

УДК 796:006.91

Бишевец Н.Г., Сергиенко К.М., Синіговец І.В, Бровина В.С.

**СУЧАСНІ МЕТОДИ АНАЛІЗУ ДАНИХ В СПОРТІ НА ПРИКЛАДІ  
ПОКАЗНИКІВ ДОВЖИНИ ТІЛА ЧОЛОВІЧИХ ЗБІРНИХ КОМАНД  
СВІТУ З ВОЛЕЙБОЛУ**

**Анотація.** У статті підіймається питання щодо широких можливостей, які відкривають комп'ютерні технології при обробці результатів наукових досліджень і пропонується перевірку нормальності розподілу спостережуваних даних, встановлення рівності генеральних дисперсій, а також перевірку гіпотези про рівність середніх проводити за допомогою засобів програми Excel.

**Ключові слова.** Підготовка, обробка, аналіз, інформація, дисперсія, розподіл, критерій, показник, асиметрія, ексцес, метод, формула.

Бышевец Н.Г., Сергиенко К.Н., Синиговец И.В, Бровина В.С.

**Современные методы анализа данных в спорте на примере показателей  
длины тела мужских сборных команд мира по волейболу.**

В статье поднимается вопрос о широких перспективах, которые открывают компьютерные технологии при обработке результатов научных исследований и предлагается проверку нормальности распределения наблюдаемых данных, установления равенства генеральных дисперсий, а также проверку гипотезы о равенстве средних проводить с помощью средств программы Excel.

**Ключевые слова.** Подготовка, обработка, анализ, информация, дисперсия, распределение, критерий, показатель, асимметрия, эксцесс, метод, формула.

Byshevets. N.G., Sergienko, K.N., Sinigovets I.V., Brovina V.S.

**Modern methods of the analysis of the data in sports on the examples  
increased results of men's combined volleyball teams of the world.**

In the article draws attention to the wide possibilities, which open computer technologies in processing results of scientific researches in industry and it is offered verification of normal distribution of the observed data, established the equality of general dispersions, and also verification of hypothesis about the equality of middle to conduct by means of the program Excel.

**Key words.** Preparation, treatment, analysis, information, dispersion, distribution, criterion, index, asymmetry, excess, method, formula.

**Постановка проблеми та її зв'язок з важливими науковими чи практичними завданнями.** Підготовка висококваліфікованих спортсменів передбачає застосування науково-обґрунтованих підходів до побудови тренувального процесу, систематичний контроль за самопочуттям спортсменів, аналіз динаміки спортивних досягнень, визначення оптимальних навантажень тощо. Сучасні реалії розвитку спортивної науки полягають в тому, що результати тренувальної і змагальної діяльності, а також дані медичних обстежень підлягають математичній та статистичній обробці, що, на думку фахівців, потребує відповідної математичної підготовки спеціаліста галузі фізичного виховання та спорту [6]. Дійсно, недостатні знання з математичної статистики зводять нанівець зусилля фахівців, направлені на обґрунтування положень та висновків, отриманих в ході наукових досліджень і унеможливають застосування результатів експериментів в практиці фізичного виховання та спорту.

**Аналіз останніх досліджень і публікацій.** Усунення прогалин в математико-статистичній підготовці студентів фізкультурних вузів проводиться в кількох напрямках. З одного боку, фахівці пропонують розширення спеціальних знань студентів за рахунок збільшення навчального навантаження шляхом впровадження додаткових дисциплін, направлених на оволодіння сучасних методів аналізу спортивних даних. Так, Л.П. Сергієнко сформульовано концепцію та зміст програми навчального курсу «Методи наукових досліджень у фізичній культурі», який передбачає вивчення студентами теоретичних основ організації і проведення наукових досліджень у фізичній культурі, спорті і фізичній реабілітації і включає вивчення методів математичної статистики в спорті [7]. З іншого боку, огляд науково-методичної і спеціальної літератури показав, що наразі з метою спрощення сприйняття математичних відомостей, автори пропонують обробку даних спортивних досліджень проводити з використанням комп'ютерних

технологій. Існують свідчення про експериментальне впровадження в учбовий процес студентів фізкультурних вузів комп'ютерної програми по дисципліні "Комп'ютерна техніка і математичні методи в спорті" [8]. В свою чергу С.В. Калитка зі співавторами провів систематизацію найбільш поширених методів аналізу даних, що застосовуються у фізичному вихованні та спорті та визначив етапи здійснення обробки результатів досліджень за допомогою програми Excel [4].

При обробці спортивних даних, перевірка гіпотез про рівність середніх часто здійснюється дослідниками за допомогою t-критерію Стьюдента. Однак аналіз наукових праць показав, що наряду з докладним розглядом етапів застосування t-критерію Стьюдента, відсутні повні відомості про умови його використання [3, 4, 6]. В наслідку такого однобокого підходу до аналізу даних відбувається викривлення результатів досліджень і хибних умовиводів. На некоректне використання t-критерію Стьюдента при обробці біомедичних даних вказує В.П. Леонов [5]. Провівши аналіз численної кількості наукових робіт в області біології та медицини, В.П. Леонов наочно продемонстрував некоректність його застосування у переважній кількості робіт та прийшов до висновку про ненадійність отриманих результатів. За свідченнями автора, в більшості підручників, в яких розглядається t-критерій Стьюдента, не приділяється увага обмеженням цього критерію, а також не вказується на наслідки їх порушення [5].

На нашу думку, у наукових роботах в галузі фізичного виховання і спорту склалась не менш несприятлива ситуація, що і обумовило вибір теми дослідження.

**Мети та завдання роботи.** Дослідження було направлене на вивчення сучасних методів обробки спортивної інформації на прикладі ростових показників найкращих волейбольних збірних команд світу. В ході дослідження нами були використані такі методи дослідження як аналіз літературних джерел та методи математичної статистики. Об'єктом дослідження стали методи обробки даних наукових та методичних

досліджень в галузі фізичного виховання і спорту, а предметом дослідження – специфіка встановлення рівності середніх показників.

**Виклад основного матеріалу дослідження.** Ґрунтовне вивчення обмежень при застосуванні t-критерію Стюдента показало наявність двох вимог його використання: нормальність розподілу вибірових даних та рівність генеральних дисперсій, якими часто нехтують дослідники [5]. Отже, для вживання t-критерію Стюдента необхідно дотримуватися наступних умов:

1. Нормальність розподілу вибірових даних.
2. Рівність генеральних дисперсій в порівнюваних групах (умова однорідності дисперсій або умова гомоскедастичності).

Розглянемо етапи встановлення нормальності розподілу вибірових даних у програмі Excel на прикладі ростових показників найкращих волейбольних збірних команд світу.

Одним з розповсюджених методів наближеної перевірки гіпотези про нормальність розподілу є використання показників асиметрії і ексцесу. **Асиметрія** — це показник, що відображає міру скошеності кривої розподілу емпіричних даних у порівнянні з диференціальною функцією нормального розподілу. **Ексцес** — показник, що відображує наявність піку кривої розподілу в порівнянні з диференціальною функцією нормального розподілу. Значення асиметрії (A) і ексцесу (E) розраховуються за наступними формулами:

$$A \approx \frac{1}{nS^3} \sum_{i=1}^n (x_i - \bar{x})^3 ; \quad E \approx \frac{1}{nS^4} \sum_{i=1}^n (x_i - \bar{x})^4 - 3 \quad (1).$$

У програмі Excel є вбудовані статистичні функції: для розрахунку A застосовують функцію СКОС, а для визначення E – функцію ЕКСЦЕСС. Отже, на першому кроці аналізу даних, після їх внесення в електронну таблицю, у комірки B22 та B23 вносимо відповідні формули СКОС(B3:B20) і ЕКСЦЕСС(B3:B20).

Для того, щоб відповісти на питання про нормальність розподілу показників довжини тіла висококваліфікованих волейболістів перевіримо справедливість наступних співвідношень:

$$|E| \leq 5\sqrt{D(E)}; \quad |A| \leq 3\sqrt{D(A)} \quad (2).$$

Якщо дані співвідношення справджуються, розподіл вважатимемо нормальним, а в іншому випадку гіпотезу нормальності відхилимо. За допомогою вбудованої функції ABS() визначимо абсолютне значення (модуль) А і Е. При цьому розрахунок дисперсії асиметрії та ексцесу визначимо за формулами:

$$D(A) = \frac{6(n-1)}{(n+1)(n+3)}; \quad D(E) = \frac{24(n-2)(n-3)n}{(n-1)^2(n+3)(n+5)} \quad (3).$$

У комірку B26 введемо наступну формулу:

$$=ЕСЛИ(ABS(СКОС(B2:B19))<=3*КОРЕНЬ((6*(\$A\$20-1))/(\$A\$20+1)/(\$A\$20+3))+ И(ABS(ЭКСЦЕСС(B2:B19)) <=5*КОРЕНЬ(24*(\$A\$20-2)*(\$A\$27-3)*\$A\$20/(\$A\$20-1)/(\$A\$20-1)/(\$A\$20+3)/(\$A\$20+5))); "так"; "ні") \quad (4).$$

В результаті проведених дій було встановлено, що показники довжини тіла волейболістів, які входять до кращих збірних команд світу, підпорядковується нормальному закону (рис.1).

Якщо аналізу підлягають дві групи даних, то рівність генеральних дисперсій в порівнюваних групах можна встановити за допомогою процедури *Двухвыборочный F-тест для дисперсии* вбудованого пакету *Анализ данных* програми Excel з подальшим встановленням достовірності розходжень між груповими середніми за допомогою процедури *Двухвыборочный t-тест для с одинаковыми дисперсиями* або *Двухвыборочный t-тест для с различными дисперсиями* в залежності від результату дисперсійного аналізу.

Як вказують фахівці, розповсюдженою помилкою при перевірці гіпотези про рівність середніх є неправомірність використання t-критерію Стьюдента для перевірки гіпотез про рівність трьох і більше групових середніх. В цьому випадку спеціалісти рекомендують використовувати загальну лінійну

модель, реалізовану в процедурі *Однофакторный дисперсионный анализ*, іншими словами – метод ANOVA, головною умовою застосування є також нормальність розподілу результативної ознаки в розглядуваних сукупностях [1]. Результати однофакторного дисперсійного аналізу представлені на рисунку (рис. 2).

### **Інтерпретація результатів дисперсійного аналізу:**

**df** – кількість ступенів вільності; всього 197, з них 10 пов'язано зі змінами показників довжини тіла між спортсменами різних команд, інші 187 – зі змінами показників довжини тіла між спортсменами однієї команди;

**SS** – загальна оцінка мінливості склала  $8882,4 \text{ см}^2$ . При цьому сума квадратів відхилень середньої довжини тіла спортсменів від загальної середньої склала  $1853,7 \text{ см}^2$ , а сума квадратів відхилень середньої довжини тіла спортсменів від групових середніх –  $7028,7 \text{ см}^2$ .

**MS** – випадкова варіація ознаки, тобто розсіяння показників довжини тіла спортсменів в одній команді становить  $37,59 \text{ см}^2$ , а систематична варіація, тобто розсіяння середніх показників довжини тіла спортсменів з різних команд –  $185,4 \text{ см}^2$ .

**F** (F-критерій) складає 4,93 і більший за F-критичне. При даних об'ємах вибірок цьому значенню F-критерію відповідає досягнутий рівень значущості  $p=0,00$ , що підтверджує статистичну значущість відмінності середньої довжини тіла волейболістів з різних країн ( $p<0,05$ ).

Таким чином, довжина тіла волейболістів достовірно відрізняється між командами різних країн світу. У середньому найменша довжина тіла спостерігається у спортсменів з Японії, найбільша – у спортсменів з Росії, а довжина тіла, найближча до загальної середньої – у спортсменів з Бразилії.

**Висновки.** В результаті проведених досліджень нами отримані наступні висновки:

- на сьогодні авторами не ставиться під сумнів необхідність достатніх математико-статистичних знань у фахівців галузі фізичного виховання і спорту;

- поглиблення знань майбутніх фахівців відбувається як за рахунок впровадження спеціальних навчальних курсів так і за рахунок застосування комп'ютерних технологій при обробці спортивних даних;
- одним із найпоширеніших критеріїв, що застосовуються під час математико-статистичної обробки матеріалів наукової і методичної діяльності фахівця галузі фізичного виховання і спорту, є t-критерій Стьюдента, однак нехтування дослідниками необхідними умовами призводить до некоректності застосування авторами даного критерію та викликає серйозні сумніви у задекларованих висновках;
- умовами застосування t-критерію Стьюдента являється нормальний розподіл досліджуваних показників, а також рівність генеральних дисперсій в порівнюваних групах;
- перевірка гіпотез про рівність трьох і більше групових середніх передбачає застосування методу ANOVA;
- проведений аналіз показав, що середня довжина тіла висококваліфікованих волейболістів достовірно відрізняється в залежності від країни, яку вона представляє ( $p < 0,05$ ).

В ході дослідження нами, запропоновано проводити перевірку означених умов, а також проводити порівняння групових середніх за допомогою можливостей програми Excel на прикладі ростових показників найкращих волейбольних збірних команд світу.

**Перспективи подальших досліджень у даному напрямку.** Подальше дослідження планується направити на вивчення можливостей автоматизації застосування непараметричних критеріїв в ході аналізу спортивної інформації.

### **Література.**

1. Ермолаев О.Ю. Математическая статистика для психологов : [учебник] / О.Ю. Ермолаев. – Москва : Флинта. - 2003. - 336 с.
2. Железняк Ю. Д. Основы научно-методической деятельности в физической культуре и спорте : учеб. пособие [для студ. высш. пед. учеб.

заведений] / Ю. Д. Железняк, П. К. Петров. - Москва : Издательский центр «Академия». - 2002. - 264 с.

3. Заляжных В. В. Статистические методы контроля и управления качеством [Электронный ресурс] : учебник / Заляжных В. В. – 2010. -

Режим доступа к учебнику : <http://www.statmetkach.com/index.html>

4. Калитка С.В. Математико-статистическая обработка материалов научных исследований в студенческих научных работах / С.В. Калитка, А.Н. Ситовский, Л.С. Бойко // Молодіжний науковий вісник. - 2007. - С. 9 - 12.

5. Леонов В.П. Ошибки статистического анализа биомедицинских данных / В.П. Леонов // Международный журнал медицинской практики/ - 2007. - № 2. - С.19-35.

6. Містулова Т. Є. Математичні методи в теорії та практиці спорту : [навч. посіб.] / Т. Є. Містулова. – Київ : Наук. світ. - 2004. - 90 с.

7. Сергієнко Л. П. До проблеми планування вивчення навчальної дисципліни «Методи наукових досліджень у фізичній культурі» / Л. П. Сергієнко // Теорія та методика фізичного виховання. - 2009. - № 2. – С. 26 – 34.

8. Филенко Л. Компьютерные обучающие программы при самостоятельной подготовке студентов высших учебных заведений физической культуры / Л. Филенко, Ю. Горбатенко // Теорія та методика фізичного виховання. - 2008. - № 4. - С. 70 - 73.



	A	B	C	D	E	F	G	H	I	J	K	L
1												
2		<b>Україна</b>	<b>Китай</b>	<b>Бразилія</b>	<b>Алжир</b>	<b>Росія</b>	<b>Японія</b>	<b>США</b>	<b>Болгарія</b>	<b>Куба</b>	<b>Аргентина</b>	<b>Італія</b>
3	1	204	196	199	193	216	185	203	206	193	202	202
4	2	202	193	209	192	196	190	193	194	191	196	197
5	3	191	200	199	188	200	193	198	200	198	190	201
6	4	198	188	193	192	201	194	206	189	195	204	178
7	5	200	199	200	196	198	187	193	192	199	195	202
8	6	206	198	184	195	203	175	198	201	197	180	191
9	7	203	192	192	197	217	197	198	197	193	200	194
10	8	192	190	200	193	200	197	191	198	208	202	197
11	9	200	195	200	194	197	196	206	191	191	195	194
12	10	202	202	198	193	201	200	193	198	196	190	197
13	11	194	191	192	188	198	187	208	192	199	194	190
14	12	203	192	195	192	194	193	198	192	192	200	190
15	13	208	200	203	187	205	190	203	198	198	188	196
16	14	206	200	201	194	200	208	198	196	206	186	197
17	15	197	195	183	186	204	196	200	202	200	190	196
18	16	197	204	199	188	187	192	193	212	200	183	202
19	17	192	197	192	190	200	168	196	200	198	196	205
20	<b>18</b>	200	192	202	191	204	195	206	216	197	195	198
21	<b>196,4</b>	<b>199,7</b>	<b>195,8</b>	<b>196,7</b>	<b>191,6</b>	<b>201,2</b>	<b>191,3</b>	<b>198,9</b>	<b>198,6</b>	<b>197,3</b>	<b>193,7</b>	<b>195,9</b>
22	асиметрія (А)	-0,29	0,08	-0,61	-0,22	0,75	-1,03	0,24	1,07	0,76	-0,38	-1,34
23	ексцес (Е)	-0,84	-0,94	0,56	-0,85	1,91	2,16	-1,14	1,06	0,73	-0,49	3,15
24	модуль А	0,29	0,08	0,61	0,22	0,75	1,03	0,24	1,07	0,76	0,38	1,34
25	модуль Е	0,84	0,94	0,56	0,85	1,91	2,16	1,14	1,06	0,73	0,49	3,15
26	Розподіл нормальний?	так	так	так	так	так	так	так	так	так	так	так

Рис. 1. Встановлення нормальності розподілу довжини тіла висококваліфікованих волейболістів у програмі Excel

Однофакторный дисперсионный анализ				
ИТОГИ				
Группы	Счет	Сумма	Среднее	Дисперсия
Україна	18	3595	199,7	26,1
Китай	18	3524	195,8	20,3
Бразилія	18	3541	196,7	42,3
Алжир	18	3449	191,6	10,1
Росія	18	3621	201,2	48,6
Японія	18	3443	191,3	81,2
США	18	3581	198,9	28,4
Болгарія	18	3574	198,6	51,2
Куба	18	3551	197,3	21,4
Аргентина	18	3486	193,7	45,5
Італія	18	3527	195,9	38,3

Дисперсионный анализ						
Источник вариации	SS	df	MS	F	P-Значение	F критическое
Между группами	1853,697	10	185,37	4,93	0,00	1,88
Внутри групп	7028,667	187	37,59			
Итого	8882,364	197				

Рис.2. Результати однофакторного дисперсійного аналізу

## Авторські довідки

Синіговець Ігор Васильович	кандидат наук з фізичного виховання та спорту, доцент кафедри спорту Чернігівського національного педагогічного університету імені Т.Г. Шевченка (м. Чернігів)
Бишевець Наталія Григорівна	старший викладач, кафедри вищої математики та економіко-математичного моделювання Академії муніципального управління (м. Київ).
Сергієнко Костянтин Миколайович	кандидат наук з фізичного виховання і спорту, доцент кафедри кінезіології Національного університету фізичного виховання і спорту України (м. Київ)
Бровіна Вікторія Сергіївна	викладач кафедри комп'ютерних наук Чернігівського державного інституту інформації, бізнесу і права (м. Чернігів)