0,5 – эффект высокий.

При анализе выборок, содержащих от 3 до 20 элементов, неизбежны потери информации, уменьшение или устранение которых может быть достигнуто посредством метода точечных распределений (МТР). Согласно МТР, группировка данных заменяет-

ся рассмотрением каждого измерения в качестве центра распределения с некоторым заданным законом. Данный метод предусматривает предварительное определение эмпирической функции или эквивалентной выборки. Данный подход обеспечивает увеличение точности параметров малых выборок и корректность прогнозирования закона распределения генеральной совокупности. Реализация многомерного анализа на основе применения МТР позволяет исследователю создавать адекватные математические модели даже при относительно небольшом числе строк данных [6].

Заключение

Относительно небольшой объем выборочных данных, формируемых исследователями в сфере физической культуры и спорта, затрудняет статистический анализ на основе использования параметрических критериев. Необходимость подтверждения нормального распределения определяет выбор математической модели критерия согласия или критерия симметрии на основе заявленных уровней точности и достоверности.

Широко используемые непараметрические критерии имеют меньшую мощность в сравнении с ранговыми непараметрическими критериями. Алгоритм использования ранговых критериев, основанный на построении эмпирических функций распределения, обеспечивает их большую чувствительность при анализе малых выборок. В этом случае допускается обработка порядковых данных, что является существенным преимуществом в спортивных исследованиях. Большинство математических моделей расчета этих статистик могут быть автоматизированы, в открытых информационных ресурсах присутствуют онлайн-калькуляторы, применение которых является бесплатным и позволяет получить результаты в графической, формальной и словесной интерпретации.

Тотальный учет данных (в том числе с низким коэффициентом линейной корреляции) и определение ряда важных характеристик на основе учета формы регрессии при проведении корреляционного анализа малых массивов возможны в процессе линеаризации функций. Данная процедура может быть использована в случае нелинейной зависимости или при построении полиномиальной регрессии с недостаточной или некорректной аппроксимацией.

В современных исследованиях в области физической культуры и спорта ES (величина эффективности) вытесняет достоверность p , определенную Фишером для обоснования состоятельности интер-

претации корреляционного анализа малых выборок. Настоящий параметр находит широкое применение при оценке результатов тренировочного процесса, построенного на основе использования экспериментальных методик.

Использование современных информационных технологий позволяет разрабатывать нейроны, эквивалентные непараметрическим критериям, и применять метод точечных распределений, обеспечивающий корректность статистического анализа малых массивов данных.

Список литературы

1. Шаталов К.В., Кириллова А.В. Применение критерия Стьюдента для оценки результатов межлабораторных сравнительных испытаний // Эталоны. Стандартные образцы. 2016. № 1. С. 42–48.
2. Абдрахманова И.В., Лушик И.В. Особенности применения критерия Манна – Уитни для проверки статистических гипотез в физической культуре и спорте // Современные наукоемкие технологии. 2022. № 4. С. 128–132.
3. Иванов А.П., Иванов А.И., Малыгин А.Ю. Альбом из девяти классических статистических критериев для проверки гипотезы нормального или равномерного распределения данных малых выборок // Надежность и качество сложных систем. 2022. № 1 (37). С. 20–29.
4. Парыгина С.А. Об особенностях применения некоторых ранговых непараметрических критериев проверки гипотез // Вестник Башкирского университета. 2016. № 2. С. 252–255.
5. Воронов И.А. Линеаризующие преобразования и «размер эффекта» при проведении экспериментов с невысоким объемом выборки в области психологии спорта // Ученые записки университета имени П.Ф. Лесгафта. 2020. № 1 (179). С. 379–385.
6. Долгов Ю.А., Долгов А.Ю., Столяренко Ю.А. Анализ выборок малого объема и их применение // Вестник Приднестровского университета. Серия: Физико-математические и технические науки. Экономика и управление. 2013. № 3 (45). С. 79–90.