Введение в статистические и математические методы педагогических измерений

Вадим Аванесов
testolog@mail.ru

Опубликовано в ж. «Педагогические Измерения № 4 2005 г.

В №2 журнала «Педагогические Измерения» была первая напечатана статья по статистическим и математическим методам педагогических измерений. В № 4 журнала излагается продолжение этой темы. Рассмотрены методы вычисления показателей вариации, асимметрии, эксцесса, проверки статистических гипотез, расчета классического коэффициента корреляции Пирсона, понятия и методы матричной алгебры.

В обеих статьях представлена часть авторского курса, подготовленного для начинающих тестологов, студентов гуманитарных вузов, а также на преподавателей гуманитарных учебных дисциплин, интересующихся педагогическими измерениями.

5. ВЫЧИСЛЕНИЕ ОСНОВНЫХ СТАТИСТИК (ПРОДОЛЖЕНИЕ)

5.2. *Расчёт показателей вариации.* Помимо средних арифметических, для характеристики групп испытуемых важен расчет показателей *вариации.* Это показатели позволяют оценить меру рассеяния данных относительно средней арифметической. В практике обычно используется четыре основных показателя.

Первый показатель вариации - это *сумма квадратов отклонений от средней арифметической.* Сущность этого показателя легко понять из примеров. Вначале возьмем трех испытуемых, имеющих мало различающиеся баллы: 1, 2, и 3. Следовательно, мы вправе ожидать и сравнительно низкий показатель вариации. Баллы представлены во втором столбце табл.2. В последней строке этого столбца представлена сумма баллов, $\Sigma X_i = (1 + 2 + 3) = 6$. Среднее арифметическое считается по известной формуле $M = \dfrac{\sum X_i}{N} = \dfrac{6}{3} = 2$. В третьем столбце табл. 2 средняя арифметическая M используется для вычитания из каждого балла испытуемого. В итого сумма отклонений баллов от среднего арифметического балла $\Sigma(X_i – M)$, как и следовало ожидать, равна нулю. Это и есть главное свойство средней арифметической: сумма отклонений баллов испытуемых от средней арифметической - по любому тесту и в любой группе - равна нулю. Если при конкретных расчетах это свойство не подтверждается, то это происходит из-за неправильного расчета самой средней, ошибки округления или счётной ошибки.

Табл.3

Испытуемые	Баллы	$(X_i – M)$	$(X_i – M)^2$
1. Иванова	1	1 – 2 = -1	1
2. Петров	2	2 – 2 = 0	0
3. Сидоров	3	3 – 2 = 1	1
Σ:	6	0	2

Как видно из последней строки табл. 3, сумма баллов по всем шести испытуемым равна шести. Сумма всех отклонений от средней арифметической равна, как и следовало ожидать, равна нулю. Сумма квадратов отклонений от средней арифметической отличается от нуля. Она всегда положительна.

Теперь возьмем другой пример, где баллы испытуемых отличаются по своим значениям, или иначе, они варьируют больше, чем в предыдущем примере. Следовательно, мы вправе ожидать более высокое значение суммы квадратов отклонений от среднего арифметического значения.

Табл.4.

ФИО испытуемых	Баллы	$(X_i – M)$	$(X_i – M)^2$
1. Митин	1	1 – 6 = -5	25
2. Семина	5	5 – 6 = -1	1
3. Богданова	12	12 – 6 = 6	36
Σ:	18	0	62

Как видно из последней строки табл. 4, сумма баллов по всем шести испытуемым заметно выросла, но еще больше выросла сумма квадратов отклонений от среднего арифметического. $\Sigma(X_i - M)^2$ можно представить в виде символа SS_x и считать по другой, более удобной формуле.

Второй, по счёту, но не по важности, показатель вариации – это *дисперсия* исходных тестовых баллов, или иначе, *варианса*.

В учебниках по статистике доказывается, что

$$SS_x = \Sigma(X_i - M)^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{N} \qquad [4]$$

Расчет значения дисперсий для ранее введенных данных в экспериментальной и контрольной группах находится так: каждое значение надо возвести в квадрат и квадраты сложить (ΣX^2), затем возвести в квадрат сумму тестовых баллов, разделить на число человек в группе и вычесть частное от суммы квадратов:

$$SS_1 = \Sigma X_1^2 - \frac{(\Sigma X_1)^2}{N} \qquad SS_2 = \Sigma X_2^2 - \frac{(\Sigma X_2)^2}{N}$$

$$SS_1 = 2106 - \frac{(150)^2}{11} = 6,054 \qquad SS_2 = 841 - \frac{(85)^2}{9} = 4,778$$

Затем находится *стандартная* мера вариации тестовых баллов, которая называется *дисперсией.* Дисперсия обозначается символом s^2 и рассчитывается по формуле

$$s^2 = \frac{SS}{N-1} \qquad (5)$$

Подставляя данные, получаем значения дисперсии в каждой группе:

$$s_1^2 = \frac{SS_1}{N-1} = \frac{60,545}{11-1} = 6,054 \qquad s_2^2 = \frac{SS_2}{N-1} = \frac{38,222}{9-1} = 4,778$$

Для тестовых заданий, в которых используется только дихотомическая оценка (1 или 0) дисперсия определяется по сравнительно простой формуле:

$$s_j^2 = p_j q_j \qquad (6)$$

где p_j и q_j, - доли правильных и неправильных ответов в каждом задании (j).

Третий показатель вариации называется *стандартное отклонение*. Эта мера вариации исходных тестовых баллов удобна тем, что в отличие от дисперсии, она выражается в тех же единицах, в которых выражен измеряемое латентное свойство личности. Стандартное отклонение обозначается символом s; оно определяется по формуле

$$s = \sqrt{\frac{SS}{N-1}} \qquad (7)$$

Соответственно, в группах, стандартные отклонения равны:

$$S_1 = \sqrt{\frac{SS_1}{N-1}} = \sqrt{\frac{60{,}545}{11-1}} = 2{,}460; \quad S_2 = \sqrt{\frac{SS_2}{N-1}} = \sqrt{\frac{38{,}222}{9-1}} = 2{,}186$$

При небольших массивах данных эти расчеты легко выполняются с помощью калькуляторов. При большом числе испытуемых, оцениваемых к тому же не по одному, а по множеству признаков, данные лучше обрабатывать на ЭВМ, по стандартной программе расчета SS, s^2, s.

6. Определение коэффициента асимметрии и эксцесса исходных тестовых баллов.

Коэффициент асимметрии (As) указывает на наличие или отсутствие смещения результатов от так называемого нормального распределения, вправо или влево. При смещении возникает левосторонняя или правосторонняя асимметрия. Расчет коэффициента асимметрии проведём на новом примере данных с оценками 25 испытуемых.

Эти данные представлены в табл. 5.

N П/п	X	$X-\overline{X}$	$(X-\overline{X})^2$	$(X-\overline{X})^3$	$(X-\overline{X})^4$
1	1	-3	9	-27	81
2	1	-3	9	-27	81
3	2	-2	4	-8	16
4	2	-2	4	-8	16
5	2	-2	4	-8	16
6	2	-2	4	-8	16
7	2	-2	4	-8	16
8	2	-2	4	-8	16
9	3	-1	1	-1	1
10	3	-1	1	-1	1
11	3	-1	1	-1	1
12	3	-1	1	-1	1
13	3	-1	1	-1	1
14	4	0	0	0	0
15	4	0	0	0	0
16	4	0	0	0	0
17	4	0	0	0	0
18	5	1	1	1	1
19	5	1	1	1	1
20	6	2	4	8	16

21	6	2	4	8	16
22	7	3	9	27	81
23	8	4	16	64	256
24	9	5	25	125	625
25	9	5	25	125	625
$\sum$:			132	252	1884

Процесс нахождения коэффициентов асимметрии и эксцесса выполняется в несколько этапов.

1. Вначале находят среднюю арифметическую, по формуле:

$\overline{X} = 4; \quad S_x = 2.345; \quad n = 25$

$$A_s = \frac{\sum\left(X - \overline{X}\right)^3}{n * S_x{}^3} = \frac{252}{25 * 2.345^3} = 0.781\text{, где}$$

A_s – коэффициент асимметрии. Помимо коэффициента рассчитывается его ошибка:

$$M_{A_s} = \sqrt{\frac{6}{n}} = 0.49$$

Достоверность коэффициента асимметрии определяется по формуле:

$$t_{A_s} = \frac{A_s}{M_{A_s}} = \frac{0.781}{0.49} = 1.595$$

Если $t_{A_s} < 3.0$, то гипотеза о достоверности асимметрии (H_1) отвергается и принимается нулевая гипотеза (H_0) о нормальности распределения.

Из соображений полноты характеристики данных помимо коэффициента а симметрии A_s рассчитывается коэффициент эксцесса (E_x), указывающий на вытянутость вверх (в случае положительного значения) или вогнутости вниз (в случае отрицательного значения) кривой нормального распределения.

По данным последнего примера:

$$E_x = \frac{\sum\left(X - \overline{X}\right)^4}{n * S_x{}^4} - 3 = \frac{1884}{25 * 30.25} - 3 = -0.509$$

Ошибка $E_x - M_{E_x} = 2 * \sqrt{\frac{6}{n}} = 2 * \sqrt{\frac{6}{25}} = 0.980$

$$t_{E_x} = \frac{E_x}{M_{E_x}} = \frac{-0.509}{0.980} = 0.519$$

Если $t_{E_x} < 3.0$, то гипотеза H_1 отвергается и принимается гипотеза H_0, свидетельствующая об отсутствии достоверного эксцесса.

7. Проверка гипотез

Исследовательская гипотеза – это научное предположение, вытекающее из соображений теоретического или практического характера. Применительно к проведению экспериментов, это обычно гипотезы о преимуществе того или иного нового метода, по сравнению с ранее применявшимся методом. В процессе статистического анализа исследовательская гипотеза переводится на язык статистической науки и заново формулируется, по меньшей мере, в виде двух статистических гипотез.

Первая, основная, называется *нулевой гипотезой* (H_o), в которой исследователь говорит о своей исходной позиции. Априори как бы декларируется, что предполагаемый метод не обладает какими-либо преимуществами, и потому с самого начала исследователь психологически готов занять честную научную позицию: различия между новым и старым методами заранее, до начала эксперимента, объявляются равными нулю.

Во второй, *альтернативной гипотезе* (H_1) делается предположение о преимуществе нового метода. Иногда выдвигается несколько альтернативных гипотез. Например, гипотеза о преимуществе прежнего метода (H_2). Альтернативные гипотезы принимаются только тогда, когда допускается возможность опровержения нулевой гипотезы. При проведении экспериментов различия, скажем, в средних арифметических экспериментальной и контрольной групп связаны с вероятностью риска отвергнуть нулевую и принять альтернативную гипотезу. Чем больше различия средних, тем больше вероятность достоверных различий, а не просто случайных флуктуаций в значениях выборочных средних.

Чем больше испытуемые в группе отличаются по уровню развития изучаемого признака, тем больше у них проявляются различия в тестовых баллах. Группы с большой вариацией называются гетерогенными (разнородными) по составу: соответственно, по мере уменьшения вариации группа все более становится гомогенной. Например, педагогу удобное работать в гомогенной группе, так как здесь легче подобрать формы и методы, эффективные если не для всех, то для большинства учащихся. В гетерогенных же группах метод, хороший для одних, нередко оказывается плохим для других.

Среднее арифметическое и стандартное отклонение является лучшей характеристикой группы по измеряемому признаку. Первое являете обобщенным показателем достигнутого группой уровня в среднем, в виде одного числа, как меры центральной тенденции. Второе, наряду с дисперсией, является общепринятым показателем вариации.

7.1. Уровни значимости. Риск отвергнуть нулевую гипотезу, когда она верна, связывают с одним из трех принятых *уровней значимости* статистического вывода:

Первый уровень – 5% риска допустить ошибку в выводах. В научных текстах изображают иногда в виде символов р = 5°/о или $\alpha \leq 0.05$ (если представлено в долях). Что означает: допускается риск ошибки в выводе в пяти случаях из ста теоретически возможных таких же экспериментов, при строго случайном отборе испытуемых для каждого эксперимента;

Второй уровень – 1%, т. е. соответственно допускается риск ошибиться в одном случае из ста ($\alpha \leq 0.01$, при тех же требованиях);

Третий уровень - 0,1%, т. е. допускается риск ошибиться в одном случае из тысячи ($\alpha \leq 0.001$). Последний уровень значимости предъявляет очень высокие требования к обоснованию достоверности результатов эксперимента и потому редко используется.

7.2. Различие случайное и закономерное. При сравнении средних арифметических экспериментальной и контрольной групп важно не только определить, какая средняя больше, но и насколько больше. Чем меньше разница между ними, тем более приемлемой окажется нулевая гипотеза об отсутствии статистически значимых (достоверных) различий. В отличие от мышления на уровне обыденного сознания, склонного воспринимать полученную в результате опыта разность средних арифметических (в нашем примере M_1=13.636,

M_2=9.444, M_1> M_2) как основание для вывода. Исследователь, знакомый с логикой статистического вывода, не будет торопиться в таких случаях. Он, скорее всего, сделает предположение о случайности различий, выдвинет нулевую гипотезу об отсутствии достоверных различий в результатах экспериментальной и контрольной групп, и лишь после опровержения нулевой гипотезы примет альтернативную.

Таким образом, вопрос о различиях в рамках научного мышления переводится в другую плоскость. Дело не только в различиях (они почти всегда есть), а в величине этих различий, и отсюда - в определении той разницы границы, после которого можно сказать: да, различия неслучайны, они статистически достоверны. А значит, испытуемые этих двух групп принадлежат после эксперимента уже не к одной (как раньше), а к двум различным генеральным совокупностям, и что уровень подготовленности учащихся, потенциально принадлежащих этим совокупностям, существенно отличается.

7.3. Стандартная ошибка средней арифметической. Стандартное отклонение используется для расчета так называется *статистической ошибки средней арифметической* $s_м$, которая является показателем точности определения выборочной средней по отношению к средней арифметической генеральной совокупности μ.

Ошибка средней арифметической $s_м$ не имеет ничего общего с ошибочными расчетами или решениями. Она является, напротив, мерой точности и определяется из выражения

$$s_м = \sqrt{\frac{SS}{N(N-1)}}. \qquad (8)$$

Соответственно,

$$s_{м_1} = \sqrt{\frac{60{,}545}{11(11-1)}} = 0{,}742 \quad \text{и} \quad s_{м_2} = \sqrt{\frac{38{,}222}{9(9-1)}} = 0{,}729$$

7.4. t – критерий Стьюдента. Доказательство статистической достоверности экспериментального влияния существенно отличается от доказательства в математике и формальной логике. В математике выводы носят более универсальный характер. Статистические же доказательства не являются столь строгими и окончательными - в них всегда допускается риск ошибиться в выводах, и потому статистическими методами не доказывается окончательно правомерность того или иного вывода, а показывается мера правдоподобности принятия той или иной гипотезы.

7.5. Оценка генеральных параметров. Для того чтобы показать границы этих различий, используются так называемые *оценки генеральных параметров.* Например, на основе расчета выборочной средней арифметической можно сделать оценку средней арифметической в генеральной совокупности. Последняя называется параметром и обозначается символом μ. Параметр μ определяется из предположения о том, что средняя арифметическая выборочного распределения (M) является лучшей, чем какая-либо другая оценка средней арифметической генеральной совокупности (μ), плюс некоторая погрешность оценки, обозначаемая символом [Δ]. Символически это выражается так:

$$\mu = M \pm \Delta.$$

При оценке значения параметра на основе значения статистики вопрос сводится к расчету средней арифметической выборочной совокупности и к определению ширины так называемого *доверительного интервала* Δ. В зависимости от принимаемого уровня значимости величина доверительного интервала может быть большей или меньшей. При пятипроцентном уровне риска допустить ошибку в выводе интервал Δ определяется из выражения $\Delta = \pm(1{,}96)\cdot(s_м)$, где 1,96—константа, табличное число, используемое при достаточно большом числе испытуемых, а $s_м$ - так называемая ошибка средней арифметической. При однопроцентном уровне риска ошибиться в выводе этот интервал естественным образом расширится: $\Delta = \pm 2{,}58\cdot б_м$, где 2,58 - также табличное число. Из этих рассуждений следует, что средняя арифметическая генеральной совокупности примет, скорее всего, одно из значений в указанных интервалах:

$$\mu = M \pm (1{,}96)\,(s_м) \quad (5\%) \quad \text{и} \quad \mu = M \pm 2{,}58\, s_м \quad (1\%)$$

В статистических таблицах I распределения Стьюдента даются значения для малых выборок. При прочих равных условиях, чем ўже доверительные интервалы, тем меньше вероятность пересечения распределений двух выборок. А это в немалой степени зависит и от количества испытуемых. Поэтому надо стремиться иметь в экспериментальной и контрольной группах примерно по тридцать и более человек—там, где это возможно. Чем меньше испытуемых, тем труднее показать достоверность различий там, где они действительно есть.

Поскольку доверительные интервалы сравниваемых средних арифметических нередко совпадают, встает вопрос о допустимой границе их совпадения. Решение о такой границе получается в процессе использования так называемого t - критерия Стьюдента:

$$t = \frac{M_1 - M_2}{s_{M_1 - M_2}}. \qquad (9)$$

где M_1 и M_2 - средние арифметические в экспериментальной и контрольной группах, а $s_{M_1 - M_2}$ называется *стандартная ошибка разности средних арифметических.* Значение этой, хотелось бы подчеркнуть, статистической ошибки $s_{M_1 - M_2}$ находится из формулы

$$s_{M_1 - M_2} = \sqrt{\frac{SS_1 + SS_2}{N_1 + N_2 - 2}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)} \qquad (10)$$

$$s_{M_1 - M_2} = \sqrt{\frac{60{,}545 + 38{,}222}{11 + 9 - 2}\left(\frac{1}{11} + \frac{1}{9}\right)} = 1{,}053.$$

Далее находится t - значение, которое надо будет сравнить с критерием - теоретическими значениями t-распределения Стьюдента (прилагается в таблицах к учебникам статистики). Для нашего примера

$$t = \frac{13{,}636 - 9{,}444}{1{,}053} = 3{,}981$$

После чего сравниваем полученное в эксперименте значение t с табличным значением t - распределения Стьюдента. При этом учитывается так называемые степени свободы (df), равные числу испытуемых минус два. В символической записи ($N_1 + N_2 - 2$). В примере $N_1 = 11$ чел., $N_2 = 9$ чел., df равно 11 + 9 - 2; всего 18. Табличное значение t при df=18 равняется 2,1, если принять возможность риска сделать ошибочное суждение в пяти случаях из ста (p = 5°/o или $\alpha \le 0.05$). Если полученное в эксперименте абсолютное значение t превышает табличное, то есть основания принять альтернативную гипотезу (H_1) о том, что учащиеся экспериментальной группы показывают в среднем более высокий уровень знаний. В эксперименте t = 3.981. Табличное t = 2.10; 3,981>2,10, откуда следует вывод о преимуществе экспериментального обучения.

Здесь могут возникнуть такие вопросы:

1. Что, если полученное в опыте значение t окажется меньше табличного? Ответ на этот вопрос довольно прост: тогда альтернативная гипотеза отвергается, а нулевая гипотеза принимается.
2. Если полученное в эксперименте t значение Стьюдента оказалось больше табличного (как в нашем случае t =3.981, при табличном значении t = 2.10), можно ли утверждать о доказанном преимуществе экспериментального метода?

Ответ: вряд ли. Не столько доказано, сколько показано, потому что с самого начала допускается риск ошибиться в пяти случаях из ста ($\alpha \le 0.05$). Наш эксперимент мог быть одним из этих пяти случаев. Но 95% возможных случаев говорит в пользу альтернативной гипотезы, а это достаточно убедительный аргумент в статистическом доказательстве.

3. Что делать, если в контрольной группе результаты окажутся выше, чем в экспериментальной? Сделаем, например, M_2 средней арифметической экспериментальной группы, а M_1 – контрольной.

$$t = \frac{9{,}444 - 13{,}636}{1{,}053} = -3{,}981$$

В таких случаях принимается во внимание абсолютное значение t (по модулю). Поскольку абсолютное значение t = 3,981 больше табличного значения t = 2,110, то следует вывод, что новый метод не выявил никакого преимущества. Наоборот, подтвердил своё преимущество традиционный метод. Вследствие чего отвергаются нулевая (H_o) и первая альтернативная гипотезы (H_1) и принимается вторая альтернативная гипотеза ((H_2) о преимуществе традиционного метода.

Применение t-критерия Стьюдента в некоторых случаях имеет свои особенности. Рассмотрим, например, широко распространенный в практике случай, когда эксперимент проводится в одной группе. Например, изучал уровень ориентации учащихся на художественно-эстетические ценности. С целью активизации формирования этой ориентации в экспериментальной группе проводились беседы, выставки детских рисунков, были организованы и посещения музеев и картинных галерей, проведены встречи с музыкантами, художниками и др. Закономерно встаёт вопрос: какова эффективность проведенной работы? С целью поиска ответа на этот вопрос до начала и после эксперимента давался тест.

7.6. Расчет коэффициентов корреляции. Результаты тестирования до начала и после окончания эксперимента приводятся в таблице 2. Таблица для расчета коэффициента корреляции. Табл.6

Испытуемые	Баллы до эксп. (X)	Баллы после эксп. (Y)	XY	X^2	Y^2
Иванов	14	18	252	196	324
Новикова	20	19	380	400	361
Сидорова	15	22	330	225	484
Пирогов	11	17	187	121	289
Агапов	16	24	384	256	576
Суворова	13	21	273	169	441
Рыжиков	16	25	400	256	625
Серова	19	26	494	361	676
Быстрова	15	24	360	225	576
Топоров	9	15	135	81	225
Σ:	148	211	3195	2290	4577
M	14,8	21,1			
SS	99,6	124,90			

Из методических соображений в этой таблице приводятся результаты небольшого числа испытуемых.

Первый шаг. По формуле (3) -находим суммы квадратов отклонений по X и по У.

$$SS_x = \Sigma X^2 - \frac{(\Sigma X)^2}{N} = 2290 - \frac{(148)^2}{10} = 99{,}6$$

$$SS_y = \Sigma Y^2 - \frac{(\Sigma Y)^2}{N`} = 4577 - \frac{(211)^2}{10} = 124{,}90$$

Второй шаг. Находим сумму произведений Х и У, скорректированную на средние значения (SP_{xy}), по формуле

$$SP_{xy} = \sum XY - \frac{\sum X \sum Y}{N}$$

$$SP_{xy} = 3195 - \frac{(148)(211)}{10} = 72{,}200$$

Третий шаг. Находим коэффициент корреляции по формуле

$$r = \frac{SP_{xy}}{\sqrt{SS_x\ SS_y}}$$

$$r = \frac{72{,}2}{\sqrt{(99{,}6)\ (124{,}90)}}$$

Если связь между начальными и конечными результатами подтверждается, а в примере табл. 2 это так, то стандартная ошибка разности средних арифметических рассчитывается не по ранее приведенной формуле, а по другой формуле

$$s_{M_1-M_2} = \sqrt{s^2_{Mx} + s^2_{My}}\ \text{--}2r(s_{Mx})\ (s_{My})\quad)$$

Как видно из формулы (12), необходимо рассчитать стандартные ошибки средних арифметических, их квадраты и коэффициент корреляции; r уже рассчитан. По формуле (7) находим стандартные ошибки средних арифметических:

$$s_M = \sqrt{\frac{SS}{N(N-1)}}$$

Стандартные ошибки средних арифметических для каждой группы равны

$$s_{M_x} = \sqrt{\frac{99{,}600}{10\,(10-1)}} = 1{,}052 \quad \text{и} \quad s_{M_y} = \sqrt{\frac{124{,}900}{10\,(10-1)}} = 1{,}178$$

Соответственно, квадраты этих ошибок, требуемые для подстановки в формулу (12) равны

$$s^2_{Mx} = (1{,}052)^2 = 1{,}107 \qquad s^2_{My} = (1{,}176)^2 = 1{,}388$$

Эти данные подставляем в формулу (11) . Тем самым находится стандартная ошибка разности средних арифметических двух зависимых совокупностей

$$s_{M_1-M_2} = \sqrt{1{,}107 + 1{,}383 - 2(0{,}647)(1{,}052)(1{,}178)} = 0{,}944$$

Теперь есть все необходимое для получения значения t. Данные подставляем в формулу:

$$t = \frac{M_x - M_y}{s_{M_1 - M_2}} = \frac{14{,}800 - 21{,}100}{0{,}944} = -6{,}677.$$

Поскольку учитывается только абсолютное значение $t = |6{,}674|$, то видно; оно довольно заметно превышает табличное, теоретическое значение t - распределения Стьюдента, t =2,10, с учетом степеней свободы $df = N_1+N_2 - 2$). Откуда следует возможность принятия альтернативной гипотезы (H_1) о достоверных различиях средних арифметических. После чего делается вывод об эффективности экспериментального воздействия.

Для того чтобы меру влияния Х на У выразить в процентах, r^2 умножают на 100. В нашем примере г=0,647, r^2 =0,419, т. примерно 42%. В случае двух и большего числа независимых переменных рассчитываются коэффициенты множественной корреляции и регрессии.

Статистические пакеты персональных компьютеров позволяют ставить и решать задачи оценки связи, влияния и прогноза. Выше уже приводились формулы для оценки связи посредством расчета коэффициентов корреляции. При наличии множества изучаемых признаков результаты корреляционного анализа компьютер выдает в виде корреляционной матрицы (R).

Сам по себе факт применения ЭВМ для расчёта данных не является гарантией качества получаемых результатов. В исследованиях еще довольно часто ЭВМ используется для точных расчетов заведомо неточных исходных данных. Вот почему еще более актуализировалась проблема использования такой информации, которая отвечала бы известным критериям надежности (точности) и валидности (пригодности, соответствия получаемых результатов цели применения теста). В некоторых изданиях и в практике часто можно слышать устаревшее словоупотребление «надежность теста» вместо правильного «надежность тестовых результатов», и «валидность теста» вместо более правильного «валидность тестовых результатов».

8. ЭЛЕМЕНТЫ МАТРИЧНОЙ АЛГЕБРЫ[1]

8.1. *Определение матриц и числовых векторов.* Проверка тестовых результатов по этим двум критериям и разработка теста в целом проводится посредством применения матриц данных. Данные возникают как результат взаимодействия множества испытуемых (n) и множества заданий (m)

Матрица – это форма компактного и удобного представления результатов тестирования. Матрицей называется множество чисел, упорядоченных в форме таблицы данных, по n числу строк и по m числу столбцов.

Изучение матриц обычно начинается с понятий «вектор-строка» и «вектор-столбец». Объединение нескольких таких векторов в одно целое образует матрицу. Рассмотрим случай, когда один испытуемый отвечает, например, на десять заданий. Оценки, полученные испытуемым, представляются как строка тестовых баллов. В математике это называется вектор-строка: $X_{1x10} = (1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0)$. В этом векторе легко видеть, что $x_{11} = 1$, $x_{12} = 1$; …$x_{18} = 0$; $x_{19} = 0$ и т.п.

Оценки за выполнение каждого задания теста не обязательно принимаются равными единице или нулю. При тестировании эта оценка удобна, а потому широко распространена в практике тестирования. Правда, эта оценка далеко не лучшая, а потому её давно пора заменить на более дифференцирующие оценки.

Лучше, когда оценки испытуемых дифференцированы по уровню трудности заданий, а также по степени правильности и полноты ответа. Это вопрос шкалирования заданий и шкалирования ответов. Шкальные баллы иногда округляются до целых значений. Тогда результаты одного испытуемого по трем, например, заданиям теста представляются как $X_{1x3} = (4\ \ 2\ \ 1)$. Это тоже вектор-строка.

Теперь посмотрим, как записываются оценки множества испытуемых, полученных в одном задании теста. Например, оценки

---

[1] Использован материал из монографии: В.С. Аванесов «Тесты в социологическом исследовании». Москва, Наука, 1982. Digitized, Jul. 28, 2011, the University of California. 196 pages/

четырех испытуемых по одному заданию (или тесту) записываются так:

$X_{4x1} = \begin{pmatrix} 3 \\ 2 \\ 4 \\ 1 \end{pmatrix}$. Получается вектор-столбец.

Если собрать все баллы испытуемых по всем заданиям теста в одну таблицу, то получим матрицу тестовых результатов. Например, в матрице X_{4x3} представлены оценки четырех испытуемых по трем заданиям теста

$$X_{4x3} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

8.2 *Состав матрицы.* Тестовые результаты всегда мыслятся в форме матриц. При этом каждая матрица, как и человек, имеет имя. Например, матрица

$$X_{4x3} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ имеет имя } X_{4x3}$$

Это имя содержит название (X) и размер**4x3.** Размер матрицы X выражается символами. Здесь использованы символы 4X3. Первый (*4)* указывает на число строк (испытуемых) в матрице, а второй (*3*) –на число столбцов (заданий теста). Знак ***X*** означает, что матрица имеет размер 4 на 3 (4X3).

Строки матрицы представляют результаты испытуемых. Каждая строка отражает баллы каждого испытуемого. Соответственно, число строк в точности равно числу испытуемых (n). На первом месте стоит испытуемый номер один, на втором – №2, на третьем месте - №3, и на четвертом №4.

Столбцы матрицы представляют задания одного теста (или баллы нескольких тестов). Три столбца примера матрицы:

$X_{4x3} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ означают, что дано всего три задания, которые обычно располагаются по номерам: на первом месте задание №1, на втором - №2 и т.д.

Элементы матрицы – это числа, отражающие оценки каждого испытуемого, по каждому заданию.

Каждый элемент матрицы имеет своё уникальное имя. Если обозначить символом i номер испытуемого, а j номер задания, то элемент матрицы x_{ij} представляет общее имя балла испытуемого i по заданию j. Например, если первый, по счету, испытуемый ответил неправильно по восьмому заданию теста, то элемент с общим именем x_{ij} получает конкретное имя x_{18}, и принимается равным нулю. Это записывается

$x_{18} = 0$.

В примере матрицы: $X_{4x3} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

элементы $x_{12} = 1$, $x_{23} = 0$ и $x_{31} = 4$. Все элементы матриц отграничены скобками.

Матрицы могут отличаться одна от другой составом элементов, размерами и структурой.

Структура матрицы определяется названиями строк и столбцов, порядком расположения строк и столбцов элементов. А также связью элементов или их зависимостью.

При выполнении матричных расчетов большую роль играет вычисление таких структурных показателей матрицы, сумма диагональных элементов матрицы (trace) и значение её определителя.

Тестовая информация сохраняется в матрицах как результаты взаимодействия множеств испытуемых с множеством заданий.

Естественно, все получаемые исходные данные зависят от состава и от структуры этих двух множеств.

Состав испытуемых интересен с точки зрения уровня развития, знаний и т.п. у испытуемых.

Матричная алгебра - это название той части математики (алгебры), где изучаются виды и свойства различных матриц, определяются допустимые операции с ними.

Сама математика определяется как наука о множествах различных объектов, о количественных и качественных отношениях между ними.

8.3. *Виды матриц.*

Если число строк матрицы не равны числу её столбцов, то такую матрицу назовем прямоугольной. Примеры прямоугольных матриц:

$$X_{4x3} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad X_{3x4} = \begin{pmatrix} 3 & 2 & 1 & 5 \\ 2 & 1 & 0 & -2 \\ 4 & 1 & 0 & -5 \end{pmatrix}$$

Обычно в тестовой матрице число испытуемых всегда больше числа заданий. При этом рекомендуется иметь более чем пятикратное превышение числа испытуемых над числом заданий. Только в этом случае статистические выводы считаются более обоснованными. Из этого требования становится очевидным, что настоящая тестовая матрица, полученная по эмпирическим данным, всегда имеет больше испытуемых, чем заданий, а потому она всегда имеет вид прямоугольной матрицы. В учебных текстах из дидактических соображений обычно пользуются небольшими матрицами. Посмотрим другой пример небольшой тестовой матрицы[2] X_{13x10}

[2] Аванесов В.С. Композиция тестовых заданий. М.: Центр тестирования. 2002. стр. 160.

1	1	1	0	1	1	1	1	1	1
1	1	0	1	1	1	1	1	1	0
1	1	1	1	0	1	1	0	1	0
1	1	1	1	0	1	0	1	0	0
1	1	1	1	1	1	0	0	0	0
1	1	1	1	0	0	1	0	0	0
1	1	0	1	1	0	1	0	0	0
1	1	1	1	1	0	0	0	0	0
1	0	1	0	1	1	0	0	0	0
0	1	1	0	0	0	0	1	0	1
1	1	1	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0

В этой матрице по строкам предполагается расположение номеров испытуемых, а по столбцам – номера заданий. В явном виде это расположение выражено в таблице 7. С данными этой таблицы можно произвести ряд статистических расчетов.

Применение формул для определения коэффициентов корреляции даёт возможность получить корреляционную матрицу тестовых результатов. Эта матрица позволяет проверить структуру разрабатываемого теста, если под структурой понимать связь ответов испытуемых между заданиями и с сумой полученных баллов.

Если матрица содержит одинаковое (равное) число строк и столбцов (n=m), то она называется квадратной. Примером квадратной матрицы является корреляционная матрица по данным X_{13x10}.

$$R = \begin{pmatrix} 1,000 & -0,123 & -0,192 & 0,312 & 0,267 & 0,267 & 0,228 & -0,433 & 0,158 & -0,677 \\ -0,123 & 1,000 & 0,178 & 0,461 & -0,033 & -0,033 & 0,337 & 0,284 & 0,234 & 0,182 \\ -0,192 & 0,178 & 1,000 & 0,051 & -0,051 & 0,283 & -0,158 & 0,083 & -0,030 & 0,284 \\ 0,312 & 0,461 & 0,051 & 1,000 & 0,238 & 0,238 & 0,415 & -0,051 & 0,141 & -0,461 \\ 0,267 & -0,033 & -0,051 & 0,238 & 1,000 & 0,381 & 0,220 & 0,051 & 0,225 & 0,033 \\ 0,267 & -0,033 & 0,283 & 0,238 & 0,381 & 1,000 & 0,220 & 0,386 & 0,592 & 0,033 \\ 0,228 & 0,337 & -0,158 & 0,415 & 0,220 & 0,220 & 1,000 & 0,158 & 0,693 & 0,101 \\ -0,433 & 0,284 & 0,083 & -0,051 & 0,051 & 0,386 & 0,158 & 1,000 & 0,426 & 0,640 \\ 0,158 & 0,234 & -0,030 & 0,141 & 0,225 & 0,592 & 0,693 & 0,426 & 1,000 & 0,272 \\ -0,677 & 0,182 & 0,284 & -0,461 & 0,033 & 0,033 & 0,101 & 0,640 & 0,272 & 1,000 \end{pmatrix}$$

По главной диагонали корреляционной матрицы стоят единицы. Это означает, что корреляция тестовых баллов задания с тем же заданием всегда равна единице. Сумма элементов матрицы, стоящих на главной диагонали корреляционной матрицы теста называется английским словом trace (что иногда переводится неудачным словом «след матрицы»). В нашем случае trace = 10. Для повышения точности выделения латентных факторов в диагональные элементы вместо единиц рекомендуется ставить коэффициенты надежности тестов, или, что еще лучше, значения квадратов коэффициентов множественной корреляции каждого теста со всеми остальными тестами.

Если к матрице X_{13x10} добавить справа значения исходных тестовых баллов, а снизу – число правильных ответов на каждое задание, то получится расширенная матрица тестовых результатов X_{14x11}

$$X_{14x11}=\begin{pmatrix}1&1&1&0&1&1&1&1&1&1&9\\1&1&0&1&1&1&1&1&1&0&8\\1&1&1&1&0&1&1&0&1&0&7\\1&1&1&1&0&1&0&1&0&0&6\\1&1&1&1&1&1&0&0&0&0&6\\1&1&1&1&0&0&1&0&0&0&5\\1&1&0&1&1&0&1&0&0&0&5\\1&1&1&1&1&0&0&0&0&0&5\\1&0&1&0&1&1&0&0&0&0&4\\0&1&1&0&0&0&0&1&0&1&4\\1&1&1&0&0&0&0&0&0&0&3\\1&1&0&0&0&0&0&0&0&0&2\\1&0&0&0&0&0&0&0&0&0&1\\12&11&9&7&6&6&5&4&3&2&65\end{pmatrix}$$

Важно знать, что расширенная матрица, образованная в результате сложения строк и (или) столбцов, не подвергается факторному анализу по причине зависимости значений элементов последней строки и последнего столбца от значений остальных элементов матрицы.

Если в матрице X_{14x11} добавленные снизу строку и справа столбец отделить линиями, то мы получим пример блочной матрицы.

$$\left(\begin{array}{cccccccccc|c}
1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 9 \\
1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 8 \\
1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 7 \\
1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 6 \\
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 6 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 5 \\
1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 5 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 5 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 4 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 4 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\hline
12 & 11 & 9 & 7 & 6 & 6 & 5 & 4 & 3 & 2 & 65
\end{array}\right)$$

Последнему столбцу блочной матрицы можно дать имя Y . Этот вектор-столбец можно назвать зависимой переменной. Остальные столбцы матрицы X_1, X_2, ….X_{10} называются независимые векторы.

Числовой вектор Y_i можно представить как линейную функцию от множества переменных X_J: $Y_i = c_1X_1 + c_2X_2 + c_3X_3 +\ldots +c_mX_m$.,
где cj являются шкальным баллом задания. В множественном регрессионном анализе c_j называется весом каждого задания. Чем больше значение c_j у какого-либо задания, тем больший вклад этого задания в общую сумму баллов. Если значения c_j для всех заданий принимается равным единице, то тогда Y_i становится элементарной суммой баллов испытуемого по всем заданиям теста.

Если к матрице добавить слева идентификаторы испытуемых – фамилии или номера, то матрица превращается в таблицу данных (Табл.7)

Таблица данных   Табл. 7

№№ испытуемых	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Исходные тестовые баллы испытуемых Y_i
1.	1	1	1	0	1	1	1	1	1	1	9
2.	1	1	0	1	1	1	1	1	1	0	8
3.	1	1	1	1	0	1	1	0	1	0	7
4.	1	1	1	1	0	1	0	1	0	0	6
5.	1	1	1	1	1	1	0	0	0	0	6
6.	1	1	1	1	0	0	1	0	0	0	5
7.	1	1	0	1	1	0	1	0	0	0	5
8.	1	1	1	1	1	0	0	0	0	0	5
9.	1	0	1	0	1	1	0	0	0	0	4
10.	0	1	1	0	0	0	0	1	0	1	4
11.	1	1	1	0	0	0	0	0	0	0	3
12.	1	1	0	0	0	0	0	0	0	0	2
13.	1	0	0	0	0	0	0	0	0	0	1
R_j	**12**	**11**	**9**	**7**	**6**	**6**	**5**	**4**	**3**	**2**	**65 5**

Элементы таблицы отражают результаты тестирования: за каждый правильный ответ на соответствующее задание студенты получают один балл, за неправильный ответ - ноль. В таблице по строкам располагаются фамилии или номера испытуемых, по столбцам - номера заданий.

Применение формулы (2) для данных последнего столбца (У) таблицы дает средний тестовый балл испытуемых:

$$M_y = \frac{\sum Y_i}{N} = \frac{65}{13} = 5$$

Сложение этих элементов по строкам и по столбцам имеет педагогический смысл. Из последнего столбца таблицы видно, что больше правильных ответов у первого испытуемого, а меньше - у последнего. Элементы этого столбца указывают на значение тестового балла X, полученного каждым (любым) испытуемым, что обозначается символом статистик p_j и q_j.

Коррелирование данных таблицы 7 без значений строки R_j даёт расширенную корреляционную матрицу (табл. 8). Иногда значения предпоследнего столбца возводят в квадрат, что делается для оценки меры связи ответов испытуемых на задания с суммой баллов.

Табл.8.

ЗАДАНИЯ	1	2	3	4	5	6	7	8	9	10	r_{jy}	r^2_{jy}
1	1.000	0.123	-0.192	0.312	0.267	0.267	0.228	-0.433	0.158	-0.677	0.132	0.017
2	0.123	1.000	0.178	0.460	-0.033	-0.033	0.337	0.284	0.233	0.182	0.488	0.238
3	-0.192	0.178	1.000	0.051	-0.051	0.283	-0.158	0.083	-0.030	0.284	0.305	0.093
4	0.312	0.460	0.051	1.000	0.238	0.238	0.415	-0.051	0.141	-0.460	0.494	0.244
5	0.267	-0.033	-0.051	0.238	1.000	0.381	0.220	0.051	0.225	0.033	0.494	0.244
6	0.267	-0.033	0.283	0.238	0.381	1.000	0.220	0.386	0.592	0.033	0.706	0.498
7	0.228	0.337	-0.158	0.415	0.219	0.219	1.000	0.158	0.693	0.101	0.651	0.424
8	-0.433	0.284	0.083	-0.051	0.051	0.386	0.158	1.000	0.426	0.639	0.534	0.285
9	0.158	0.233	-0.030	0.141	0.225	0.592	0.693	0.426	1.000	0.272	0.752	0.565
10	-0.677	0.182	0.284	-0.460	0.033	0.033	0.101	0.639	0.272	1.000	0.293	0.086

В табл. 8 особое внимание обращается на значения предпоследнего столбца. Задания, имеющие корреляции около 0,3, обычно не включаются в тест. В нашем примере непригодным для теста можно считать первое задание, а в число заданий сомнительной или спорной ценности можно отнести третье задание.

Если в квадратной матрице все элементы, кроме диагональных, равны нулю, то такая матрица называется диагональной. Пример диагональной матрицы (D):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

Матрица, у которой все элементы равны нулю, называется нулевой (O)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Матрица, у которой все элементы ниже диагонали равны нулю, называется верхней треугольной матрицей.

$$\begin{pmatrix} 1 & 2 & 4 & 2 \\ 0 & 5 & 1 & 6 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Матрица, у которой все элементы выше диагонали равны нулю, называется нижней треугольной матрицей.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 4 & 1 & -1 & 0 \\ 2 & 6 & 8 & 1 \end{pmatrix}$$

Довольно часто корреляционные матрицы записываются в виде верхней или нижней треугольной матрицы.

Диагональная матрица, у которой все элементы главной диагонали равны единице, а остальные равны нулю, называется единичной матрицей (I). Например:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Единичная матрица играет очень важную роль в различных расчетах, и в том числе – в определении тест-образующих свойств заданий.

Матрицы называются равными, если каждый элемент одной матрицы равен соответствующему элементу другой матрицы. Возьмем две матрицы, A и B.

$$A_{4x3} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad B_{4x3} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Равенство двух матриц имеет место в случае, когда элементы a_{ij} матрицы $A_{n\,x\,m}$ и элементы $в_{ij}$ $B_{n\,x\,m}$ равны ($a_{ij} = в_{ij}$).

9. Применение статистических и математических методов на различных этапах тестового процесса.

Применение статистических методов является частью *тестового процесса*[3]. Можно выделить четыре основных этапа.

Первым этапом является разработка *заданий в тестовой форме.* Как в

[3] Понятие «тестовый процесс определено в статье: Аванесов В.С. Определение исходных понятий теории педагогических измерений// Педагогические Измерения, №2 2005г.

искусстве, овладение формой является необходимым, но недостаточным условием для разработки качественного теста. На первом этапе требуется подготовка в области преподаваемого предмета, знание тестовых форм, владение логикой и умение трансформировать фрагменты содержания учебной дисциплины в содержание заданий.

На втором этапе разрабатываются *тестовые задания.*

В отличие от распространенной в России и на Западе вопросно-ответной формы заданий, в качестве основы используется логика высказываний. Преимущества этой основы вытекают из отличия высказываний от вопросов: истинность или ложность высказывания легко определяется по логическим правилам, в то время как вопросы сами по себе ни истинны, ни ложны. Предлагаемая основа открывает возможности эффективной компьютеризации контроля знаний.

Для успешной деятельности на втором этапе от разработчиков потребуется, дополнительно, некоторая подготовка в области применения статистических методов, обработки и интерпретации данных.

На третьем этапе отбираются задания и создают тесты, повышают качество и эффективность теста. Наличие достаточного числа тестовых заданий позволяет перейти к разработке *теста как системы*, обладающей целостностью, составом и структурой. Путь к достижению этого идеала лежит через трудности создания качественных тестов, Разработка тестов начинается с анализа содержания преподаваемых знаний и овладения принципами формулирования тестовых заданий.

На третьем этапе от разработчика тестов потребуется некоторая математико-статистическая подготовка, знание основных тестовых теорий и методов многомерного статистического анализа, опыт правильной интерпретации тестовых результатов. Кроме того, потребуется умение тактично обсуждать с авторами явные и скрытые дефекты их заданий в тестовой форме, психологическая готовность участников тестового процесса к совместному поиску аргументов.

На четвертом этапе применяется одна из моделей математического измерения так называемой латентной переменной. На этом этапе уточняется содержание интересующей переменной, определяется шкала измерения, применяется компьютерная программа для определения и уточнения параметров заданий и уровня подготовленности испытуемых.

Как уже отмечалось в первой статье данного номера журнала. лучшей моделью педагогического измерения следует признать модель Г. Раша. Она позволяет получить значения интересующих параметров личности и заданий на одной и той же интервальной шкале.

И в заключение несколько слов об уместности использования статистики и ЭВМ. Изучение каких-либо методов статистических расчетов без понимания вопроса, где эти методы можно применять, а где нет, ведет обычно к многочисленным ошибкам. Особенно важно вначале выяснить, на какой шкале измерения получены данные. В зависимости от этого подбирается статистический метод расчета данных, который применим в одной шкале и неприменим для данных в другой шкале. Тем самым мы затронули важный вопрос соотношения измерения и статистики. Но это предмет другой статьи.