УДК 621.317.088

Викулин И.М., Горбачев В.Э., Коробицын Б.В., Криськив С.К.
Vikulin I.M., Gorbachev V.E., Korobitzin B.V., Kryskiv S.K.

ОЦЕНКА ПРИГОДНОСТИ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ И ИСКЛЮЧЕНИЕ АНОМАЛЬНЫХ ЗНАЧЕНИЙ

THE MARK OF USEFUL MEASUREMENT RESULTS AND EXCEPTION ANOMALOUS VALUES

Аннотация. В статье рассмотрены вопросы выявления аномальных результатов измерений. Показано применение различных критериев аномальности. Проведено сравнение критериев и даны рекомендации по их применению.

Summary. The articles deals with the problems of detection of anomalous measurement results. The applications of various criteria of anomalousness are demonstrated. The comparison of the criteria is carried out and recommendations on their applications are given.

Измерение – одно из самых распространенных действий, выполняемых как на производстве, так и в научных исследованиях во всех областях человеческой деятельности. Непрерывно растущие требования к точности измерений удовлетворяются не только за счет повышения точности приборов, но и путем применения более эффективных методов обработки результатов измерений.

В литературе вопросам обработки результатов измерений уделяется достаточно много внимания. Многие источники, специально посвященные методам обработки измерительной информации, изложены слишком сложно. Другие, содержащие описание практических приемов обработки результатов измерений, не дают их сравнительной характеристики и рекомендаций по применению того или иного приема.

Измерениям нередко присущи большие отклонения некоторых результатов. При этом сразу нельзя утверждать, что такие отклонения являются частью выборки из генеральной совокупности или же промахами.

Существует целый ряд методов, предназначенных для обработки результатов, распределенных произвольным образом, так называемые робастные методы [1]. Но их применение, как правило, сопряжено с дополнительными сложными вычислениями, не всегда доступными в метрологической практике. Однако в классической статистике также имеются методы выявления аномальных результатов измерений – промахов. При этом используется двухступенчатая процедура: выявление и исключение промахов и последующее применение к исправленным результатам классических процедур оценивания.

Целью настоящей работы является сравнительный анализ различных методов выявления аномальных значений (промахов) применительно к лабораторной практике и выработка рекомендаций по их применению.

В производственной практике число измерений обычно не превышает десяти. В связи с этим методы определения функции распределения недостаточно селективны, чтобы, например, нормальное распределение можно было бы отличить от равномерного.

В общем случае исходят из того, что погрешности измерений обусловлены множеством причин, действия которых суммируются. Из предельной теоремы теории вероятностей следует заключение, что по крайней мере в принципе распределение должно быть нормальным. Значимость нормального распределения определяется тем, что для случайных погрешностей почти всегда его можно принять. Только в том случае, когда имеется много измеренных значений или в других обоснованных случаях, оправдано говорить о других распределениях [2].

Рассмотрим случай обработки результатов измерений, число которых больше двух, но меньше девяти по правилам, рекомендуемым для малых выборок [3].

Реальные данные редко соответствуют теоретическому распределению, так как могут быть засорены выбросами или промахами. В этих случаях лишь медиана является устойчивой оценкой, применимой в широкой области.

Статистическим критерием значимости для выявления аномальных результатов в малой выборке служит критерий Q:

$$Q = \left(x_n^* - x_{n-1}^*\right)/R, \tag{1}$$

где x_n^* – аномальное значение, x_{n-1}^* – значение результата, соседнего с x_n^* в вариационном ряду; R – размах выборки – разность между значениями крайних членов вариационного ряда. В табл. 1 приведены значения Q - критерия в зависимости от объема выборки и необходимой доверительной вероятности P. Если экспериментальное значение $Q_{эксп}$, определенное по формуле (1), превосходит значение *P,* то результат x_n^* отклоняется с вероятностью, равной доверительной. В ином случае аномальные результаты в выборке считают отсутствующими.

Таблица 1 – Значение критерия Q для различных значений доверительной вероятности

Объем выборки, n	P		
	0,90	0,95	0,99
3	0,89	0,94	0,99
4	0,68	0,77	0,89
5	0,56	0,64	0,76
6	0,48	0,56	0,70
7	0,43	0,51	0,64
8	0,40	0,48	0,58

Пример 1. Пусть результаты измерений представлены в виде вариационного ряда: 1,17; 1,20; 1,23; 1,30; 1,35; 1,43; 1,73. Определим по Q - критерию, является ли значение 1,73 при $P = 0{,}95$ аномально большим.

Согласно (1)

$$Q_{\text{эксп}} = \frac{x_8^* - x_7}{x_8^* - x_1} = \frac{1{,}73 - 1{,}43}{1{,}73 - 1{,}17} = 0{,}537 .$$

Табличные значения Q – критерия для $n = 8$ и $P = 0{,}95$ равно 0,48. Так как $0{,}537 > 0{,}48$, результат 1,73 следует считать аномальным. Среднее значение $\overline{x}$ для восьми результатов равно 1,33, а после исключения аномального значения $\overline{x} = 1{,}27$. Медиана для восьми результатов равна 1,26, а для семи – 1,23. Очевидно, корректнее оценить результат измерений медианой, равно 1,23.

Проверка x_7 по Q – критерию подтверждает нормальность этого результата.

Для достаточно больших объемов выборки $(n > 8)$ значения среднего $\overline{x}$ и среднеквадратичного отклонения (СКО) S могут быть вычислены уже с большой вероятностью. Поэтому статистический критерий при вычислении аномальных результатов должен включать величины $\overline{x}$ и S.

Рассмотрим методику обнаружения аномальных результатов (промахов), изложенную А.Н. Зайделем [4] (вариант 1). Если мы делаем ряд одинаковых измерений, подверженным случайным ошибкам, то в этом ряду могут встретиться измерения и с очень большими случайными ошибками. Но большие ошибки имеют малую вероятность, и мы склонны приписать такой результат промаху и отбросить его как заведомо неверный. Однако такой метод отбрасывания результатов, которые нам кажутся слишком сильно выпадающими из других измерений, порочен. Так легко получить завышенную и фиктивную точность измерений. Вероятность того, что из десяти измерений хотя бы одно будет случайно отличаться от среднего более чем на $3S$, будет 0,03 или 3%, а при 100 измерениях – 30%.

Обычно число измерений не очень велико и редко больше 10…20. При этом точное значение S неизвестно, следовательно, отбрасывать измерения, отличающиеся от среднего более чем на $3S$ нельзя.

Для оценки вероятности β случайного появления выскакивающих значений в ряду n измерений (для $n < 25$) на основании результатов, даваемых теорией вероятностей, была составлена табл. 2. При n*,* большем 25, можно для расчетов такого рода положить $S_n = \sigma$ и оценку β делать, пользуясь соотношением

$$\beta \approx \alpha^n . \tag{2}$$

Здесь α – доверительная вероятность, определяемая для нормального распределения. Для применения табл. 2 мы вычисляем среднее арифметическое $\bar{x}$ и СКО S из всех измерений, включая подозреваемое $x_к$, которое, на наш взгляд, недопустимо велико или мало.

Вычисляем относительное уклонение этого измерения от среднего арифметического, выраженное в долях СКО:

$$\upsilon_{\text{макс}} = \left| \frac{\bar{x} - x_к}{S} \right|. \tag{3}$$

По табл. 2 находим, какой вероятности β соответствует полученное значение $\upsilon_{\text{макс}}$. Разумеется, следует договориться, при каких значениях β мы будем отбрасывать измерения.

Табл. 2 составлена так, что наименьшее помещенное в ней значение β равно 0,01. Оставлять измерения, вероятность появления которых меньше этой величины, обычно нецелесообразно.

Таблица 2 – Оценка выскакивающих результатов $\upsilon_{\text{макс}} = \left| \frac{\bar{x} - x_к}{S} \right|$ в ряду n измерений для вероятности β

n	β			
	0,1	0,05	0,025	0,01
3	1,41	1,41	1,41	1,41
4	1,65	1,69	1,71	1,72
5	1,79	1,87	1,92	1,96
6	1,89	2,00	2,07	2,13
7	1,97	2,09	2,18	2,27
8	2,04	2,17	2,27	2,37
9	2,10	2,24	2,35	2,46
10	2,15	2,29	2,41	2,54
11	2,19	2,34	2,47	2,61
12	2,23	2,39	2,52	2,66
13	2,26	2,43	2,56	2,71
14	2,30	2,46	2,60	2,76
15	2,33	2,49	2,64	2,80
16	2,35	2,52	2,67	2,84
17	2,38	2,55	2,70	2,87
18	2,40	2,58	2,73	2,90
19	2,43	2,60	2,75	2,93
20	2,45	2,62	2,78	2,96
21	2,47	2,64	2,80	2,98
22	2,49	2,66	2,82	3,01
23	2,50	2,68	2,84	3,03
24	2,52	2,70	2,86	3,05
25	2,54	2,72	2,88	3,07

Таблица 3 – Ряд измерений длины l

№ измерения	l
1	258,5
2	255,4
3	256,6
4	256,7
5	257,0
6	256,5
7	256,7
8	255,3
9	256,0
10	266,0
11	256,3
12	256,5
13	256,0
14	256,3
15	256,9

Возьмем в качестве примера ряд измерений некоторой длины (табл. 3). В нашем примере $l_{10} = 266$, $\upsilon_{\text{макс}}$ получается равным

$$\frac{266{,}0 - 257{,}1}{2{,}6} = 3{,}42 \cdot$$

Наибольшее значение $\upsilon_{макс}$ для $n = 15$, приведенное в табл. 2, равно 2,80, чему соответствует $\beta = 0{,}01$. Так как с ростом $\upsilon_{\text{макс}}$ значение β уменьшается, то при $\upsilon_{\text{макс}} = 3{,}42$ β должно быть значительно меньше 0,01. Такие значения β отсутствуют в таблице. Из того, что $\beta << 0{,}01$, следует, что результат 266,0 надо отбросить, считая его промахом.

В оставшемся ряду представляется также подозрительным результат 255,5. Для него $\upsilon_{\text{макс}}$ получается равным

$$\frac{258{,}5-256{,}5}{2}=1{,}0\,.$$

Из табл. 2 видно, что этому значению соответствует $\beta > 0{,}1$ и результат 258,5, разумеется, нужно оставить.

Рассмотрим теперь применение критерия Шовене [5] к задаче обнаружения аномальных результатов измерений (вариант 2). Предположим, мы делаем n измерений $x_1,\dots,x_n$ одной и той же величины x. Учитывая все n измерений, вычислим $\bar{x}$ и S. Если один из результатов измерений отличается от $\bar{x}$ настолько, что представляется подозрительным $\left(x_{\text{под}}\right)$, то сначала вычислим

$$t_{\text{под}}=\frac{x_{\text{под}}-\bar{x}}{S}, \tag{4}$$

где $t_{\text{под}}$ – число СКО, на которое $x_{\text{под}}$ отличается от $\bar{x}$. Затем находим из табл. 4 вероятность P в долях (вне $t_{\text{под}}\cdot S$) того, что нормальное измерение будет отличаться от $\bar{x}$ на $t_{\text{под}}$ или более СКО:

$$P\,(\text{вне } t_{\text{под}}\cdot S)=1-P/100\,. \tag{5}$$

Таблица 4 – Вероятность нахождения величины $x_{\text{под}}$ в пределах $\left[\bar{x}-t_{\text{под}}\cdot S;\bar{x}+t_{\text{под}}\cdot S\right]$

P,% / t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,00	0,80	1,60	2,39	3,19	3,99	4,78	5,58	6,38	7,17
0,1	7,97	8,86	9,55	10,34	11,13	11,92	12,71	13,50	14,28	15,07
0,2	15,85	16,63	17,41	18,19	18,97	19,74	20,51	21,28	22,05	22,82
0,3	23,58	24,34	25,10	25,86	26,61	27,37	28,12	28,86	29,61	30,35
0,4	31,08	31,82	32,55	33,28	34,01	34,73	35,45	36,16	36,88	37,59
0,5	38,59	38,99	39,69	40,39	41,08	41,77	42,45	43,13	43,81	44,48
0,6	45,15	45,81	46,47	47,13	47,78	48,43	49,07	49,71	50,35	50,98
0,7	51,61	52,23	52,85	53,46	54,07	54,67	55,27	55,87	56,46	57,05
0,8	57,03	58,21	58,78	59,35	59,91	60,47	61,02	61,57	62,11	62,65
0,9	63,19	63,72	64,24	64,76	65,28	65,79	66,29	66,80	67,29	67,78
1,0	68,27	68,75	69,23	69,70	70,17	70,63	71,09	71,54	71,99	72,43
1,1	72,87	73,30	73,73	74,15	74,57	74,99	75,40	75,80	76,20	76,60
1,2	76,99	77,37	77,75	78,13	78,50	78,87	79,23	79,59	79,95	80,29
1,3	80,64	80,98	81,32	81,65	81,98	82,30	82,62	82,93	83,24	83,55
1,4	83,85	84,15	84,44	84,73	85,01	85,29	85,57	85,84	86,11	86,38
1,5	86,64	86,90	87,15	87,40	87,64	87,89	88,12	88,36	88,59	88,82
1,6	89,04	89,26	89,48	89,69	89,90	90,11	90,31	90,51	90,70	90,90
1,7	91,09	91,27	91,46	91,64	91,81	91,99	92,16	92,33	92,49	92,65
1,8	92,81	92,97	93,12	93,28	93,42	93,57	93,71	93,85	93,99	94,12
1,9	94,26	94,39	94,51	94,64	94,76	94,88	95,00	95,12	95,23	95,34
2,0	95,45	95,56	95,66	95,76	95,86	95,96	96,06	96,15	96,25	96,34
2,1	96,43	96,51	96,60	96,68	96,76	96,84	96,92	97,00	97,07	97,15
2,2	97,22	97,29	97,36	97,43	97,49	97,56	97,62	97,68	97,74	97,80
2,3	97,86	97,91	97,97	98,02	98,07	98,12	98,17	98,22	98,27	98,32
2,4	98,36	98,40	98,45	98,49	98,53	98,57	98,61	98,65	98,69	98,72
2,5	98,76	98,79	98,83	98,86	98,89	98,92	98,95	98,98	99,01	99,04
2,6	99,07	99,09	99,12	99,15	99,17	99,20	99,22	99,24	99,26	99,29
2,7	99,31	99,33	99,35	99,37	99,39	99,40	99,42	99,44	99,46	99,47
2,8	99,49	99,50	99,52	99,53	99,55	99,56	99,58	99,59	99,60	99,61
2,9	99,63	99,64	99,65	99,66	99,67	99,68	99,69	99,70	99,71	99,72
3,0	99,73	-	-	-	-	-	-	-	-	-
3,5	99,95	-	-	-	-	-	-	-	-	-
4,0	99,994	-	-	-	-	-	-	-	-	-
4,5	99,994	-	-	-	-	-	-	-	-	-
5,0	99,99994	-	-	-	-	-	-	-	-	-

Далее P умножаем на n (полное число измерений), чтобы получить

$$n\,(\text{хуже, чем } t_{\text{под}}) = n \cdot P\,(\text{вне } t_{\text{под}} \cdot S).$$

Полученное значение n – число ожидаемых измерений, которые дают столь же плохие результаты, как $x_{\text{под}}$. Если n меньше $1/2$, то $x_{\text{под}}$ не удовлетворяет критерию Шовене и отбрасываются.

После отбрасывания результата, не удовлетворяющего критерию Шовене, естественно, надо пересчитать $\bar{x}$ и S по оставшимся данным. При этом получится значение S, которое меньше первоначального, и может случиться так, что с новым значением S некоторые другие результаты измерений не будут удовлетворять критерию Шовене. Однако большинство специалистов считает, что критерий Шовене не должен применяться второй раз с использованием пересчитанных значений $\bar{x}$ и S.

Рассмотрим на примере данных табл. 3 применение критерия Шовене

$$t_{\text{под}1} = \frac{266{,}0 - 257{,}1}{2{,}6} = 3{,}42\,.$$

Из табл. 4 находим P (вне $t_{\text{под}} \cdot S$) равное 0,0016. Тогда

$$n\ (\text{хуже, чем } t_{\text{под}}) = 15 \cdot 0{,}0016 = 0{,}024\,.$$

Критерий Шовене утверждает, что если ожидаемое число измерений, столь же плохих, как и подозрительный результат, меньше $1/2$, то подозрительный результат следует исключить. В нашем случае это так.

Попробуем применить критерий Шовене к исправленному ряду измерений

$$t_{\text{под}2} = \frac{258{,}5 - 256{,}5}{2} = 1{,}0\,.$$

$$P\,(\text{вне } t_{\text{под}} \cdot S) = 0{,}317,\ n\,(\text{хуже, чем } t_{\text{под}}) = 14 \cdot 0{,}317 = 4{,}44\,.$$

Таким образом, значение $t_{под2}$ не является аномальным и не отбрасывается.

Рассмотрим теперь вариант 3 критерия оценки пригодности измерений, предложенный в справочном руководстве по метрологии [6].

Вычислим показатель анормальности

$$V_к = \left| \frac{x_к - \bar{x}}{S} \right|. \tag{5}$$

Затем, задавшись вероятностью γ, для данного объема выборки n из табл. 5 находят параметр β (здесь, как обычно, $\gamma = 0{,}95$).

Таблица 5 – Значения β для $\gamma = 0{,}95$ и различного числа измерений

n	β	n	β	n	β	n	β	n	β
3	1,15	9	2,11	15	2,41	21	2,58	27	2,68
4	1,46	10	2,18	16	2,44	22	2,59	28	2,69
5	1,67	11	2,23	17	2,48	23	2,61	29	2,69
6	1,82	12	2,29	18	2,50	24	2,63	30	2,70
7	1,94	13	2,33	19	2,53	25	2,65		
8	2,03	14	2,37	20	2,56	26	2,66		

Если результат $x_к$ принадлежит данной нормальной совокупности, то с вероятностью γ можно утверждать, что абсолютное значение показателя анормальности $V_к$ не превысит β. Следовательно, критерием анормальности является условие $V_к \geq \beta$. Если это условие соблюдается, вероятность данного результата наблюдения $x_к$ меньше $1-\gamma$. Следовательно, он анормален и должен быть исключен из данной выборки (после чего значения $\bar{x}, S$ и $V_к$ должны быть вычислены снова).

Рассмотрим теперь на примере данных табл. 5 применение последнего критерия:

$$V_{\kappa} = \left| \frac{266{,}0 - 257{,}1}{2{,}6} \right| = 3{,}42 \,.$$

По табл. 5 найдем $\beta = 2{,}41$.

Налицо выполнение условия анормальности этого измерения, и оно должно быть отброшено. Применение критерия анормальности к исправленной выборке дает:

$$V_{\kappa} = \frac{258{,}5 - 256{,}5}{2} = 1{,}0 \,.$$

Этому соответствует значение $\beta = 2{,}37$, так как $V_{\kappa 2} < \beta_2$, результат с вероятностью $\gamma = 0{,}95$ считается нормальным.

Примечание. Числовые данные примера 1, обработанные по схеме 1, 2 и 3 вариантов, показали аномальность того же измерения, что и Q - критерий.

Из рассмотренных здесь вопросов выявления аномальных результатов измерений видно, что при малом числе измерений наиболее удобно пользоваться Q - критерием. При достаточно большом числе измерений возможности обработки расширяются, и выбор конкретного варианта определяется стоящей задачей. Вариант 1 позволяет оперативно учитывать требуемую доверительную вероятность результата и применим для любого числа измерений, которое больше двух. Вариант 2 несколько более громоздок в применении и содержит условно введенное число измерений, равное $1/2$. Вариант 3 наиболее удобен в применении для доверительной вероятности 0,95, что, впрочем, чаще всего и требуется.

Литература

1. *Грановский В.А., Сирая Т.Н.* Методы обработки экспериментальных данных при измерениях. – Л.: Энергоатомиздат. Ленингр. отд-ние, 1989. – 288 с.
2. *Тойберг П.* Оценка точности результатов измерений / Пер. с нем. - М.: Энергоатомиздат, 1988. – 88 с.
3. *Романенко В.Н., Орлов А.Г., Никитина Г.В.* Книга для начинающего исследователя-химика. – Л.: Химия, 1987. – 280 с.
4. *Зайдель А.Н.* Ошибки измерений физических величин. – Л.: Наука, Ленинград. отд., 1974. – 108 с.
5. *Тейлор Дж.* Введение в теорию ошибок. – М: Мир, 1985. – 272 с.
6. *Рего К.Г.* Метрологическая обработка результатов технологических измерений: справ. Пособие. – К.: Техника, 1987. – 128 с.