

В ПОРЯДКЕ ОБСУЖДЕНИЯ

TO BE DISCUSSED

Статья поступила в редакцию 19.01.2016

DOI 10.20915/2077-1177-2016-0-1-42-49

УДК 543.06:658.562

ПРИМЕНЕНИЕ КРИТЕРИЯ СТЬЮДЕНТА ДЛЯ ОЦЕНКИ РЕЗУЛЬТАТОВ МЕЖЛАБОРАТОРНЫХ СРАВНИТЕЛЬНЫХ ИСПЫТАНИЙ

К.В. Шаталов, А.В. Кириллова

Результаты межлабораторных сравнительных испытаний, как правило, представляют собой выборки малого объема. Поэтому для оценки качества результатов испытаний лабораторий, участвующих в межлабораторных сравнительных испытаниях, предложено использовать критерий Стьюдента, который связывает между собой три основных характеристики малой выборки: ширину доверительного интервала, соответствующую ему доверительную вероятность и объем выборки. На примерах межлабораторных сравнительных испытаний дизельного топлива показано, что использование критерия Стьюдента позволяет более точно оценивать качество результатов испытаний лабораторий, в частности выявлять и оценивать результаты, находящиеся вне доверительного интервала аттестованного значения образца для контроля. Расчет усредненного значения статистики Стьюдента для совокупности результатов испытаний позволяет оценивать качество работы лаборатории, ранжировать их, формировать рейтинг лабораторий на объективной основе.

Ключевые слова: межлабораторные сравнительные испытания, испытательные лаборатории, оценка результата испытаний, оценка качества работы лаборатории, аттестованное значение, доверительный интервал, критерий Стьюдента, Z-индекс.

✓ **Ссылка при цитировании:** Шаталов К.В., Кириллова А.В. Применение критерия Стьюдента для оценки результатов межлабораторных сравнительных испытаний // Стандартные образцы. 2016. № 1. С. 42–49. DOI 10.20915/2077-1177-2016-0-1-.....

Авторы:

ШАТАЛОВ К.В.

Начальник отдела квалификационных испытаний топлив и масел ФАУ «25 ГосНИИ химмотологии Минобороны России», доцент базовой кафедры «Химмотология горюче-смазочных материалов» ФГБОУВПО «РГУ нефти и газа имени И.М. Губкина», канд техн. наук
Российская Федерация, 121467, Москва,
ул. Молодогвардейская, 10
Тел.: 8 (499) 149-90-90
E-mail: 1499090@mail.ru

КИРИЛЛОВА А.В.

Инженер отдела квалификационных испытаний топлив и масел ФАУ «25 ГосНИИ химмотологии Минобороны России»
Российская Федерация, 121467, Москва,
ул. Молодогвардейская, 10
E-mail: nastya_kir@bk.ru

Главной задачей любой испытательной лаборатории является получение точных и достоверных результатов. Лаборатория должна гарантировать качественное проведение испытаний и располагать процедурами демонстрации своей способности, стабильно получать результаты испытаний с требуемой точностью [1]. В международной и отечественной практике признано, что наиболее действенной формой подтверждения испытательной лабораторией качества результатов измерений является участие в межлабораторных сравнительных испытаниях (далее – МСИ) [2–5].

Важным элементом МСИ является оценка результатов испытаний, представленных лабораториями-участниками. Наиболее часто для этого используют результаты расчета Z -индекса, который представляет собой относительную разность результата испытания и аттестованного значения образца для контроля, выраженную в единицах среднего квадратического отклонения погрешности [3–12]:

$$Z = \frac{X - C}{\sigma}, \quad (1)$$

где X – результат испытаний;

C – аттестованное значение образца для контроля;

σ – среднее квадратическое отклонение погрешности.

Допустимые значения Z -индекса устанавливают на основе закона нормального распределения – приемлемыми (удовлетворительными) считаются результаты, находящиеся в интервале $\pm 2\sigma$ от аттестованного значения образца для контроля, неприемлемыми (неудовлетворительными) считаются результаты, отклоняющиеся от аттестованного значения более чем на $\pm 3\sigma$ [2–7].

В ряде случаев применение Z -индексов не позволяет дать однозначную оценку результатов межлабораторных сравнительных испытаний. Во-первых, классическая теория ошибок, основанная на нормальном распределении, неприменима для обработки малого числа измерений, каковыми и являются результаты межлабораторных сравнительных испытаний [13]. Во-вторых, нет однозначного ответа, каким образом определять значения σ . Нормативные документы, которыми руководствуются провайдеры МСИ, предлагают различные способы определения среднего квадратического отклонения погрешности. Их наиболее полный перечень изложен в ГОСТ Р ИСО 13528 (раздел 6):

– на основе приписанного значения характеристики погрешности методики измерений;

– на основе значения желаемого для достижения лабораториями – участниками МСИ;

– на основе значения, полученного из общей модели воспроизводимости метода измерений (например, кривой Хорвица);

– на основе значения, вычисленного исходя из приписания приписанных характеристик прецизионности;

– на основе значения, вычисленного по результатам, полученным в ходе МСИ.

Однако необходимо учитывать, что значительное число методик измерений (например, нефти и нефтепродуктов) не имеют приписанных характеристик погрешности, а при их вычислении на основе приписанных характеристик прецизионности не учитывается систематическая погрешность методики. А значения среднего квадратического отклонения погрешности, вычисленного на основании экспериментальных данных, могут существенно различаться в зависимости от применяемого алгоритма вычисления, что приводит к существенным отклонениям в значениях Z -индексов. Так, используя алгоритмы, изложенные в нормативных документах [2, 4, 14], для одной и той же выборки экспериментальных данных были получены оценки σ и Z -индексов, отличающиеся более чем в 2 раза (табл. 1).

Из приведенного примера видно, что оценка результатов МСИ (а следовательно, и оценка работы той или иной лаборатории) на основе применения Z -индекса не является однозначной и может являться источником разного рода конфликтных ситуаций.

В математической статистике оценку малых выборок принято проводить на основе распределения Стьюдента, которое связывает между собой три основных характеристики выборочной совокупности: ширину доверительного интервала, соответствующую ему доверительную вероятность и объем выборки [13, 15]. В частности, на основе распределения Стьюдента можно получить надежную оценку доверительного интервала для аттестованного значения образца для контроля:

$$C \pm t_{p,f} \cdot \frac{S}{\sqrt{n}}, \quad (2)$$

где $t_{p,f}$ – табличное значение коэффициента Стьюдента, зависящее от заданной доверительной вероятности P и числа степеней свободы $f = n - 1$;

S – среднее квадратическое отклонение выборки из n результатов.

Для оценки результатов межлабораторных сравнительных испытаний целесообразно использовать следующую формулу расчета критерия Стьюдента:

$$t_i = \frac{|X_i - C|}{\sqrt{\frac{S^2}{N} + \frac{\Delta_{ar}^2}{3}}}, \quad (3)$$

где t_i – значение статистики Стьюдента для i -й лаборатории;

Таблица 1

Влияние алгоритма вычисления на значения Z-индекса

Шифр лаборатории	Кинематическая вязкость диз. топлива при 20 °С, мм²/с	Значения Z-индекса при вычислении σ по алгоритму		
		РМГ 61	ГОСТ Р ИСО 13528 прил. С	Р 50.2.011 прил. И
1 ИЛ	2,7776	1,36	1,54	3,04
2 ИЛ	2,7530	0,09	0,10	0,20
3 ИЛ	2,7510	0,21	0,23	0,46
4 ИЛ	2,7590	0,26	0,30	0,59
5 ИЛ	2,7490	0,32	0,37	0,73
6 ИЛ	2,7290	1,50	1,70	3,36
7 ИЛ	2,7650	0,62	0,70	1,38
8 ИЛ	2,7510	0,21	0,23	0,46
9 ИЛ	2,7240	1,80	2,03	4,02
10 ИЛ	2,7800	1,50	1,70	3,36
11 ИЛ	2,7550	0,03	0,03	0,06
12 ИЛ	2,7370	1,03	1,17	2,31
13 ИЛ	2,7700	0,91	1,03	2,04
14 ИЛ	2,7630	0,50	0,57	1,12
15 ИЛ	2,7540	0,03	0,03	0,07
Среднее значение	2,7545			
Расчетное значение σ		0,01698	0,01501	0,00759

X_i – результат испытаний i -й лаборатории;

C – аттестованное значение образца для контроля;

S^2 – выборочная дисперсия результатов МСИ;

N – количество лабораторий-участников МСИ;

$\Delta_{ат}$ – погрешность аттестованного значения образца для контроля [14].

Расчетное значение статистики Стьюдента t_i сравнивают с табличным значением критерия Стьюдента $t_{p,f}$ при числе степеней свободы $f = N - 1$ для доверительной вероятности $P = 0,95$ (табл. 2).

Если расчетное значение статистики Стьюдента меньше или равно табличному значению критерия Стьюдента: $t_i \leq t_{p,f}$, то результат i -й лаборатории считаем удовлетворительным.

Если расчетное значение статистики Стьюдента больше табличного значения критерия Стьюдента: $t_i > t_{p,f}$, то результат i -й лаборатории считаем неудовлетворительным.

Проведем сравнение эффективности рассмотренных способов оценки результатов испытаний на примере межлабораторных сравнительных испытаний, прове-

Таблица 2

Значения критерия Стьюдента для доверительной вероятности $P = 0,95$ [14]

$f = N - 1$	$t_{p,f}$	$f = N - 1$	$t_{p,f}$	$f = N - 1$	$t_{p,f}$
1	12,71	10	2,23	19	2,09
2	4,30	11	2,20	20	2,09
3	3,18	12	2,18	21	2,08
4	2,78	13	2,16	22	2,07
5	2,57	14	2,15	23	2,07
6	2,45	15	2,14	24	2,06
7	2,37	16	2,12	25	2,06
8	2,31	17	2,11	26	2,06
9	2,26	18	2,10	27	2,05

денных в 2015 году с лабораториями горючего Минобороны России. Рассмотрим результаты определения кинематической вязкости дизельного топлива при 20 °С по ГОСТ 33–2000 (табл. 3).

Обработку результатов, представленных лабораториями, в частности определение аттестованного значения, проводили по принятому в Минобороны РФ алгоритму [16]. Доверительный интервал аттестованного значения рассчитали по формуле (2).

В связи с тем что методика определения кинематической вязкости нефтепродуктов (ГОСТ 33–2000) не содержит приписанной характеристики погрешности, для расчета Z-индексов дополнительно провели определение среднего квадратического отклонения погрешности МСИ согласно РМГ 61–2010 [14].

Расчетные значения Z-индексов для всех пятнадцати лабораторий – участников МСИ находятся в интервале от 0,03 до 1,8, что меньше допустимого значения 2,0. Следовательно, согласно рекомендациям, изложенным

в документах [3–7], качество результатов испытаний для всех лабораторий, участвовавших в МСИ, необходимо признать удовлетворительным. Однако результаты семи из пятнадцати лабораторий – 1 ИЛ (2,7776 мм²/с); 6 ИЛ (2,7290 мм²/с); 7 ИЛ (2,7650 мм²/с); 9 ИЛ (2,7240 мм²/с); 10 ИЛ (2,7800 мм²/с); 12 ИЛ (2,7370 мм²/с); 13 ИЛ (2,7700 мм²/с) – выходят за пределы доверительного интервала для аттестованного значения 2,7456–2,7634 мм²/с, и возникает вопрос: можно ли считать такие результаты удовлетворительными?

Вычисления статистики Стьюдента показывают, что для шести лабораторий расчетное значение превышает табличное значение критерия Стьюдента – 1 ИЛ ($t_{\text{расчет.}} = 4,0$); 6 ИЛ ($t_{\text{расчет.}} = 4,42$); 9 ИЛ ($t_{\text{расчет.}} = 5,28$); 10 ИЛ ($t_{\text{расчет.}} = 4,41$); 12 ИЛ ($t_{\text{расчет.}} = 3,03$); 13 ИЛ ($t_{\text{расчет.}} = 2,68$). Следовательно, качество результатов испытаний практически для всех лабораторий, чьи результаты выходят за пределы доверительного интервала, следует признать неудовлетворительным.

Таблица 3

Оценка результатов МСИ по показателю «кинематическая вязкость» с использованием Z-индекса и критерия Стьюдента

Шифр лаборатории	Кинематическая вязкость при 20 °С, мм ² /с	С использованием Z-индекса			С использованием критерия Стьюдента		
		Z-индекс	Критерий	Оценка	$t_{\text{расчет.}}$	Критерий	Оценка
1 ИЛ	2,7776	1,36	≤ 2	удовл.	4,00	$> t$ табл.	неудовл.
2 ИЛ	2,7530	0,09	≤ 2	удовл.	0,26	$\leq t$ табл.	удовл.
3 ИЛ	2,7510	0,21	≤ 2	удовл.	0,61	$\leq t$ табл.	удовл.
4 ИЛ	2,7590	0,26	≤ 2	удовл.	0,78	$\leq t$ табл.	удовл.
5 ИЛ	2,7490	0,32	≤ 2	удовл.	0,95	$\leq t$ табл.	удовл.
6 ИЛ	2,7290	1,50	≤ 2	удовл.	4,42	$> t$ табл.	неудовл.
7 ИЛ	2,7650	0,62	≤ 2	удовл.	1,82	$\leq t$ табл.	удовл.
8 ИЛ	2,7510	0,21	≤ 2	удовл.	0,61	$\leq t$ табл.	удовл.
9 ИЛ	2,7240	1,80	≤ 2	удовл.	5,28	$> t$ табл.	неудовл.
10 ИЛ	2,7800	1,50	≤ 2	удовл.	4,41	$> t$ табл.	неудовл.
11 ИЛ	2,7550	0,03	≤ 2	удовл.	0,09	$\leq t$ табл.	удовл.
12 ИЛ	2,7370	1,03	≤ 2	удовл.	3,03	$> t$ табл.	неудовл.
13 ИЛ	2,7700	0,91	≤ 2	удовл.	2,68	$> t$ табл.	неудовл.
14 ИЛ	2,7630	0,50	≤ 2	удовл.	1,47	$\leq t$ табл.	удовл.
15 ИЛ	2,7540	0,03	≤ 2	удовл.	0,09	$\leq t$ табл.	удовл.
Аттестованное значение образца для контроля, C					2,7456		
Погрешность аттестованного значения образца для контроля, $\Delta_{\text{ат}}$					0,0070		
Доверительный интервал аттестованного значения					2,7456–2,7634		
Среднее квадратичное отклонение результатов МСИ, S					0,01597		
Выборочная дисперсия результатов МСИ, S^2					0,0003		
Среднее квадратическое отклонение погрешности МСИ, σ_{Δ}					0,0170		
Значение критерия Стьюдента, $t_{\text{табл}}$					2,15		

Приведенный пример наглядно показывает преимущества применения критерия Стьюдента для оценки качества результатов испытаний лабораторий. В случае отсутствия в методике испытаний приписанной характеристики погрешности расчет Z -индекса дает завышенную оценку лаборатории, не позволяет выявить результаты, находящиеся вне доверительного интервала аттестованного значения. Использование критерия Стьюдента позволяет более объективно оценивать качество результатов испытаний лабораторий, в частности выявлять и оценивать результаты, находящиеся вне доверительного интервала аттестованного значения.

В подавляющем большинстве случаев в ходе МСИ лаборатории проводят испытания образцов для контроля по ряду показателей качества, и требуется дать оценку качества работы лаборатории по совокупности результатов испытаний, полученных при проведении МСИ. Однозначной рекомендации в данном случае быть не может, так как необходимо оценивать совокупность результатов, полученных различными методиками и по различным показателям качества, говоря другими словами, необходимо оценивать результаты, принадлежащие различным выборочным (генеральным) совокупностям. Используя метод экспертных оценок и базируясь на более чем 50-летнем опыте проведения МСИ с лабораториями горячего Минобороны России, считаем, что объективная оценка качества работы лаборатории по совокупности результатов испытаний, полученных при проведении МСИ, может быть получена путем вычисления среднего значения статистики Стьюдента по всем показателям качества для того или иного образца для контроля:

$$t = \frac{\sum_{j=1}^n t_j}{n}, \quad (4)$$

где t_j – значение статистики Стьюдента для j -го показателя качества;

n – количество показателей качества.

Среднее значение статистики Стьюдента t сравнивают с табличным значением критерия Стьюдента $t_{p,f}$ при числе степеней свободы $f = n - 1$ для доверительной вероятности $P = 0,95$ (табл. 1).

Если среднее значение статистики Стьюдента меньше или равно табличному значению критерия Стьюдента:

$t > t_{p,f}$, то качество работы лаборатории считаем удовлетворительным.

Если среднее значение статистики Стьюдента больше табличного значения критерия Стьюдента: $t > t_{p,f}$, то качество работы лаборатории считаем неудовлетворительным.

Оценка качества работы лаборатории по совокупности результатов испытаний с использованием среднего значения статистики Стьюдента на примере межлабораторных сравнительных испытаний, проведенных в 2015 году с лабораториями горячего Минобороны России, представлена в табл. 4.

Использование усредненного значения статистики Стьюдента для оценки совокупности результатов испытаний также позволяет ранжировать лаборатории по качеству работы, формировать рейтинг лабораторий на объективной основе. В частности, для примера, представленного в табл. 4, можно ранжировать лаборатории следующим образом: 2 ИЛ ($t = 0,83$); 5 ИЛ ($t = 0,95$); 3 ИЛ ($t = 0,98$); 4 ИЛ ($t = 1,09$); 11 ИЛ ($t = 1,16$); 8 ИЛ ($t = 1,30$); 13 ИЛ ($t = 1,49$); 10 ИЛ ($t = 1,53$); 6 ИЛ ($t = 1,57$); 15 ИЛ ($t = 1,63$); 14 ИЛ ($t = 2,69$); 7 ИЛ ($t = 3,02$); 1 ИЛ ($t = 3,65$); 12 ИЛ ($t = 4,88$); 9 ИЛ ($t = 11,14$).

Таким образом, в связи с тем что статистическую оценку малых выборок принято проводить на основе распределения Стьюдента, для оценки качества результатов испытаний, например нефтепродуктов, лабораториями, участвующими в межлабораторных сравнительных испытаниях, предложено использовать критерий Стьюдента, который связывает между собой три основных характеристики выборочной совокупности: ширину доверительного интервала, соответствующую ему доверительную вероятность и объем выборки. Это позволяет более объективно и однозначно оценивать качество результатов испытаний, получаемых той или иной лабораторией, в частности выявлять и оценивать результаты, находящиеся вне доверительного интервала аттестованного значения образца для контроля. Расчет усредненного значения статистики Стьюдента для совокупности результатов испытаний позволяет оценивать качество работы лаборатории, ранжировать их, формировать рейтинг лабораторий на объективной основе.

ЛИТЕРАТУРА

1. ГОСТ ИСО/МЭК 17025–2009 Общие требования к компетентности испытательных и калибровочных лабораторий. М.: Стандартиформ, 2012. 27 с.
2. Р 50.4.006–2002 Рекомендации по аккредитации. Межлабораторные сравнительные испытания при аккредитации и инспекционном контроле испытательных лабораторий. Методика и порядок проведения. М.: Стандартиформ, 2009. 41 с.

Таблица 4

Оценка качества работы лабораторий по результатам МСИ образца дизельного топлива с использованием критерия Стьюдента

Шифр лабора- тории	Значения статистики Стьюдента для показателей качества													Оценка на основе статистики Стьюдента		
	Плотность при 20 °С, кг/м³	Плотность при 15 °С, кг/м³	Температура вспышки в закрытом тигле, °С	Температура 40 % отгона, °С	Температура 50 % отгона, °С	Температура 90 % отгона, °С	Температура 95 % отгона, °С	Температура помутнения, °С	Температура застывания, °С	Кинематическая вязкость при 20 °С, мм²/с	Концентрация фактиче- ских смол, мг/100 см³	Кислотность, мг КОН/100 см³	Значе- ние pH	Сред- нее	Таблич- ное значе- ние	Оценка
1 ИЛ	1,67	4,85	3,99	2,51	2,16	4,04	4,62	4,28	6,55	4,00	2,34	4,46	1,95	3,65	2,18	неудовл.
2 ИЛ	0,36	0,59	0,87	0,65	0,36	0,32	0,11	2,20	1,05	0,26	1,41	1,57	0,98	0,83	2,18	удовл.
3 ИЛ	0,30	1,01	0,10	1,56	1,16	1,69	0,43	0,35	2,00	0,61	0,46	1,26	1,80	0,98	2,18	удовл.
4 ИЛ	0,52	0,47	0,10	0,65	1,37	0,32	0,28	2,43	1,39	0,78	4,21	1,00	0,64	1,09	2,18	удовл.
5 ИЛ	0,30	1,01	0,68	1,11	0,96	1,19	0,66	0,35	1,39	0,95	0,46	1,16	2,15	0,95	2,18	удовл.
6 ИЛ	0,30	1,01	0,10	2,05	1,65	2,17	2,65	1,50	0,17	4,42	3,29	0,35	0,75	1,57	2,18	удовл.
7 ИЛ	3,15	6,97	1,65	2,51	2,16	1,55	12,10	0,35	1,39	1,82	0,46	2,51	2,60	3,02	2,18	неудовл.
8 ИЛ	0,69	1,65	1,46	2,01	1,87	1,56	0,90	0,35	2,00	0,61	0,46	1,26	2,03	1,30	2,18	удовл.
9 ИЛ	1,78	3,67	7,68	2,51	3,16	0,94	1,07	4,05	0,17	5,28	108,00	3,56	2,92	11,14	2,18	неудовл.
10 ИЛ	1,61	3,13	0,10	0,79	1,25	1,67	1,07	0,58	0,17	4,41	0,46	3,22	1,44	1,53	2,18	удовл.
11 ИЛ	0,03	0,47	1,46	0,65	0,36	0,92	1,29	2,20	0,44	0,09	2,39	1,76	2,96	1,16	2,18	удовл.
12 ИЛ	0,63	0,06	1,65	1,11	2,88	12,14	8,39	7,06	2,27	3,03	14,50			4,88	2,23	неудовл.
13 ИЛ	0,20	1,01	3,21	0,65	0,65	0,92	1,07	0,35	1,05	2,68	0,09	5,97	1,58	1,49	2,18	удовл.
14 ИЛ	3,15	6,97	1,46	3,36	3,38	4,05	3,27	1,50	1,05	1,47	2,34	0,75	2,27	2,69	2,18	неудовл.
15 ИЛ	0,13	0,47	3,02	2,46	2,37	2,81	1,69	0,35	1,05	0,09	3,29	1,86	1,57	1,63	2,18	удовл.

3. Р 50.2.011–2005 Рекомендации по метрологии. Государственная система обеспечения единства измерений. Проверка квалификации испытательных (измерительных) лабораторий, осуществляющих испытания веществ, материалов и объектов окружающей среды (по составу и физико-химическим свойствам), посредством межлабораторных сличений. М.: Федеральное агентство по техническому регулированию и метрологии, 2005. 54 с.
4. ГОСТ Р ИСО 13528–2010 Статистические методы. Применение при экспериментальной проверке компетентности посредством межлабораторных сравнительных испытаний. М.: Стандартинформ, 2012. 61 с.
5. ISO/IEC Guide 43-1:1997 Proficiency testing by interlaboratory comparisons. Part 1. Development and operation of proficiency testing schemes.
6. Kuselman I., Fajgelj A. IUPAC/CITAC Guide: Selection and use of proficiency testing schemes for a limited number of participants – chemical analytical laboratories (IUPAC Technical Report) // Pure Appl. Chem. 2010. Vol. 82. № 5. P. 1099–1135.
7. Thompson M., Ellison S.L.R., Wood R. The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report) // Pure Appl. Chem. 2006. Vol. 78. № 1. P. 145–196.
8. РМГ 58–2003 ГСИ. Оценка качества работы испытательной лаборатории пищевых продуктов и продовольственного сырья. Методика внешнего контроля точности результатов испытаний. М.: Изд-во стандартов, 2004. 8 с.
9. Трубаева Л.В., Лоханина С.Ю. Межлабораторные сравнительные испытания как средство подтверждения технической компетентности лабораторий Удмуртской республики, проводящих испытания нефти // Вестник Удмурт. ун-та. 2011. № 1. С. 99–110.
10. Голынец О.С., Собина Е.П. Проведение межлабораторных сравнительных испытаний комбикорма // Стандартные образцы. 2007. № 1. С. 24–32.
11. Межлабораторные сравнительные испытания сигарет / С.К. Кочеткова [и др.] // Стандартные образцы. 2011. № 3. С. 34–47.
12. Степановских В.В., Колпакова Е.К., Комина И.Г. Межлабораторные сравнительные испытания металлургических материалов // Стандартные образцы. 2012. № 4. С. 57–63.
13. Налимов В.В. Применение математической статистики при анализе вещества. М.: Физматиздат, 1960. 430 с.
14. РМГ 61–2010 Государственная система обеспечения единства измерений. Показатели точности, правильности, прецизионности методик количественного химического анализа. Методы оценки. М.: Стандартинформ, 2012. 62 с.
15. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. М.: ФИЗМАТЛИТ, 2006. 816 с.
16. Шаталов К.В., Сорокова И.И. Межлабораторные сравнительные испытания нефтепродуктов в испытательных лабораториях Минобороны России // Тонкие химические технологии. 2014. Т. 9. № 2. С.104–111.

APPLICATION OF STUDENT T-TEST FOR EVALUATION OF INTERLABORATORY COMPARATIVE TESTS RESULTS

K.V. Shatalov, A.V. Kirillova

FAE «The 25th State Research Institute of Chemmotology
of Ministry of Defense of the Russian Federation»
10, Molodogvardeyaskaya str., Moscow, Russia, 121467
Tel.: 8 (499) 149-90-90. E-mail: 1499090@mail.ru

The results of interlaboratory comparative tests typically represent a small sample size. Therefore, to evaluate the quality of test results of laboratories participating in interlaboratory comparative tests, it is proposed to use Student t-test, which connects three main characteristics of small sample: the width of the confidence interval, the corresponding confidence probability and sample size. The examples of interlaboratory comparative tests of diesel show, that the use Student test allows to more objectively evaluate the quality of tests results of laboratories, in particular to identify and assess the results outside the confidence interval of the certified value of the sample for inspection. The calculation of average values of Student statistic for the aggregate test results allows to evaluate the quality of the laboratory, to range them, to form a ranking of laboratories on an objective basis.

Key words: interlaboratory comparative tests, testing laboratories, evaluation of tests result, quality assessment of laboratory workflow, certified value, confidence interval, Student t-test, Z-index.

- ✓ **When quoting reference:** Shatalov K.V., Kirillova A.V. *Primenenie kriteriia St'iudenta dlia otsenki rezul'tatov mezhlaboratornykh sravnitel'nykh ispytaniy* [Application of Student t-test for evaluation of interlaboratory comparative tests results]. *Standartnye obrazcy – Reference materials*, 2016, No. 1, pp. 42–49. (In Russian). DOI 10.20915/2077-1177-2016-0-1-42-49.

REFERENCE

1. GOST ISO/MEK 17025–2009 Obshhie trebovaniya k kompetentnosti ispytel'nykh i kalibrovochnykh laboratorij [Interstate Standard ISO/MEK 17025–2009 General requirements for the competence of testing and calibration laboratories]. Moscow, Standartinform Publ., 2012, 27 p. (In Russian).
2. R 50.4.006–2002 Rekomendatsii po akkreditatsii. Mezhlaboratornye sravnitel'nye ispytaniia pri akkreditatsii i inspektsionnom kontrole ispytel'nykh laboratorij. Metodika i poriadok provedeniia [Recommendations on accreditation P 50.4.006–2002. Interlaboratory comparison tests for the accreditation and conformity surveillance of testing laboratories. Procedure and order of carrying out the work]. Moscow, Standartinform Publ., 2009, 41 p. (In Russian).
3. R 50.2.011–2005 Rekomendatsii po metrologii. Gosudarstvennaia sistema obespecheniia edinstva izmerenij. Proverka kvalifikatsii ispytel'nykh (izmeritel'nykh) laboratorij, osushchestvliaiushchikh ispytaniia veshhestv, materialov i ob'ektov okruzhaiushchej sredy (po sostavu i fiziko-khimicheskim svoystvam), posredstvom mezhlaboratornykh slichenij [Recommendations on metrology P 50.2.011–2005. State system of ensuring unity of measurements. Check of qualification of the test (measuring) laboratories which are carrying out tests of substances, materials and objects of environment (on structure and physical and chemical properties), by means of interlaboratory checkings]. Moscow, Federal'noe agentstvo po tekhnicheskomu regulirovaniyu i metrologii Publ., 2005, 54 p. (In Russian).
4. GOST R ISO 13528–2010 Statisticheskie metody. Primenenie pri éksperimental'noj provere kompetentnosti posredstvom mezhlaboratornykh sravnitel'nykh ispytaniy [GOST R ISO 13528–2010 Statistical methods for use in proficiency testing by interlaboratory comparisons]. Moscow, Standartinform Publ., 2012, 61 p. (In Russian).
5. ISO/IEC Guide 43-1:1997 Proficiency testing by interlaboratory comparisons. Part 1. Development and operation of proficiency testing schemes.
6. Kuselman I., Fajgelj A. IUPAC/CITAC Guide: Selection and use of proficiency testing schemes for a limited number of participants – chemical analytical laboratories (IUPAC Technical Report). *Pure Appl. Chem.*, 2010, vol. 82, No 5, pp. 1099–1135.
7. Thompson M., Ellison S.L.R., Wood R. The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report). *Pure Appl. Chem.*, 2006, vol. 78, No 1, pp. 145–196.
8. RMG 58–2003 GSI. Otsenka kachestva raboty ispytel'noj laboratorii pishchevykh produktov i prodovol'stvennogo syr'ia. Metodika vneshnego kontrolya tochnosti rezul'tatov ispytaniy [Recommendations on Interstate standardization 58–2003 GSI. Evaluation of the working quality of foodstuff testing laboratory. Methods for development of the external control of test results accuracy]. Moscow, Izd-vo standartov Publ., 2004, 8 p. (In Russian).
9. Trubacheva L.V., Lokhanina S.Yu. Mezhlaboratornye sravnitel'nye ispytaniia kak sredstvo podtverzhdeniia tekhnicheskoy kompetentnosti laboratorij Udmurtskoj Respubliki, provodiashchikh ispytaniia nefti [Interlaboratory comparative researches as means of confirmation of technical competence of the laboratories of the Udmurt Republic which are carrying out oil tests]. *Vestnik Udmurtskogo universiteta – Bulletin of the Udmurt university*, 2011, No. 1, pp. 99–110. (In Russian).
10. Golynets O. S., Sobina E.P. Provedenie mezhlaboratornykh sravnitel'nykh ispytaniy kombikorma [Carrying out interlaboratory comparative researches of compound feed]. *Standartnye obrazcy – Reference material*, 2007, No.1, pp. 24–32. (In Russian).
11. Kochetkova S.K. [et al.]. Mezhlaboratornye sravnitel'nye ispytaniia sigaret [Interlaboratory comparative researches of cigarettes]. *Standartnye obrazcy – Reference material*, 2011, No. 3, pp. 34–47. (In Russian).
12. Stepanovskikh V.V., Kolpakova E.K., Komina I.G. Mezhlaboratornye sravnitel'nye ispytaniia metallurgicheskikh materialov [Interlaboratory comparative researches of metallurgical materials]. *Standartnye obrazcy – Reference material*, 2012, No. 4, pp. 57–63. (In Russian).
13. Nalimov V. V. *Primenenie matematicheskoy statistiki pri analize veshhestva* [Application of mathematical statistics in the analysis of substance], Moscow, Fizmatizdat Publ., 1960, 430 p. (In Russian).
14. Shatalov K.V., Sorokovova I.I. Mezhlaboratornye sravnitel'nye ispytaniia nefteproduktov v ispytel'nykh laboratoriyakh Minoborony Rossii [Interlaboratory comparative researches of oil products in test laboratories of the Ministry of Defence of the Russian Federation]. *Tonkie khimicheskie tekhnologii – Thin chemical technologies*, 2014, vol. 9, No. 2, pp. 104–111. (In Russian).
15. RMG 61–2010 Gosudarstvennaia sistema obespecheniia edinstva izmerenij. Pokazateli tochnosti, pravil'nosti, pretsizionnosti metodik kolichestvennogo khimicheskogo analiza. Metody otsenki [Recommendations on Interstate standardization 61–2010 State system of ensuring unity of measurements. Indicators of accuracy, correctness, pretsizionnost of techniques of the quantitative chemical analysis. Assessment methods], Moscow, Standartinform Publ., 2012, 62 p. (In Russian).
16. Kobzar A.I. *Prikladnaia matematicheskaya statistika. Dlia ingenerov i nauchnykh rabotnikov*. [Applied mathematical statistics for engineers and scientists] Moscow, Fizmatizdat Publ., 2006, 816 p. (in Russian).