Методы исключения результатов с грубыми погрешностями.

Вопросы:

1. Основные понятия и определения методов исключения результатов с грубыми погрешностями.

2. Критерии проверки результатов с грубыми погрешностями: Ирвина, Романовского, вариационного размаха, Диксона, Райта, Смирнова, Шовене.

1. Основные понятия и определения методов исключения результатов с грубыми погрешностями.

Грубые погрешности (промахи) относятся к числу погрешностей, изменяющимся случайным образом при повторных наблюдениях. Они явно превышают по своему значению погрешности, оправданные условиями проведения эксперимента. Под промахом понимается значение погрешности, отклонение которого от центра распределения существенно превышает значение, оправданное объективными условиями измерения. Поэтому с точки зрения теории вероятности появление промаха маловероятно.

Причинами грубых погрешностей могут быть неконтролируемые изменения условий измерений, неисправность, ошибки оператора и др.

Для исключения грубых погрешностей применяют аппарат проверки статистических гипотез.

В метрологии используются статистические гипотезы, под которыми понимают гипотезы о виде неизвестного распределения, или о параметрах известных распределений.

Примеры статистических гипотез:

1) рассматриваемая выборка (или ее отдельный результат) принадлежит генеральной совокупности;

2) генеральная совокупность распределена по нормальному закону;

3) дисперсии двух нормальных совокупностей равны между собой.

В двух гипотезах сделано предположение о виде неизвестного распределения и принадлежности отдельных (подозрительных) результатов данному виду распределения, а в третьей - о параметрах двух известных распределений. Наряду с выдвинутой гипотезой рассматривают и противоречащую ей гипотезу. Нулевой (основной) называют выдвинутую гипотезу. А конкурирующей (альтернативной) называют ту, которая противоречит нулевой.

При выдвижении и принятии гипотезы могут иметь место следующие четыре случая:

1) гипотеза принимается, причем и в действительности она правильная;

2) гипотеза верна, но ошибочно отвергается. Возникающую при этом ошибку называют **ошибкой первого рода**, а вероятность ее появления называют **уровнем значимости** и обозначают $q(\alpha)$;

3) гипотеза отвергается, причем в действительности она неверна;

4) гипотеза неверна, но ошибочно принимается. Возникающую при этом ошибку называют **ошибкой второго рода**, а вероятность ее появления обозначают β .

Величину $1-\beta$, т. е. вероятность, что гипотеза будет отвергнута, когда она ошибочна, называют **мощностью критерия**.

Следует заметить, что в нормативной документации по статистическому контролю качества продукции и учебниках по управлению качеством вероятность признать негодной партию годных изделий (т. е., совершить ошибку первого рода) называют “риском производителя”, а вероятность принять негодную партию – “риском потребителя”.

Все статистические критерии являются случайными величинами, принимающими определенные значения (таблицы критических значений).

Областью принятия гипотезы (областью допустимых значений) называют совокупность значения критерия, при которых гипотезу принимают. Критической называют совокупность значений критерия, при которых нулевую гипотезу отвергают. Область принятия гипотезы и критическая область разделены критическими точками, в качестве которых и выступают табличные значения критериев.

Область непринятия гипотезы может быть односторонней (правосторонней или левосторонней) и двух сторонней.

Правосторонней называют критическую область, определяемую неравенством $K_{набл} > k_{кр}$, где $k_{кр}$ – положительное число.

- Левосторонней называют критическую область, определяемую неравенством $K_{набл} < k_{кр}$, где $k_{кр}$ – отрицательное число.

Двусторонней называют критическую область, определяемую неравенствами $K_{набл} > k_1$; $K_{набл} < k_2$ где $k_2 > k_1$. Если критичсекие точки симметричны относительно нуля, двусторонняя критическая область определяется неравенствами: $K_{набл} < - k_{кр}$, $K_{набл} > k_{кр}$, или равносильным неравенством $|K_{\text{набл}}| > k_{кр}$.

Основной принцип проверки статистических гипотез формулируется следующим образом: если наблюдаемое (опытное) значение критерия принадлежит критической области – гипотезу отвергают, если наблюдаемое значение критерия принадлежит области принятия гипотезы – гипотезу принимают.

Проверку статистической гипотезы проводят для принятого уровня значимости q(принимается равным 0,1; 0,05; 0,01 и т. д.). Так принятый уровень значимости q=0,05 означает, что выдвинутая нулевая статистическая гипотеза может быть принята с доверительной вероятностью P = 0,95. Или есть вероятность отвергнуть эту гипотезу (совершить ошибку первого рода), равная P =0,95.

- Нулевая статистическая гипотеза подтверждает принадлежность проверяемого “подозрительного” результата измерения (наблюдения) данной группе измерений.

Формальным критерием аномальности результата наблюдений (а, следовательно, и основанием для принятия конкурирующей гипотезы: “подозрительный” результат не принадлежит данной группе измерений) при этом служит граница, отнесенная от центра распределения на величину tS , т. е

$$\left|x_{i\text{под}} - \overline{X_{\text{ц.р.}}}\right| \geq tS,$$

где $x_{i\text{под}}$ – результат наблюдения, проверяемый на наличие грубой погрешности;

t – коэффициент, зависящий от вида и закона распределения, объема выборки, уровня значимости.

Таким образом, границы погрешности зависят от вида распределения, объема выборки и выбранной доверительной вероятности.

- При обработке уже имеющихся результатов наблюдений произвольно отбрасывать отдельные результаты не следует, так как это может привести к фиктивному повышению точности результата измерений. Группа измерений (выборка) может содержать несколько грубых погрешностей и их исключение производят последовательно, по одному.

Все методы исключения грубых погрешностей (промахов) могут быть разделены на два основных типа:

а) методы исключения при известном генеральном СКО;

б) методы исключения при неизвестном генеральном СКО.

В первом случае $X_{\text{ц.р.}}$ и СКО вычисляется по результатам всей выборки, во втором случае из выборки перед вычислением удаляются подозрительные результаты.

В случае ограниченного числа наблюдений и (или) сложности оценки параметров закона распределения рекомендуется исключать грубые погрешности, используя приближенные коэффициенты вида распределения. При этом исключаются значения $x_i < x_{r-}$ и $x_i > x_{r+}$, где x_{r-} , x_{r+} – границы промахов, определяемые выражениями:

- $$x_{r-} = X - S \cdot \left[1 + A \cdot \sqrt{\frac{1}{\gamma^2} - 1}\right];$$

$$x_{r+} = X - S \cdot \left[1 + A \cdot \sqrt{\frac{1}{\gamma^2} - 1}\right].$$

где A – коэффициент, значение которого выбирается в зависимости от заданной доверительной вероятности в диапазоне от 0,85 до 1,30;

γ – контрэксцесс, значение которого зависит от формы закона распределения величины (ЗРВ).

После исключения промахов операции по определению оценок центра распределения и СКО результатов наблюдений и измерений необходимо повторить.

Поскольку на практике чаще встречаются измерения при неизвестном СКО (ограниченное число наблюдений), то рассмотрены следующие критерии проверки подозрительных (с точки зрения погрешностей) результатов наблюдений: Ирвина, Романовского, вариационного размаха, Диксона, Смирнова, Шовене.

Поскольку критериальные требования (коэффициенты), определяющие границу, за которой находятся “грубые” (в смысле погрешностей) результаты наблюдений у разных авторов различны, то проверку следует выполнять сразу по нескольким критериям (рекомендуется использовать не меньше трех, из рассматриваемых ниже). Окончательное заключение о принадлежности “подозрительных” результатов рассматриваемой совокупности наблюдений следует делать по большинству критериев.

2. Критерии проверки результатов с грубыми погрешностями: Ирвина, Романовского, вариационного размаха, Диксона, Райта, Смирнова, Шовене.

- **2.1 Критерий Ирвина.**

Для полученных экспериментальных данных определяют коэффициент по формуле:

$$\lambda = \frac{(x_{n+1} - x_n)}{S}$$

где x_{n+1}, x_n – наибольшие значения случайной величины;

S – среднее квадратическое отклонение, вычисленное по всем значениям выборки.

Затем этот коэффициент сравнивается с табличным значением λ_q, возможные значения которого приведены в таблице 1.

- Таблица 3.1 – Критерий Ирвина λ_q .

Число измерений n	Уровень значимости	
	q=0,05	q=0,01
1	2	3
2	2,8	3,7
3	2,2	2,9
10	1,5	2,0
20	1,3	1,8
30	1,2	1,7
50	1,1	1,6
100	1,0	1,5
400	0,9	1,3
1000	0,8	1,2

Если $\lambda > \lambda_q$, то нулевая гипотеза не подтверждается, т. е. результат - ошибочный, и он должен быть исключен при дальнейшей обработке результатов наблюдений.

• **2.2 Критерий Романовского.**

Конкурирующая гипотеза о наличии грубых погрешностей в подозрительных результатах подтверждается, если выполняется неравенство:

$$\left|x_{i\text{под}} - \overline{X_{\text{ц.р.}}}\right| \geq t_p S$$

где t_p – квантиль распределения Стьюдента при заданной доверительной вероятности с числом степеней свободы $k = n - k_n$ (k_n - число подозрительных результатов наблюдений). Фрагмент квантилей для распределения Стьюдента представлен в таблице 2.

Точечные оценки распределения $X_{\text{ц.р.}}$ и СКО S результатов наблюдений вычисляется без учета k_n подозрительных результатов наблюдений.

Таблица 3.2 – Критерий Стьюдента t_p (квантили Стьюдента).

Доверительная вероятность p	Число степеней свободы k												
	3	4	5	6	8	10	12	18	22	30	40	60	120	
0,90	2,35	2,13	2,01	1,94	1,86	1,81	1,78	1,73	1,72	1,70	1,68	1,67	1,66	1,64
0,95	3,18	2,78	2,57	2,45	2,31	2,23	2,18	2,10	2,07	2,04	2,02	2,00	1,98	1,96
0,99	5,84	4,60	4,03	3,71	3,36	3,17	3,06	2,98	2,82	2,75	2,70	2,86	2,62	2,58

2.3 Критерий вариационного размаха.

Является одним из простых методов исключения грубой погрешности измерений (промаха). Для его использования определяют размах вариационного ряда упорядоченной совокупности наблюдений ($x_1 \le x_2 \le \ldots \le x_k \le \ldots \le x_n$) :

$$R_n = x_n - x_1.$$

Если какой-либо член вариационного ряда, например x_k, резко отличается от всех других, то производят проверку, используя следующее неравенство:

$$\bar{X} - z \cdot R_n < x_k < \bar{X} + z \cdot R_n,$$

где $\bar{X}$ – выборочное среднее арифметическое значение, вычисленное после исключения предполагаемого промаха;

z – критериальное значение.

Нулевую гипотезу (об отсутствии грубой погрешности) принимают, если указанное неравенство выполняется. Если x_k не удовлетворяет этому условию, то этот результат исключают из вариационного ряда.

Коэффициент z зависит от числа членов вариационного ряда n, что представлено в таблице 3.

Таблица 3.3 – Критерий вариационного размаха.

n	5	6	7	8-9	10-11	12-15	16-22	23-25	26-63	64-150
z	1,7	1,6	1,5	1,4	1,3	1,2	1,1	1,0	0,9	0,8

2.4 Критерий Диксона

Критерий основан на предположении, что погрешности измерений подчиняются нормальному закону и проверка гипотезы о принадлежности нормальному закону распределения. При использовании критерия вычисляют коэффициент Диксона (наблюдаемое значение критерия) для проверки наибольшего или наименьшего экстремального значения в зависимости от числа измерений. В таблице 4 приведены формулы для вычисления коэффициентов. Коэффициенты r_{10}, r_{11} применяют, когда имеется один выброс, а r_{21} и r_{22} - когда два выброса. Требуется первоначальное упорядочение результатов измерений (объема выборки). Критерий применяется, когда выборка может содержать более одной грубой погрешности.

Таблица 3.4 – Формулы коэффициентов Диксона.

Число измерений n (обьем выборки)	Коэффициент Диксона	Для наименьшего экстремального параметра	Для наибольшего экспериментально го параметра
1	2	3	4
3-7	r_{10}		
8-10	r_{11}		
11-13	r_{21}		
14-25	r_{22}		

Вычисленные для выборки по формулам значения коэффициентов Диксона r сравнивают с принятым (табличным) значением критерия Диксона r_q (таблица 5).

Нулевая гипотеза об отсутствии грубой погрешности выполняется, если выполняется неравенство $r < r_q$.

Если $r > r_q$, то результат признается грубой погрешностью и исключается из дальнейшей обработки.

Таблица 5 – Критериальные значения коэффициентов Диксона (при принятом уровне значимости *q*).

Статистика	Число измерений	r_q при уровне значимости q			
		0,1	0,05	0,02	0,01
1	2	3	4	5	6
r_{10}	3	0,886	0,941	0,976	0,988
	4	0,679	0,765	0,846	0,899
	5	0,557	0,642	0,729	0,780
	6	0,482	0,560	0,644	0,698
	7	0,434	0,507	0,586	0,637
r_{11}	8	0,479	0,554	0,631	0,683
	9	0,441	0,512	0,587	0,636
	10	0,409	0,477	0,551	0,597
r_{21}	11	0,517	0,576	0,538	0,679
	12	0,490	0,546	0,605	0,642
	13	0,467	0,521	0,578	0,615
r_{22}	14	0,462	0,546	0,602	0,641
	15	0,472	0,525	0,579	0,616
	16	0,452	0,507	0,559	0,595
	17	0,438	0,490	0,542	0,577
	18	0,424	0,475	0,527	0,561
	19	0,412	0,462	0,514	0,547
	20	0,401	0,450	0,502	0,535
	21	0,391	0,440	0,491	0,524
	22	0,382	0,430	0,481	0,514
	23	0,374	0,421	0,472	0,505
	24	0,367	0,413	0,464	0,497
	25	0,360	0,406	0,457	0,489

2.5 Критерии "3 σ " , Райта.

Критерий “правило трех сигм” является одним из простейших для проверки результатов, подчиняющихся нормальному закону распределения. Сущность правила трех сигм: если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.

На практике правило трех сигм применяют так: если распределение изучаемой случайной величины неизвестно, но условие, указанное в приведенном правиле, выполняется, то есть основания предполагать, что изучаемая величина распределена нормально; в противном случае она не распределена нормально. С этой целью для выборки (включая подозрительный результат) вычисляется центр распределения и оценка СКО результата наблюдений. Результат, который удовлетворяет условию

$\left|x_{i\text{под}} - \overline{X_{\text{ц.р.}}}\right| \geq 3S$, считается имеющим грубую погрешность и удаляется, а ранее вычисленные характеристики распределения уточняются.

Этому критерию аналогичен критерий Райта, основанный на том, что если остаточная погрешность больше четырех сигм, то этот результат измерения является грубой погрешностью и должен быть исключен при дальнейшей обработке. Оба критерия надежны при числе измерений больше 20…50. Их правомочно применять, когда известна величина генерального среднеквадратического отклонения (S) .

- Может оказаться, что при новых значениях $\overline{X_{\text{ц.р.}}}$ и S другие результаты попадут в категорию аномальных. Однако дважды использовать критерии грубой погрешности не рекомендуется.

2.6 Критерий Смирнова.

Критерий Смирнова используется при объемах выборки $n \geq 25$ или при известных значениях генеральных среднего и СКО. Он устанавливает менее жесткие границы грубой погрешности. Для реализации этого критерия вычисляются действительные значения квантилей распределения (наблюдаемое значение критерия) по формуле:

$$\beta = \frac{max\left|x_{i\text{под}} - \overline{X}\right|}{\overline{S}}$$

Найденное значение сравнивается с критериальным β_k, приведенным в таблице 6.

- Таблица 6 – Квантили распределения β_k.

Объем Выборки n					
	0,100	0,050	0,0010	0,005	0,001
1	2	3	4	5	6
1	1,282	1,645	2,326	2,576	3,090
2	1,632	1,955	2,575	2,807	3,290
3	1,818	2,121	2,712	2,935	3,403
4	1,943	2,234	2,806	3,023	3,480
5	2,036	2,319	2,877	3,090	3,540
6	2,111	2,386	2,934	3,143	3,588
7	2,172	2,442	2,981	3,188	3,628
8	2,224	2,490	3,022	3,227	3,622
9	2,269	2,531	3,057	3,260	3.692
10	2,309	2,568	3,089	3,290	3,719
15	2,457	2,705	3,207	3,402	3,820
20	2,559	2,799	3,289	3,480	3,890
25	2,635	2,870	3,351	3,539	3,944
30	2,696	2,928	3,402	3,587	3,988
40	2,792	3,015	3,480	3.662	4,054
50	2,860	3,082	3,541	3,716	4,108
100	3,076	3,285	3,723	3,892	4,263
250	3,339	3,534	3,946	4,108	4,465
500	3,528	3,703	4,108	4,263	4,607

2.7 Критерий Шовене.

Критерий Шовене применяется для законов, не противоречащих нормальному, и строится на определении числа ожидаемых результатов наблюдений $n_{ож}$, которые имеют столь же большие погрешности, как и подозрительный. Гипотеза о наличии грубой погрешности принимается, если выполняется условие:

$$n_{ож} \leq 0{,}5 .$$

Порядок проверки гипотезы следующий:

1) вычисляются среднее арифметическое $\overline{X}$ и СКО S результатов наблюдений для всей выборки;

2) из таблицы нормированного нормального распределения (Справочное пособие – интегральная функция нормированного нормального распределения) по величине $z = \frac{|x_{i\text{под}} - \overline{X_{\text{ц.р.}}}|}{S}$ определяется вероятность появления подозрительного результата в генеральной совокупности чисел n:

$$P\left(z \cdot S < \left|x_{i\text{под}} - \overline{X_{\text{ц.р.}}}\right|\right);$$

- 3) число ожидаемых результатов $n_{ож}$, определяется по формуле:

$$n_{ож} = n \cdot P.$$

Указанные выше критерии во многих случаях оказываются “жесткими”. Тогда рекомендуется пользоваться критерием грубой погрешности " *k*", зависящим от объема выборки *n* и принятой доверительной вероятности *P*.

Таблица 7 – Зависимость критерия грубой погрешности *k* от объема выборки *n* и доверительной вероятности *P.*

n	P=95,00	P=99,00	P=99,73	N	P=95,00	P=99,00	P=99,73
9	4,42	7,10	11,49	25	3,84	5,14	6,25
10	4,31	6,99	10,26	30	3,80	5,00	5,95
12	4,16	6,38	8,80	40	3,75	4,82	5,56
15	4,03	5,88	7,66	50	3,73	4,70	5,34
20	3,90	5,41	6,73				

- Для распределений, отличных от нормального, таких классов, как двух модальных кругловершинных композиций нормального и дискретного распределения с эксцессом ε = 1,5-3,0; островершинных двумодальных; композиций дискретного двузначного распределения и распределения Лапласа с эксцессом ε = 1,5-6,0; композиций равномерного распределения с экспоненциальным распределением эксцесса ε = 1,8-6,0 и классом экспоненциальных распределений в пределах изменения эксцесса ε = 1,8-6,0 граница грубой погрешности определяется величиной ± $(t_{\text{гр}} \cdot \sigma)$ или ± $(t_{\text{гр}} \cdot S)$, где:

$$t_{\text{гр}} = 1{,}2 + 3{,}6\,(1-\gamma) \cdot lg\frac{n}{10}$$

где γ – контрэксцесс;

$$t_{\text{гр}} = 1{,}55 + 0{,}8 \cdot \sqrt{\varepsilon - 1} \cdot lg\frac{n}{10}$$

- Погрешности в определении оценок S СКО и $t_{\text{гр}}$ являются отрицательно коррелированными, т. е. возрастание СКО S сопровождается уменьшением $t_{\text{гр}}$. Поэтому определение границ грубой погрешности для законов, отличных от нормального, с эксцессом ε ≤ 6 с помощью критерия $t_{\text{гр}}$ является достаточно точным и может широко использоваться на практике.

Оценки $\bar{X}$, S и ε должны вычисляться после исключения подозрительных результатов из выборки. После расчета границ грубой погрешности результаты наблюдений, оказавшиеся внутри границ, возвращаются, а ранее найденные характеристики распределения уточняются.

Для равномерного распределения за границы грубой погрешности можно принять величину $\pm\, 1{,}8 \cdot S$.