

## ОБ ИСПОЛЬЗОВАНИИ ЭКСПЕРТНЫХ ОЦЕНОК В ИССЛЕДОВАНИЯХ В СФЕРЕ ФИЗИЧЕСКОЙ КУЛЬТУРЫ И СПОРТА

**В.К. Пономаренко**, канд. физ.-мат. наук, доцент, **Ю.О. Волков**,  
Белорусский государственный университет физической культуры,  
**К. Городилов**,  
танцевальный клуб “Kreedo Dance”, г. Таллинн, Эстония

*В статье рассматривается пример применения коэффициента конкордации Кендалла для выявления степени согласованности мнений арбитров (экспертов). Обсуждаются некоторые возможные проблемы.*

## ON THE USE OF A PEER REVIEW IN RESEARCH IN THE SPHERE OF PHYSICAL CULTURE AND SPORT

*An example of Kendall's concordance coefficient to identify the coherence degree of referees' (experts) opinions is considered in the article. Some potential problems are discussed.*

Наука начинается с тех пор, как начинают измерять. Точная наука без меры немислима.  
Д.И. Менделеев

**Введение.** Если задуматься над смыслом слова измерять (или измерение), то несложно заметить, что первичным является понятие меры – эталона, исходя из которого и в терминах которого мы получаем итоговый результат. Очевидно также, что вследствие этого полученные данные имеют количественный характер. Вместе с тем происходящие в повседневной жизни события (явления, происшествия и т. д.) носят преимущественно качественный характер, и человек к этому хорошо приспособлен. «Как показали многочисленные опыты, человек более правильно (и с меньшими затруднениями) отвечает на вопросы качественного, например, сравнительного, характера, чем количественного. Так, ему легче сказать, какая из двух гирь тяжелее, чем указать их примерный вес в граммах» [1, с. 80]. До недавнего времени при рассмотрении качественных явлений использовались исключительно термины оценивать или оценка. Неслучайно говорят об оценке знаний, а не об их измерении. Однако примерно с середины прошлого века в связи с все нарастающим проникновением математических, количественных, методов в психологию, социологию, медико-биологические, педагогические и т.п. исследования термин «измерение» приобрел более широкий смысл, включив в себя также количественные характеристики качественных явлений. Подтверждением тому может служить, например, появление термина «педагогическая квалиметрия». В такой обобщенной, общенаучной теории измерений исследование (математическое моделирование) реального явле-

ния начинается с установления типа шкал, к которым принадлежат переменные, характеризующие исследуемое явление. В цитированной выше работе [1] рассматриваются шесть шкал измерений: шкала наименований (иначе называемая номинальной или номинативной [2, с. 12]), порядковая шкала, шкала интервалов, шкала отношений, шкала разностей и абсолютная шкала. Из них наиболее употребительными являются первые четыре. Шкала наименований и порядковая шкала предназначены для работы с качественными признаками, шкала интервалов и шкала отношений – с количественными.

В любом научном исследовании методы математической обработки результатов, полученных вследствие тех или иных измерений, проведенных в ходе исследования, во многом предопределяются свойствами обрабатываемых результатов-данных, их принадлежностью к тем или иным из перечисленных выше шкал.

Так какие же виды данных могут возникать при проведении исследований в сфере физической культуры и спорта? Ответ на этот вопрос определяется степенью участия человека в оценке (измерении) спортивного результата.

Степень эта очень вариативна:

1) от технически возможного неучастия (например, при фиксировании результатов в таких видах спорта, как плавание или беговые дисциплины легкой атлетики), когда измеряемые величины характеризуются количественно, а результаты фиксируются, как правило, исключительно техническими средствами;

2) до полной противоположности (например, при оценке результатов в обоих видах гимнастики, акробатике, фигурном катании, прыжках в воду, синхронном плавании, а также при нахождении экспертных оценок того или иного факта или явления), когда не используются никакие технические средства и устройства, а сами результаты качественного характера.

В первом случае измеряемые величины выражаются в соответствующих единицах и долях единиц Международной системы (СИ). К полученным числовым наборам можно применять любые статистические методы, обусловленные смыслом исследования и использовать все статистические функции, необходимые для его осуществления.

Во втором случае результаты выражаются в баллах, а это означает использование порядковой шкалы, и к полученным наборам значений применимы только непараметрические методы статистики [2, 3].

**Основная часть.** Поскольку, как отмечалось выше, результаты судейства во многих видах спорта, а также итоги различных опросов выражаются в баллах (являются своего рода экспертными оценками), напомним основные общепринятые методы математической обработки экспертных оценок:

- 1) проверка согласованности мнений экспертов и
- 2) усреднение их мнений.

При отсутствии согласованности производится разбиение ответов экспертов на группы по сходству мнений, и усреднение проводится внутри каждой группы [1, 4].

Рассмотрим применение (в названном ранее порядке) методов математической обработки экспертных оценок на примере задачи из дипломной работы третьего из авторов данной статьи.

Задачи дипломного исследования, в частности, включали в себя:

а) сбор данных о профессионально-значимых качествах личности руководителя танцевального клуба “Kreedo Dance“ по оценке тренеров, ему подчиненных;

б) математическую обработку (экспертизу) полученных данных и выводы, сделанные на ее основе.

Для решения задачи пункта а) проводилось анкетирование тренеров клуба. В связи с тем что при выборе качеств руководителя, подлежащих анкетированию, затруднительно определить, какое из этих качеств более значимо по сравнению с другими, было принято предположение об их равной значимости. Каждое профессионально-значимое качество оценивалось в баллах от 1 до 5. Приведенная ниже таблица 1, содержащая ответы тренеров-респондентов, выраженные в баллах, на анкету о профессионально-значимых качествах личности руководителя организации явилась основным практическим результатом в решении задачи пункта а), а ее содержимое – исходными данными для решения задачи пункта б).

Таблица 1. – Оценка профессионально-значимых качеств руководителя тренерами-респондентами

Профессионально-значимые качества личности	Ответы тренеров в баллах									
	Эксперт 1, 32 года	Эксперт 2, 24 года	Эксперт 3, 27 лет	Эксперт 4, 50 лет	Эксперт 5, 55 лет	Эксперт 6, 24 года	Эксперт 7, 23 года	Эксперт 8, 24 года	Эксперт 9, 23 года	Эксперт 10, 27 лет
1. Компетентность	4	4	5	4	5	5	4	4	4	4
2. Знание действительности	3	4	4	3	4	3	3	4	4	4
3. Понимание целей организации	4	5	5	4	5	5	5	4	5	5
4. Умение видеть проблемы	5	5	5	5	5	5	4	5	4	4
5. Умение видеть ситуацию	4	5	5	4	4	5	5	4	3	4
6. Восприимчивость к изменениям	3	4	3	3	3	4	3	3	3	4
7. Принятие решений	5	4	4	5	4	4	5	5	5	4
8. Требовательность	5	5	4	5	5	5	5	5	5	5
9. Владение информацией	4	5	4	4	5	4	4	5	4	4
10. Целеустремленность	5	5	5	5	4	4	5	5	4	5
11. Деловитость	5	5	5	5	5	5	5	5	5	5
12. Работоспособность	4	4	4	4	3	4	4	3	3	3
13. Дисциплинированность	5	5	5	5	5	5	5	5	5	5
14. Образованность	5	5	5	4	4	3	4	4	5	4

Окончание таблицы 1.

Профессионально-значимые качества личности	Ответы тренеров в баллах									
	Эксперт 1, 32 года	Эксперт 2, 24 года	Эксперт 3, 27 лет	Эксперт 4, 50 лет	Эксперт 5, 55 лет	Эксперт 6, 24 года	Эксперт 7, 23 года	Эксперт 8, 24 года	Эксперт 9, 23 года	Эксперт 10, 27 лет
15. Умение организовать коллектив	4	4	4	3	3	4	3	4	3	3
16. Опыт	4	4	4	4	5	5	4	4	4	4

Таким образом, по сути, мы имеем дело с экспертными оценками, причем в роли экспертов выступают сотрудники спортивной организации, а в роли объектов экспертизы – оцениваемые ими качества их руководителя. Рассматриваемая нами ситуация соответствует случаю однотуровой экспертизы [1], когда каждый эксперт выражает свое мнение, не общаясь и не консультируясь с другими экспертами. Он полностью независим.

Перейдем к осуществлению первого этапа – проверке согласованности мнений экспертов. Наиболее часто для этой цели применяется ранговый коэффициент конкордации Кендалла – Смита (обычно называемый просто коэффициентом конкордации Кендалла).

Для расчета коэффициента конкордации проведем следующие операции [5, с. 98–99]:

1. На основании таблицы 1 составим таблицу 2, в которой вместо баллов будут указаны ранги. Ранжирование выполняется по следующим правилам: меньшему количеству баллов присваивается меньший ранг. Наименьшему присваивается ранг 1, наибольшему  $n$  – количество оцениваемых качеств (в нашем случае  $n = 16$ ). В случае, если несколько значений равны, им начисляется ранг, представляющий собой среднее арифметическое из тех рангов, которые они получили бы, если бы не были равны.

2. Рассчитаем сумму рангов для каждого из  $n$  профессионально-значимых качеств личности.

3. Вычислим среднюю сумму рангов. Это можно сделать по формуле:

$$m \times \frac{n+1}{2}, \quad (1)$$

где  $m$  – количество экспертов (в нашем случае  $m=10$ ).

Можно просто рассчитать среднее арифметическое значение полученных сумм рангов – результат получится такой же.

4. Рассчитаем отклонение суммы рангов, полученных каждым профессионально-значимым качеством, от средней суммы рангов.

5. Определим квадраты рассчитанных в пункте 4 отклонений.

6. Просуммируем квадраты отклонений. Полученное число обозначим буквой  $S$ .

Рассчитаем коэффициент конкордации по формуле:

$$W = \frac{12S}{m^2(n^3 - n)}. \quad (2)$$

Результаты проведенных нами операций показаны в таблице 2.

Таблица 2. – Результаты ранжирования оценок тренеров и расчет коэффициента конкордации

Номер оцениваемого качества	Ранги, рассчитанные по баллам экспертов										Сумма рангов	Отклонение от средней суммы рангов	Квадрат отклонения от средней суммы рангов
	Эксперт 1, 32 года	Эксперт 2, 24 года	Эксперт 3, 27 лет	Эксперт 4, 50 лет	Эксперт 5, 55 лет	Эксперт 6, 24 года	Эксперт 7, 23 года	Эксперт 8, 24 года	Эксперт 9, 23 года	Эксперт 10, 27 лет			
1.	6	4	12,5	7	12,5	12,5	6,5	6	7,5	7	81,5	-3,5	12,25
2.	1,5	4	5	2	6	1,5	2	6	7,5	7	42,5	-42,5	1806,25
3.	6	12	12,5	7	12,5	12,5	13	6	13,5	14	109	24	576
4.	13	12	12,5	13,5	12,5	12,5	6,5	13	7,5	7	110	25	625
5.	6	12	12,5	7	6	12,5	13	6	2,5	7	84,5	-0,5	0,25
6.	1,5	4	1	2	2	5,5	2	1,5	2,5	7	29	-56	3136
7.	13	4	5	13,5	6	5,5	13	13	13,5	7	93,5	8,5	72,25
8.	13	12	5	13,5	12,5	12,5	13	13	13,5	14	122	37	1369
9.	6	12	5	7	12,5	5,5	6,5	13	7,5	7	82	-3	9
10.	13	12	12,5	13,5	6	5,5	13	13	7,5	14	110	25	625
11.	13	12	12,5	13,5	12,5	12,5	13	13	13,5	14	129,5	44,5	1980,25
12.	6	4	5	7	2	5,5	6,5	1,5	2,5	1,5	41,5	-43,5	1892,25
13.	13	12	12,5	13,5	12,5	12,5	13	13	13,5	14	129,5	44,5	1980,25
14.	13	12	12,5	7	6	1,5	6,5	6	13,5	7	85	0	0
15.	6	4	5	2	2	5,5	2	6	2,5	1,5	36,5	-48,5	2352,25
16.	6	4	5	7	12,5	12,5	6,5	6	7,5	7	74	-11	121
Средняя сумма рангов											85	S=16557	

Рассчитанный по формуле (2) коэффициент конкордации оказался равным 0,49. Обычно считается, что согласованность мнений экспертов является удовлетворительной, если  $W > 0,5$ , и хорошей, если  $W > 0,7$ . В нашем случае согласованность мнений экспертов ниже удовлетворительного уровня.

Для того чтобы определить, присутствует ли согласованность в оценках экспертов, выдвинем статистические гипотезы. Нулевая гипотеза предполагает полную рассогласованность мнений экспертов и статистическую недостоверность рассчитанного коэффициента конкордации, конкурирующая гипотеза предполагает, что во мнениях экспертов присутствует согласованность, и рассчитанный нами коэффициент конкордации статистически достоверен. Для про-

верки гипотез используем  $\chi^2$ -критерий [6]. С этой целью вычислим  $\chi^2_{\text{эмп}}$  по формуле:

$$\chi^2_{\text{эмп}} = \frac{12S}{mn(n+1)} = \frac{12 \times 16557}{10 \times 16 \times 17} = 73,05. \quad (3)$$

Эмпирическое значение  $\chi^2_{\text{эмп}}$  сравним с табличным, соответствующим принятому уровню значимости  $\alpha$  и числу степеней свободы  $k=n-1$ . Если  $\chi^2_{\text{эмп}} > \chi^2_{\text{кр}}(\alpha; k)$ , то коэффициент конкордации  $W$  существенен на выбранном уровне значимости.

Итак, на уровне значимости  $\alpha=0,05$   $\chi^2_{\text{кр}}(0,05;15)=24,996$ . Поскольку  $73,05 > 24,996$ , мы должны отвергнуть нулевую гипотезу, так как при истинности выдвинутой нулевой гипотезы полученный нами результат маловероятен (его вероятность меньше 0,05). На уровне значимости  $\alpha=0,01$   $\chi^2_{\text{кр}}(0,01;15)=30,578$ , что дает нам право с вероятностью 99 % принять конкурирующую гипотезу о статистической достоверности коэффициента конкордации. Значит, в нашем случае имеет место слабая согласованность мнений экспертов.

Проведем усреднение мнений экспертов – второй этап математической обработки экспертных оценок. С целью такого усреднения после получения оценок экспертов «...рассчитывают средние баллы и рассматривают их как интегральные (т. е. обобщенные, итоговые) оценки, выставленные объектам экспертизы коллективом опрошенных экспертов. Какими формулами пользоваться для вычисления средних величин? Ведь средних величин существует, как мы знаем, весьма много разных видов.

По традиции обычно применяют среднее арифметическое. Специалисты по теории измерений уже более 30 лет знают, что такой способ некорректен, поскольку баллы обычно измерены в порядковой шкале. Обоснованным является использование медиан в качестве средних баллов. Однако полностью игнорировать средние арифметические нецелесообразно из-за их привычности и распространенности. Поэтому представляется рациональным использовать одновременно оба метода – и метод средних арифметических баллов, и метод медиан баллов. Такая рекомендация находится в согласии с общенаучной концепцией устойчивости [7], рекомендующей применять различные методы для обработки одних и тех же данных с целью выделить выводы, получаемые одновременно при всех методах. Такие выводы, видимо, соответствуют реальной действительности, в то время как заключения, меняющиеся от метода к методу, зависят от субъективизма исследователя, выбирающего метод обработки исходных экспертных оценок» [1, с. 103–104].

Результаты расчетов средних арифметических и медиан каждого профессионально-значимого качества личности приведены в таблице 3.

Как видим из таблицы 3, в ряде случаев среднее арифметическое и медиана совпадают (качества № 7, 11, 13, 15). Максимальное отличие среднего арифметического и медианы наблюдается в качествах № 2, 12 и составляет 0,4 балла. Это позволяет сделать вывод о том, что результаты, полученные каждым методом, мало различаются между собой. Следовательно, оба метода можно при-

менять для определения интегральных оценок, выставленных экспертами объектам экспертизы.

Таблица 3. – Средние арифметические и медианы оценок тренеров

Профессионально-значимые качества личности	Среднее арифметическое	Медиана
1. Компетентность	4,3	4
2. Знание действительности	3,6	4
3. Понимание целей организации	4,7	5
4. Умение видеть проблемы	4,7	5
5. Умение видеть ситуацию	4,3	4
6. Восприимчивость к изменениям	3,3	3
7. Принятие решений	4,5	4,5
8. Требовательность	4,9	5
9. Владение информацией	4,3	4
10. Целеустремленность	4,7	5
11. Деловитость	5	5
12. Работоспособность	3,6	4
13. Дисциплинированность	5	5
14. Образованность	4,3	4
15. Умение организовать коллектив	3,5	3,5
16. Опыт	4,2	4

**Заключение.** Приведенный нами пример математической обработки экспертных оценок позволяет сделать следующие выводы:

1. Поскольку объекты экспертных оценок в сфере физической культуры и спорта чаще всего относятся к области спорта высоких достижений, а эксперты являются специалистами высокого класса, их оценки не могут быть полностью рассогласованными. В любом случае при проведении экспертизы будет иметь место согласованность мнений. В нашем случае слабая согласованность мнений вызвана тем, что по условию эксперимента профессионально-значимые качества личности руководителя спортивной организации оценивались подчиненными ему тренерами, которых невозможно было заменить другими экспертами.

2. Как было показано выше, мнения экспертов в области спорта высоких достижений не могут сильно различаться. Поэтому есть смысл интегрировать (усреднять) оценки коллектива экспертов. Для этой цели можно использовать два метода: метод среднего арифметического – некорректный с точки зрения теории измерений и метод медианы – абсолютно корректный метод. Одинаковость полученных результатов позволяет утверждать, что целесообразно использовать оба метода.

Отметим также, что в рамках данной статьи мы опирались на распространенное мнение о том, что выбор методов и приемов экспертных заключений зависит, помимо всего прочего, от степени согласованности ответов экспертов [8]. Как утверждалось выше, в исследованиях в сфере физической культуры и спорта

согласованность экспертных оценок в той или иной степени всегда присутствует. Однако, в общем случае, требование согласованности мнений экспертов – не более чем миф. На практике всегда существует лицо, принимающее решение (ЛПР) – человек, который на основании результатов работы экспертной комиссии и, возможно, с учетом другой: либо недоступной экспертам, либо даже общедоступной, но не принятой экспертами во внимание информации принимает окончательное решение. Расхождение мнений экспертов, наличие среди них специалистов, суждение которых значительно отличается от преобладающего мнения остальных экспертов, может быть обусловлено одной из двух или обеими причинами: либо часть из этого меньшинства недостаточно компетентна, либо другая его часть по компетентности и по пониманию проблемы превосходит всех членов экспертной комиссии. Вот тогда-то решающее слово за ЛПР!

1. Орлов, А. И. Организационно-экономическое моделирование : в 3 ч. / А. И. Орлов. – М. : МГТУ им. Н. Э. Баумана, 2011. – Ч. 2 : Экспертные оценки – 567 с.
2. Сидоренко, Е. В. Методы математической обработки в психологии / Е. В. Сидоренко. – СПб. : Социально-психологический центр, 1996. – 349 с.
3. Холлендер, М. Непараметрические методы статистики / М. Холлендер, Д. Вулф. – М. : Финансы и статистика, 1983. – 518 с.
4. Орлов, А. И. Прикладная статистика : учебник / А. И. Орлов. – М. : Экзамен, 2006. – 671 с.
5. Спортивная метрология : учебник для ин-тов физ. культуры / В. М. Зацюрский [и др.] ; под общ. ред. В. М. Зацюрского. – М. : Физкультура и спорт, 1982. – 256 с.
6. Алгоритмика, статистика и теория вероятностей. Существенность коэффициента конкордации [Электронный ресурс]. – Режим доступа : <http://matstats.ru/kk.html>. – Дата доступа : 05.04.2016.
7. Орлов, А. И. Устойчивость в социально-экономических моделях / А. И. Орлов. – М. : Наука, 1979. – 296 с.
8. Шмерлинг, Д. С. Экспертные оценки. Методы и применение / Д. С. Шмерлинг [и др.] // Статистические методы анализа экспертных оценок. – М. : Наука, 1977. – 284 с.

*Поступила 07.07.2016*

## **ЗАДАЧИ И СОДЕРЖАНИЕ НАУЧНЫХ ИССЛЕДОВАНИЙ ПРИ СПОРТИВНО-ОРИЕНТИРОВАННОМ ФИЗИЧЕСКОМ ВОСПИТАНИИ СТУДЕНТОВ**

**В.И. Ярмолинский**, канд. техн. наук,  
Белорусский государственный университет

*Смещение акцентов в системе общего физкультурного образования студентов в сторону спортивно-ориентированного физического воспитания требует нового взгляда на задачи и механизмы деятельности кафедр физического воспитания и спорта, призванных, прежде всего, укреплять здоровье и повышать работоспособность будущих специалистов. Автор считает, что при-*