

4. Кривенцов А. А. Основы моделирования подготовленности спортсменов : учеб. пособ. / А. А. Кривенцов. – Малаховка : КИФК, 1990. – 87 с.
5. Кузнецов В. В. Модельные характеристики легкоатлетов / В. В. Кузнецов, В. В. Петровский, Б. Н. Шустин. – К. : Здоров'я, 1979. – 88 с.
6. Моделирование в спорте : метод. рекомендации / сост. Петриченко Д. В. – Минск, 1989. – 29 с.
7. Платонов В. Н. Общая теория подготовки спортсменов в олимпийском спорте / В. Н. Платонов. – К. : Олимп. л-ра, 1997. – С. 461.
8. Штофф В. А. Моделирование и философия / В. А. Штофф. – М. ; Л., 1966. – 19 с.
9. Штофф В. А. Роль моделей в познании / В. А. Штофф. – Л. : Из-во ЛГУ, 1963. – 63 с.
10. Шустин Б. Н. Моделирование и прогнозирование в системе спортивной тренировки / Б. Н. Шустин. – М. : СААМ, 1995. – С. 226–237.

Рецензент: докт. біол. наук, проф. Мицкан Б.М.

УДК 796/799:519.22/25
ББК 75

Ірина Іванишин

ДО ПРОБЛЕМИ ВИКОРИСТАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ В НАУКОВИХ ДОСЛІДЖЕННЯХ ІЗ ФІЗИЧНОЇ КУЛЬТУРИ

У статті проаналізовано основні помилки використання статистичних методів у наукових дослідженнях із фізичної культури. Проведено порівняння відносних частот застосування авторами україно- й англомовних журналів різних методів статистики.

Ключові слова: методи математичної статистики, наукові дослідження, фізичне виховання.

В статье проанализированы главные ошибки по применению статистических методов в научных исследованиях по физической культуре. Проведено сравнение относительных частот применения авторами украинско- и англоязычных журналов различных методов математической статистики.

Ключевые слова: методы математической статистики, научные исследования, физическое воспитание.

In the article the main mistakes respect to using of the statistical methods in scientific researches on physical education are analysed. The comparative analysis of relative frequency of different statistical methods usage national and foreign scientists is carried.

Key words: statistical methods, scientific researches, physical education.

Постановка проблеми та аналіз результатів останніх досліджень. Очевидним є факт, що ефективність розв'язання дослідницьких завдань значною мірою залежить від правильного планування експериментальних досліджень і професійної інтерпретації отриманих даних. Зокрема, важливою є адекватна статистична обробка матеріалів досліджень, достеменний аналіз усіх факторів, які впливають на валідність висновків за результатами експерименту [5; 8]. Про значення правильного розуміння статистичних результатів В.М.Заціорський [6] писав ще в 1989 р., відзначаючи їх некоректне використання та примітивне розуміння експериментальних методів у вигляді тільки однофакторного експерименту. У той же час провідні науковці зарубіжних країн застосовували багатфакторні експерименти [2; 3; 17; 19; 23].

Свого часу ця проблема була актуальною і в зарубіжній медико-біологічній і соціальній науках [9; 23]. Однак на даному етапі, виходячи із сучасних публікацій, відбулися суттєві зміни і рівень використання прикладної статистики став достатньо високим. На думку авторів [13; 15], це пояснюється декількома причинами: більш раннім і широким застосуванням інформаційних технологій у медико-біологічних і

педагогічних дослідженнях, більш глибокою статистичною освітою наукових кадрів і наявністю спеціалізованих підрозділів, які професійно виконують статистичну обробку результатів досліджень. За кордоном створено факультети та організації з біостатистики [18], розроблено рекомендації щодо представлення результатів наукових досліджень [21].

Проблема неякісного застосування методів математичної статистики в наукових публікаціях із фізичної культури вже привернула увагу ВАК України, що свідчить про важливість цієї проблеми.

Мета роботи – виявити основні помилки застосування статистичних методів у наукових дослідженнях із фізичної культури.

Методи та організація дослідження. У даному дослідженні проводився наукометричний аналіз статей із п'яти українських (“Вісник Прикарпатського університету. Фізична культура”, “Теорія і методика фізичного виховання і спорту”, “Актуальні проблеми фізичної культури і спорту”, “Молода спортивна наука України”, “Педагогіка, психологія та медико-біологічні проблеми фізичного виховання і спорту”) і трьох закордонних журналів (“Physical Education and Sport Pedagogy”, “Sport and Physical Education”, “Journal of Physical Education and Sport (JPES)”), а також дисертаційних досліджень у галузі фізичного виховання і спорту, опублікованих за останні 5 років. У загальному було проаналізовано 34 україномовні, 22 англійськомовні наукові статті і 24 дисертаційні роботи.

Оцінка статей і дисертацій проводилася з точки зору адекватності статистичних методів цілям і завданням дослідження, частоти застосування різних статистичних методів і коректності їх опису.

Результати дослідження. Як показав контент-аналіз, статистичні методи використовувалися в 82,4% вітчизняних статей, у зарубіжних публікаціях ця частка склала 95,5%. Проте в 53,6% публікацій, автори яких декларували використання статистичних методів, вони взагалі ніяк не описували використовувані ними методи статистики. Про те, що методи статистики застосовувалися, можна було здогадатися тільки за виразами типу “ $p < 0,05$ ” і ($M \pm m$), оскільки вирази типу “ $p < 0,05$ ” або “ $p < 0,01$ ” можуть бути отримані тільки при використанні конкретних статистичних критеріїв. У 23 (67,6%) статтях, де використовувався вираз “ $p < 0,05$ ”, не згадується статистичний критерій, який використовувався. Навпаки, 81,8% зарубіжних авторів повідомляють, за допомогою яких саме статистичних критеріїв проводилася перевірка конкретних гіпотез. Крім того, у більшості робіт не розкривається зміст величини “ p ”, при тому, що одна частина авторів (56,5%) розуміє під “ p ” рівень значущості, а інша – рівень довірчої ймовірності.

Для багатьох вітчизняних дисертаційних досліджень характерна наявність в описі методів статистики безглузвих і абсурдних виразів, що є комбінаціями так званих “*мемів*” – кліше, запозичених з інших робіт. Детальний аналіз цих мемів виявляє їх камуфляжний характер, унаслідок чого кілька років тому авторами [16] було введено поняття “*камуфляжного мему*”. За результатами наших досліджень та досліджень, проведених авторами [9; 15; 20], можна виділити три групи кліше:

1) у 45,8% всіх робіт просто констатується факт використання статистичних методів (кліше типу “Результати опрацьовані статистично”, “Обробку даних проводили статистичними методами”, “Результати опрацьовані методом (загальноприйнятими методами) варіаційної статистики”, “Результати досліджень опрацьовані статистично загальноприйнятим способом, відмінність встановлювали на рівні ймовірності не менше 95%” і т. п.);

2) для 37,5% (9/24) робіт характерна наявність більш конкретних деталей: статистичного критерію, одного чи двох статистичних параметрів; назви пакета прог-

рам (“Статистичну обробку проводили методом довірчих інтервалів”, “Достовірність значень визначали за t-критерієм Стьюдента”, “Отримані дані опрацьовували статистично за допомогою персонального комп’ютера” та ін.). При цьому, по-перше, у 66,7% (6/9) робіт немає посилань на авторів так званих “загальноприйнятих методик” або посилання не відповідають нумерації в списку використаних джерел; по-друге, згадування персонального комп’ютера не гарантує достовірності й надійності отриманих висновків. Крім марки комп’ютера необхідно вказати й конкретний тип статистичної програми (STATGRAPHICS, CSS, SPSS, SuperCalc, STATISTICA і Microsoft Excel), що було зроблено тільки у двох вітчизняних дисертаціях і 90,9% (20/22) зарубіжних публікацій;

3) третя група робіт (16,7%) містить достатньо неграмотні поєднання термінів і критеріїв (“Групи порівнювали при довірчому коефіцієнті 95%”, “Розрахунок середньої кількості $M \pm m$ проводили за методом Стьюдента”, “Статистичну обробку даних проводили за методом Стьюдента із застосуванням χ^2 -критерію”, “Кореляційний аналіз проводили шляхом порівняння двох груп за допомогою t -критерію”.

У більшості дисертацій (58,3%) опис використаних статистичних методів займає від 8–10 речень до максимум 2-х сторінок, у 41,7% дисертацій такий опис займав 5–7 сторінок. Детальне вивчення цих описів виявило ряд недоліків і протиріч, що свідчить про некомпетентність авторів.

По-перше, щодо позначень статистичних показників слід розрізняти позначення статистичних параметрів генеральних сукупностей, а також їх оцінок, отриманих за результатами вибірових досліджень. Перші позначаються грецькими літерами, а другі – літерами латинського алфавіту [4; 5; 9]. При виборі статистичних оцінок математики рекомендують керуватись вимогами незміщеності, спроможності та ефективності [13; 14]. Математики вважають, що найбільше цим вимогам відповідають оцінки середнього арифметичного (\bar{X}), дисперсії (S_x^2), стандартного відхилення (S_x) і коефіцієнтів кореляції Пірсона (r_{xy}), а тому в наукових роботах краще за все подавати середнє (\bar{X}) як міру центральної тенденції та стандартні відхилення (S_{xy}) як показники варіативності, а при оцінці взаємозв’язку – коефіцієнт кореляції (r_{xy}). Завжди потрібно вказувати обсяг вибірки (n), а для коефіцієнта кореляції – рівень значущості ($P < 0,05$, $0,01$ чи $0,001$).

По-друге, у 28 випадках (82,4%) стандартне відхилення перевищує середнє значення, тобто коефіцієнт варіації становить більше 100%. З математичної точки зору, якщо коефіцієнт варіації більший за 100%, то дані вибірки неоднорідні, а тому непридатні для будь-якого аналізу [1; 7].

По-третє, ряд авторів декларують використання різних пакетів програм, а в роботах вказують вираз “ $p < 0,05$ ”. У зв’язку із цим слід відзначити, що в комп’ютерних пакетах програм під час перевірки статистичних гіпотез обчислюється не тільки значення статистичного критерію, але й досягнутий рівень значущості для оцінки цього критерію. Тому під час використання прикладних програм слід вказувати конкретне значення досягнутого рівня значущості статистичного критерію.

По-четверте, у переважній більшості робіт українських науковців вибрані критерії не відповідають законам розподілу вибірових статистичних даних. У 70,8% робіт був використаний t-критерій, застосований для нормально розподілених статистичних даних, хоча перевірка нормальності розподілу не виконувалася. Слід зазначити, що при відхиленнях від нормального розподілу ознаки в сукупності для обробки даних використовуються інші статистичні методи [12; 22]. Крім того, не проводилася перевірка інших необхідних умов використання цього критерію,

зокрема відсутності мультимодальності розподілу, а також умови рівності дисперсій сукупностей, які вивчаються. Тому результати досліджень, на яких ґрунтуються висновки, не заслуговують довіри. Ще однією типовою помилкою є використання t-критерію Стьюдента для множинних порівнянь (у 39,3% робіт), хоча сучасним стандартом для перевірки гіпотези про належність декількох вибірок до однієї генеральної сукупності є дисперсійний аналіз [7; 10]. Грубою помилкою є застосування стандартного t-критерію Стьюдента до зв'язаних вибірок (типу “до–після”). Із цією метою слід застосовувати критерій Стьюдента для парних порівнянь чи критерій знакових рангів Вілкоксона [11].

По-п'яте, часто спостерігаються й помилки вивчення кореляційного зв'язку між кількісними даними. У 82,1% вітчизняних публікацій застосовують параметричний коефіцієнт кореляції Браує–Пірсона. Слід відзначити, що він оцінює силу лінійного зв'язку двох кількісних ознак, проте в жодній роботі не показано лінійності цих зв'язків. Найпростішим способом грубої оцінки зв'язку є побудова кореляційного поля. Визначивши вид можливого кореляційного зв'язку, необхідно знайти коефіцієнт кореляції, напрям зв'язку й достовірність отриманого коефіцієнта ($r=0,76$, $p<0,01$).

Щодо експериментальних досліджень слід відзначити, що в зарубіжних наукових роботах використовують багатофакторні експерименти, які дозволяють у більш стислі терміни знайти оптимальне рішення у виборі тренувальних режимів чи методів навчання [2; 10; 17].

На рис. 1 наведено результати порівняння відносних частот використання статистичних методів в україномовних і англомовних наукових публікаціях.

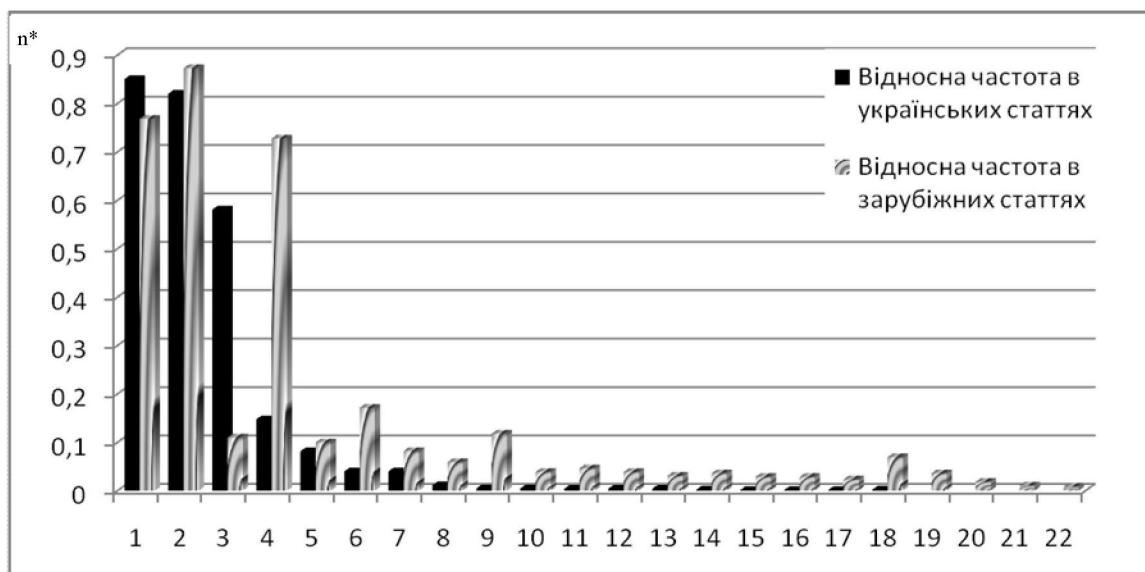


Рис. 1. Розподіл відносних частот використання статистичних методів в українських і зарубіжних статтях: 1) оцінка дескриптивних статистик; 2) використання виразів типу “ $p < \dots$ ”; 3) перевірка гіпотез про рівність середніх t-критерієм Стьюдента; 4) перевірка гіпотез з використанням непараметричних критеріїв; 5) оцінка парних коефіцієнтів кореляції Пірсона; 6) дисперсійний аналіз; 7) множинний регресійний аналіз; 8) нелінійний регресійний аналіз; 9) оцінка коефіцієнтів рангової кореляції; 10) факторний аналіз; 11) дискримінантний аналіз; 12) аналіз таблиць спряженості; 13) кластерний аналіз; 14) методи теорії планування експериментів; 15) послідовний аналіз Вальда; 16) аналіз часових рядів; 17) оцінка відношення шансів; 18) перевірка законів розподілу; 19) оцінка частинних коефіцієнтів кореляції; 20) коваріаційний аналіз; 21) багатомірне шкалювання; 22) лонглінійний аналіз

Загальна кількість статистичних методів і критеріїв, використаних в англо-мовних статтях, склала більше 50. Із цього різноманіття ми вибрали тільки найбільш відомі, об'єднавши їх у 22 групи. Після оцінки відносних частот використання того або іншого методу в публікаціях, проводилася перевірка статистичної гіпотези про рівність цих пропорцій з використанням Z -критерію [7]. Для 20 груп досягнуті рівні значущості " p " мали значення набагато менші від 5%, що дозволяє стверджувати статистично значущу відмінність цих частот.

Тільки для парних коефіцієнтів кореляції Пірсона ($p=0,1445$) і дескриптивних статистик ($p=0,0961$) були прийняті нульові гіпотези про рівність частот використання цих методів у статтях. Для решти всіх груп порівняння ці частоти відрізнялися статистично значущо ($p \ll 0,05$).

Як видно з рис. 1, зарубіжні дослідники частіше використовують багатомірні методи, ніж автори українських статей. Це говорить про те, що статистична парадигма зарубіжних дослідників містить виражений багатомірний підхід до дослідження об'єктів, які вивчаються. Однак слід відзначити, що і в зарубіжних наукових дослідженнях мало використовуються методи, розроблені за останні 10–15 років, а саме: багатомірне шкалювання та лонглінійний аналіз [19; 23].

Велика кількість вищеназваних недоліків у текстах наукових статей та методів обробки результатів досліджень свідчить як мінімум про чотири вірогідні причини цього явища.

Перша причина – низька статистична культура дослідників.

Друга причина – відсутність у структурах НДІ і вузів спеціалізованих лабораторій біостатистики, покликаних забезпечити дослідникові кваліфікований статистичний аналіз спостережень.

Третя причина – відсутність галузевої нормативної бази, що регламентує етап статистичного аналізу.

Четверта причина – відсутність кваліфікованої статистичної експертизи в редакціях журналів, дисертаційних й експертних радах ВАК України, що говорить про відсутність у них фахівців, які володіють даними технологіями.

Висновок

Проведений аналіз показує недостатній рівень підготовленості наукових кадрів із питань статистичної обробки результатів наукових досліджень. Відсутність цих знань призводить до систематичних помилок у публікаціях, які дають підстави вважати результати цих досліджень недостовірними. Можна відзначити суттєве відставання наших науковців від зарубіжних у питаннях використання багатомірних методів математичної статистики і теорії планування експерименту.

1. Баева Т. Е. Применение статистических методов в педагогическом исследовании : учеб.-метод. пособие для студ. и аспирантов ин-тов физ. культуры / Т. Е. Баева, С. Н. Бекасова, В. А. Чистяков ; под общей ред. М. В. Прохоровой ; НИИХ Санкт-Петербургского ун-та. – С. Пб., 2001. – 82 с.
2. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей / Бююль Ахим, Ефель Петер ; [пер. с нем.]. – С. Пб. : ДиаСофтЮП, 2001. – 608 с.
3. Гласс Дж. Статистические методы в педагогике и психологии / Дж. Гласс, Дж. Стэнли. – М. : Прогресс, 1976. – 496 с.
4. Дубина И. Н. Математические основы эмпирических социально-экономических исследований : учебное пособие / Дубина И. Н. – Барнаул : Изд-во Алт. ун-та, 2006. – 267 с.
5. Закирьянов К. Х. Статистические и экспериментальные методы в педагогике, психологии и физической культуре : [учеб. пособ.] / К. Х. Закирьянов, Л. И. Орехов. – Алматы : КазАСТ, ННПЦФК, 2002. – 112 с.

6. Заціорський В. М. Осторожно: статистика! / В. М. Заціорський // Теорія і практика фізическої культури. – 1989. – № 2. – С. 52–55.
7. Измерения и вычисления в спортивно-педагогической практике : [учебн. пособ. для вузов физ. культ.] / В. П. Губа, М. П. Шестаков, Н. Б. Бубнов, М. П. Борисенко. – М. : Физкультура и спорт, 2006. – 220 с.
8. Кизько А. П. Принципиальные вопросы корректности результатов экспериментального исследования в области физического воспитания и спорта / А. П. Кизько // Теория и практика физической культуры. – 2004. – № 1. – С. 59–61.
9. Ланг Т. Двадцать ошибок статистического анализа, которые вы сами можете обнаружить в биомедицинских статьях / Т. Ланг // Международный журнал медицинской практики. – 2005. – № 1. – С. 21–31.
10. Ланг Т. А. Как описывать статистику в медицине : [руководство для авторов, редакторов и рецензентов] / Т. А. Ланг, М. Сесик ; [пер. с англ. ; под ред. В. П. Леонова]. – М. : Практическая медицина, 2010. – 488 с.
11. Лапач С. Н. Статистика в науке и бизнесе / Лапач С. Н., Чубенко А. В., Бабич П. Н. – К. : МОРИОН, 2002. – 640 с.
12. Орехов Л. И. Управление, контроль, измерение, статистические и экспериментальные методы в педагогике, психологии и физической культуре : [учеб. пособ.] / Орехов Л. И., Караваева Е. Л., Асмолова Л. А. – Алматы : КазАСТ, ННПЦФК, ВКТУ, 2004. – 169 с.
13. Орехов Л. И. О необходимости соответствия статистических и экспериментальных методов современным требованиям / Л. И. Орехов, Е. Л. Караваева // Теория и практика физической культуры. – 2005. – № 3. – С. 46–49.
14. Орлов А. И. Эконометрика / Орлов А. И. – М. : ЭКЗАМЕН, 2004. – 576 с.
15. Применение статистики в медицине и биологии : анализ публикаций 1990–1997 гг. / В. П. Леонов, П. В. Ижевский ; ГНЦ РФ, Ин-т биофизики. – М., 1998. – 10 с. – Деп. в ВИНТИ 23.01.98Х9 179-В98.
16. Смирнов Ю. И. Спортивная метрология / Ю. И. Смирнов, М. М. Полевщиков. – М. : Академия, 2000. – 232 с.
17. Факторный, дискриминантный и кластерный анализ / [Ким Дж.-О., Мьюллер Ч. У., Клекка У. Р. и др.]. – М. : Финансы и статистика, 1989. – 215 с.
18. Curriculum guidelines for undergraduate programs in statistical science [Electronic resource] : American Statistical Association. – 2002. – Available at http://www.amstat.org/education/Curriculum_Guidelines.html.
19. Gifi A. Nonlinear multivariate analysis / Gifi A. – New York : John Wiley & Sons, 1990. – 579 p.
20. Hargrove W. W. Pseudoreplication: a sine qua non for regional ecology / W. W. Hargrove, J. Pickering // Landscape Ecology. – 1992. – V. 6, № 4. – P. 251–258.
21. Hopkins W. G. Guidelines on style for scientific writing / W. G. Hopkins // Sportscience. – 1999. – 3 (1). – Available at <http://www.sportsci.org/jour/9901/wghstyle.html>.
22. Research in physical education, exercise science and sport / [edit. by Charles F. Cicciarella]. – 2nd ed. – 2007. – 350 p.
23. Tracey N. D. Multivariate control charts for individual observations / N. D. Tracey, J. Young, R. L. Mason // Journal of Quality Technology. – 1992. – Issue 2. – P. 88–95.

Рецензент: канд. фіз.-мат. наук, доц. Дмитришин Р.І.