



ЛЕКЦИИ ПО ПРАКТИЧЕСКОЙ МЕТРОЛОГИИ ЛАБОРАТОРНОЙ МЕДИЦИНЫ

Лекция 3. Технологии определения аналитических характеристик аналитических систем

М.И. Прищепа, кандидат технических наук, ЗАО «АНАЛИТИКА»

В настоящей лекции будут описаны процедуры общепринятых технологий определения на практике эксплуатационных значений основных аналитических характеристик аналитических систем, а также рассмотрены вопросы точности определения этих значений.

Основными аналитическими характеристиками аналитических систем (далее – АС), как было сказано в Лекции 2, являются систематическое смещение Δ и стандартное отклонение σ , которые также называют систематической и случайной составляющей погрешности АС. Эти две составляющие при заданном уровне доверия сполна определяют максимальную величину полной погрешности $TE_{ас}$ результатов измерений содержания аналита в биопробах, получаемых с использованием данной АС.

Очень важно помнить, что один и тот же экземпляр АС, эксплуатируемый в разных внешних условиях температуры, влажности, чистоты воздуха, электропитания, электромагнитной близости другого оборудования и т.д. и т.п., будет обладать скорее всего разными эксплуатационными значениями аналитических характеристик. Иногда даже перемещение АС из одного помещения в другое значительно изменяет эти значения.

Итак, примем к сведению, что каждая АС обычно обладает уникальными значениями аналитических характеристик, характерными для конкретных условий ее эксплуатации и диапазона измерений. Теоретически оценивать эти характеристики можно **двумя методами**.

При использовании **первого метода** исходят из возможности оценивать значения аналитической характеристики АС путем суммирования определенным образом индивидуальных значений этой характеристики каждого из измерительных преобразователей (далее – ИП), входящих в состав этой АС. Иными словами, путем учета вноса аналитических характеристик отдельных ИП в составе АС в итоговое значение той или иной составляющей ее погрешности. Понятно, что для этого надо потребовать от изготовителя каждого ИП предоставлять в документации соответствующую информацию о его аналитических характеристиках при тех или иных условиях его эксплуатации. Хотя понятно, что даже если такая информация и будет представлена в документации, нет никаких гарантий, что она будет верна для конкретных условий эксплуатации ИП. Да и оценивать по ней итоговое значение каждой составляющей погрешности будет все равно достаточно не просто для сотрудников лабораторий. Заметим, что этот метод – по крайней мере в части определения систематического смещения АС – является вариантом практической реализации рекомендаций соблюдения лабораториями условий прослеживаемости измерений [1], упомянутых в начале Лекции 1, требующих оценки итоговой величины неопределенности результатов при передаче размера единицы физической величины от ее эталона к калибруемому средству измерений.

В связи со сложностью практической реализации первого метода для достижения той же цели обычно лаборатории используют так называемый **метод «чёрного ящика»** и определяют аналитические характеристики АС сразу же как цельного средства измерения, то есть без детализации на ИП, входящих в состав АС. Термин «чёрный ящик» используют для обозначения системы, внутреннее устройство и

механизм работы которой или очень сложны, или неизвестны или просто неважны в рамках решаемой задачи. Метод «чёрного ящика» - метод исследования таких систем, когда вместо свойств и взаимосвязей составных частей системы изучается ее реакция как целого на тот или иной вид воздействия, на нее направленный.

В теории измерений обычно метод «черного ящика» используют для описания работы преобразователя сигналов, поступающих на его вход, когда нет необходимости знать о его содержимом в деталях. Аналитические характеристики таких черных ящиков определяют путем сравнения формы и структуры измерительных сигналов, входящих в него и выходящих из него. Любая АС для лабораторий это и есть, по сути, тот самый «черный ящик». На вход такого «черного ящика», то есть АС, подают образец пробы пациента или образец контрольного материала (далее - КМ) или образец калибратора как есть или преобразованный, а на выходе этого «черного ящика» имеют измерительную информацию, характеризующую уровень аналита в исследуемом образце, поданном на его вход. Чтобы узнать, насколько такая измерительная информация может быть далека от фактического уровня аналита в исследуемом образце, совсем не обязательно разбираться в конструкции и механизме работы «черного ящика» и выяснять насколько аналитически точно он на каждом отдельном шаге измерительного процесса преобразовывает одну физическую величину в другую, пытаюсь таким образом обеспечивать прослеживаемость результатов проводимых измерений. Чтобы определить аналитические характеристики «черного ящика», то есть АС в целом, вполне достаточно подавать на вход «черного ящика» в течение определенного времени серии аликвот эталона состава или свойств биопробы с заранее известным содержанием искомого аналита. А затем сравнить полученные на выходе «черного ящика» результаты измерения между собой и с заранее известным уровнем аналита в образце на его входе. Конечно и тут не все так просто.

Если для каких-либо аналитов нет таких эталонов, а их пока нет для очень многих аналитов, то тогда вместо эталонов можно использовать или эрзац-эталон в виде контрольных материалов (далее – КМ) промышленного изготовления, близких по матрице к пробам пациентов, или неаттестованные, но застabilизированные по составу пробы пациентов. В последнем случае, по понятным причинам, требуется наличие также референтных методов измерения уровней искомого аналита, чтобы иметь возможность аттестовать содержание аналита в используемой биопробе. На практике применяют и другие технологии замены эталонов состава и свойств, позволяющие лаборатории оценивать правильность определения содержания аналита в образцах КМ или в пробах пациентов. Эти технологии основаны либо на участии лаборатории с исследуемой АС в программе ВОК, либо на ведении для нее контроля качества по ежедневным средним по пациентам.

Далее в этой лекции мы будем рассматривать использование только второго метода для определения истинных эксплуатационных значений аналитических характеристик АС, то есть как цельного средства измерений. При этом будем иметь в виду, что истинные эксплуатационные значения этих характеристик можно определить только теоретически путём проведения очень большого количества повторных измерений уровня аналита в аликвотах одного и того же эталона или одной и той же биопробы, что по понятным причинам реализовать достаточно сложно, вернее сказать, что это по разным объективным причинам в принципе на практике невозможно.

На практике делают только так называемые выборочные оценки этих истинных значений, проводя ограниченное, но статистически значимое количество повторных измерений уровня аналита в

аликвотах одного и того же биоматериала, для чего обычно используют контрольный материал промышленного изготовления с аттестованным содержанием искомого аналита.

У выборочных оценок с ограниченным объемом измерений есть как плюсы, так и минусы по сравнению с чрезмерно большой выборкой. Основной плюс заключается в том, что для проведения первых требуется гораздо меньше ресурсов, чем для проведения вторых. Основной минус состоит в том, что выборочные оценки всегда имеют какие-то отклонения от истинных значений. Более того, повторные выборочные оценки практически всегда отличаются друг от друга. Поэтому основная задача проведения выборочных оценок заключается в получении нужной для них точности при минимизации временных и материальных затратах на их проведение.

Правила получения выборочных оценок истинных значений систематического смещения и стандартного отклонения АС в условиях ее эксплуатации, которые будем обозначать соответственно V_n и S_n для выборок объемом n , стандартизованы и описаны в приказе Минздрава РФ № 220 [2] и в ГОСТ Р 53133.2-2008 [3]. Помимо отечественных руководящих документов рекомендации по проведению оценки эксплуатационных значений аналитических характеристик АС изложены в зарубежных стандартах, например, в руководстве EP15-A3 [4] Clinical and Laboratory Standards Institute, который является одним из ведущих международных организаций в области стандартизации. Хотя процедуры проведения такой оценки, описанные в этих документах, несколько и отличаются друг от друга, но не принципиально.

Отечественные и международные правила исходят из того, что необходимо определять основные аналитические характеристики для всех АС, которые планируется вводить в эксплуатацию, в том числе и для АС, основные части которых подвергались или ремонту, или замене. А затем сравнить их с установленными нормами, чтобы принять решение о запуске новой или обновленной АС в эксплуатацию.

Аналитические характеристики АС в пределах желаемого диапазона измерений определяют путем статистической обработки результатов повторных (или что тоже самое – многократных) измерений содержания аналита в аликвотах нескольких биоматериалов с разными уровнями искомого аналита, более или менее равномерно расположенными внутри этого диапазона. Последнее связано с тем, что эксплуатационные значения аналитических характеристик, как правило, зависят не только от условий эксплуатации АС, но и от уровня аналита в биопробе. Из теории измерений известно, что генеральная, то есть бесконечно большая совокупность результатов повторных измерений на одной и той же АС уровня аналита в аликвотах одного и того же биоматериала будут распределены по нормальному закону со своими конкретными значениями параметров распределения - математического ожидания μ и стандартного отклонения σ . Практика же показывает, что достаточно большая совокупность или, как говорят, представительная выборка повторных измерений может при определенных условиях достаточно хорошо описывать их распределение и в генеральной совокупности, позволяя только с помощью выборочных результатов достаточно точно оценивать значения показателей распределения μ и σ . Одно из этих условий состоит в отсутствии среди результатов в выборке так называемых выбросов. Для накопления нужного количества n результатов повторных измерений, или, как говорят, представительной выборки объемом n , адекватно характеризующей их распределение в генеральной совокупности, что необходимо для достаточно надежной оценки истинных значений систематического смещения и стандартного отклонения, на АС проводят так называемые установочные серии. Чтобы получить представление о поведении аналитических характеристик в пределах диапазона измерений АС, рекомендуют проводить установочные серии с использованием хотя бы двух биоматериалов с двумя

разными уровнями аналита, характерных для принятия клинических решений. Поэтому при проведении установочных серий необходимо использовать минимум два биоматериала с уровнями исследуемого аналита, характерными для нормы и патологии.

Из положений математической статистики следует, что для того, чтобы выборка результатов повторных измерений объемом n была представительной, а полученные выборочные значения аналитических характеристик являлись надежными оценками их истинных значений, число n должно быть не менее 20. Поэтому в отечественных правилах [2, 3] число n выбрано равным 20, а в документе CLSI - равным 25, вернее не менее, чем 25. Отличие процедур, описываемых в этих документах, состоит также в том, что в соответствии с отечественными правилами, установочную серию проводят в течение 20 рабочих дней, проводя по одному измерению ежедневно в каждом из двух типов КМ, а по документу CLSI установочную серию рекомендовано проводить или в течение 5 рабочих дней по 5 повторных измерений в одной аналитической серии ежедневно или в течение 7 рабочих дней по 4 повторных измерения в одной аналитической серии ежедневно.

Необходимость накопления экспериментальных данных для проведения оценки эксплуатационных значений аналитических характеристик в течение такого длительного срока связана с тем, что воспроизводимость результатов повторных измерений в течение кратких сроков эксплуатации АС всегда значительно лучше по сравнению с воспроизводимостью в течение длительных сроков ее эксплуатации. Краткосрочная воспроизводимость, то есть воспроизводимость в пределах одной аналитической серии или в пределах рабочего дня называется повторяемостью, а долгосрочная воспроизводимость, то есть воспроизводимость изо дня в день – прецизионностью. Очевидно, что данные, полученные в установочной серии в течение большего количества рабочих дней, будут более адекватно характеризовать воспроизводимость результатов изо дня в день, то есть прецизионность АС. Следует отметить, что **для решения клинических задач именно этот показатель является наиболее важным.**

Используя все n результатов повторных измерений, полученных в установочной серии, вычисляют по нижеприведенным формулам их среднеарифметическое значение $(X_{cp})_n$, которое является, как говорят, выборочной оценкой истинного или, что тоже самое, генерального значения математического ожидания μ , и стандартное отклонение S_n , которое является, как говорят, выборочной оценкой истинного или, что тоже самое, генерального значения σ .

$$(X_{cp})_n = \frac{\sum_{i=1}^n X_i}{n} = \frac{(X_1 + X_2 + X_3 + \dots + X_n)}{n} \quad (3.1)$$

$$S_n = \sqrt{\frac{\sum_{i=1}^n [X_i - (X_{cp})_n]^2}{n - 1}} \quad (3.2)$$

Выборочные оценки генерального значения σ , вычисленные по формуле (3.2), являются несмещенными оценками стандартного отклонения. Следует иметь в виду, что оценки этого статистического показателя, рассчитываемые по формуле

$$S_n = \sqrt{\frac{\sum_{i=1}^n [X_i - (X_{cp})_n]^2}{n}} \quad (3.2a)$$

вытекающей из определения дисперсии, при малых объемах выборки являются смещенными и всегда менее точно характеризуют его генеральное значение. Поэтому при малых объемах выборки, когда $n < 100$, этот показатель желательно оценивать так, чтобы его оценка была несмещенной, а именно по формуле (3.2). Обычно для расчетов несмещенных оценочных значений S_n используют формулу, аналогичную формуле (3.2), но преобразованную к виду, более удобному для проведения вычислений, а именно:

$$S_n = \sqrt{\frac{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}{n - 1}} \quad (3.3)$$

Как видно из вышеприведенных формул для вычисления значений S_n , они не могут быть отрицательными.

Таблица 3.1. Вычисление значений $(X_{cp})_n$ (формула 3.1) и S_n (формула 3.2) по результатам измерений уровня натрия в КМ с Акм = 144 ммоль/л.

Номер измерения	Результаты X_i , ммоль/л	$[X_i - (X_{cp})_n]$	$[X_i - (X_{cp})_n]^2$
1	144	-1	1
2	145	0	0
3	143	-2	4
4	145	0	0
5	145	0	0
6	147	2	4
7	145	0	0
8	146	1	1
9	145	0	0
10	145	0	0
11	145	0	0
12	146	1	1
13	145	0	0
14	146	1	1
15	143	-2	4
16	145	0	0
17	150	5	25
18	143	-2	4
19	144	-1	1
20	143	-2	4
$\sum_{i=1}^n X_i = 2900$		$\sum_{i=1}^n [X_i - (X_{cp})_n]^2 = 50$	
$(X_{cp})_n = 145$ ммоль/л		$S_n = \sqrt{50/19} = \sqrt{2,63} = 1,62$ ммоль/л	

где X_i – i -тый результат ($i=1,2,3,\dots,n$) измерения; $\sum_{i=1}^{i=n} X_i$ – сумма всех n результатов, начиная с X_1 и кончая X_n ; $(X_{ср})_n = \sum_{i=1}^{i=n} X_i/n = (X_1 + X_2 + X_3 + \dots + X_n)/n$ – среднеарифметическое значение n результатов; $\sum_{i=1}^{i=n} [X_i - (X_{ср})_n]^2$ – сумма квадратов отклонений каждого из n результатов от их среднеарифметического значения; $\sum_{i=1}^{i=n} X_i^2$ – сумма квадратов результатов; $(\sum_{i=1}^{i=n} X_i)^2$ – квадрат суммы всех результатов измерений. Примеры вычислений $(X_{ср})_n$ и S_n по вышеприведенным формулам представлены соответственно в Таблицах 3.1 и 3.2. Отличие вычислений в этих примерах состоит только в том, что в первом примере расчет значения S_n проводился по формуле (3.2), а во втором примере – по формуле (3.3). Как и следовало ожидать, значения S_n , вычисленные по формулам (3.2) и (3.3) совпадают.

Таблица 3.2. Вычисление значений $(X_{ср})_n$ (формула 3.1) и S_n (формула 3.3) по результатам измерений уровня натрия в КМ с Акм = 144 ммоль/л.

Номер измерения	Результаты X_i , ммоль/л	Сумма квадратов результатов $(X_i)^2$
1	144	20736
2	145	21025
3	143	20449
4	145	21025
5	145	21025
6	147	21609
7	145	21025
8	146	21316
9	145	21025
10	145	21025
11	145	21025
12	146	21316
13	145	21025
14	146	21316
15	143	20449
16	145	21025
17	150	22500
18	143	20449
19	144	20736
20	143	20449
$\sum_{i=1}^{i=n} X_i = 2900$	$\sum_{i=1}^{i=n} X_i^2 = 420550$	$(\sum_{i=1}^{i=n} X_i)^2 = 8410000$
$(X_{ср})_n = 145$ ммоль/л		
$S_n = \sqrt{(420550 - 8410000/20)/19} = \sqrt{50/19} = \sqrt{2,63} = 1,62$ ммоль/л		

В руководстве [4] рекомендуется после вычисления значений $(X_{ср})_n$ и S_n проверять все результаты установочной серии на наличие среди них результатов, сильно отклонившихся от выборочного среднего, то есть результатов, вышедших за границы определенных доверительных интервалов, близких к $(X_{ср})_n \pm 3 \cdot S_n$ и зависящих от выбранного уровня доверия. Такие результаты обычно принято называть грубыми ошибками или выбросами. При их наличии среди результатов,

полученных в установочной серии, они изымаются из нее, а полученные ранее значения $(X_{cp})_n$ и S_n рассчитывают заново, то есть без учета выбросов.

Из опыта известно, что результаты повторных измерений уровня аналита в одном и том же биоматериале, полученные с использованием АС с неизменными аналитическими характеристиками, всегда распределены по нормальному закону. При нормальном распределении результатов повторных измерений для выявления грубых ошибок (выбросов) обычно принято использовать статистический критерий Граббса [5].

В соответствии с положениями этого критерия сначала определяют среднее значение $(X_{cp})_n$ и стандартное отклонение S_n , используя все результаты, полученные в установочной серии, затем выявляют среди них наименьший X_{min} и наибольший X_{max} результаты, предполагая, что они могут являться выбросами, затем, используя эти данные, рассчитывают критерий Граббса G_1 и G_2 соответственно для результатов X_{min} и X_{max} по формулам:

$$G_1 = \frac{(X_{cp})_n - X_{min}}{S_n} \quad (3.4)$$

$$G_2 = \frac{X_{max} - (X_{cp})_n}{S_n} \quad (3.5)$$

И, наконец, полученные расчетные значения G_1 и G_2 сравнивают с критическим значением G_T , приведённым в Таблице 3.3, которое соответствует количеству n результатов, полученных в установочной серии, и выбранному уровню доверия (уровню значимости). Если $G_1 \leq G_T$, то результат X_{min} не считают грубой ошибкой и его сохраняют среди результатов установочной серии. Если $G_2 \leq G_T$, то результат X_{max} не считают грубой ошибкой и его сохраняют среди результатов установочной серии. Если $G_1 > G_T$ и/или $G_2 > G_T$, то тогда X_{min} и/или X_{max} считают грубыми ошибками и их исключают из результатов установочной серии как маловероятные значения.

Таблица 3.3. Критические значения G_T для критерия Граббса.

Число результатов n	Критические значения G_T при уровне значимости:	
	$\alpha=0,01$	$\alpha=0,05$
5	1,764	1,715
10	2,482	2,290
15	2,806	2,549
17	2,894	2,620
18	2,932	2,651
19	2,968	2,681
20	3,001	2,709
21	3,031	2,733
22	3,060	2,758
23	3,087	2,781
24	3,112	2,802
25	3,135	2,841
27	3,178	2,859
30	3,236	2,908
40	3,381	3,036

Таблица 3.4. Перерасчет значений $(X_{ср})_n$ (формула 3.1) и S_n (формула 3.2) после выявления по Граббсу грубой ошибки среди результатов, приведенных в Таблице 3.1.

Номер измерения	Результаты X_i , ммоль/л	$[X_i - (X_{ср})_n]$	$[X_i - (X_{ср})_n]^2$
1	144	-1	1
2	145	0	0
3	143	-2	4
4	145	0	0
5	145	0	0
6	147	2	4
7	145	0	0
8	146	1	1
9	145	0	0
10	145	0	0
11	145	0	0
12	146	1	1
13	145	0	0
14	146	1	1
15	143	-2	4
16	145	0	0
18	143	-2	4
19	144	-1	1
20	143	-2	4
	0.00		25.00
	$\sum_{i=1}^{i=n} X_i = 2750$		$\sum_{i=1}^{i=n} [X_i - (X_{ср})_n]^2 = 25$
	$(X_{ср})_n = 144,7$ ммоль/л		$S_n = \sqrt{25/18} = \sqrt{1,39} = 1,18$ ммоль/л

Далее, используя оставшиеся результаты, вычисляют для них среднеарифметическое значение и стандартное отклонение, после чего процедуру проверки наличия грубых ошибок среди оставшихся результатов повторяют заново. То есть, если среди 20 результатов повторных измерений применение критерия Граббса при выбранном 95% уровне доверия (что соответствует уровню значимости $\alpha=0,05$) выявит грубую ошибку, показывая, что сомнительный результат отклонился от среднего в большую или меньшую сторону более, чем на $2,71 \cdot S_n$, то тогда производят перерасчет среднеарифметического значения и стандартного отклонения по оставшимся 19 результатам. Далее приведен конкретный пример выявления грубой ошибки с 95% уровнем доверия ($\alpha = 0,05$) среди 20 результатов измерений из Таблицы 3.1. Сначала по формулам (3.4) и (3.5) вычисляются значения G_1 и G_2 : $G_1 = (145-143) / 1,62 = 1,235 < G_T = 2,709$ и соответственно $G_2 = (150-145) / 1,62 = 3,09 > G_T = 2,709$. Таким образом, по критерию Граббса среди результатов, приведенных в Таблице 3.1, имеется один выброс, а именно результат под номером 17, который изымается из последующего расчета $(X_{ср})_n$ и S_n . Перерасчет значений $(X_{ср})_n$ и S_n по формулам (3.1) и (3.2) приведен в Таблице 3.4.

Заметим, что оценочное значение S_n стандартного отклонения, вычисленное сначала по 20 результатам измерений, которое мы обозначим как S_1 , а затем после исключения грубой ошибки по 19 результатам измерений, которое мы обозначим как S_2 , является и в том и в другом случае выборочной оценкой истинного значения стандартного отклонения генеральной совокупности результатов повторных измерений. Для проверки того, что и одно и другое выборочные значения S_1 и S_2 характеризуют одно и тоже истинное значение стандартного отклонения σ , точнее говоря, что эти оценочные значения не отличаются статистически значимо друг от друга, в математической статистике используют критерий Фишера, который также называют F-тестом [6]. В соответствии с этим критерием сначала вычисляют значение показателя F по формуле

$$F = \frac{(S_1)^2}{(S_2)^2} \quad (3.6)$$

учитывая при этом, что $S_1 > S_2$, а затем сравнивают его с табличным критическим значением $F_{крит}$. Это критическое значение зависит от уровня значимости (уровня доверия) и от количества результатов n_1 и n_2 , которые использовались для вычисления выборочных значений. При уровне доверия, равном 95%, $S_1 = 1,62$ ммоль/л, $S_2 = 1,18$ ммоль/л, $n_1 = 20$ и $n_2 = 19$ из таблиц критических значений для статистики Фишера находим, что $F_{крит} = 2,21$. И поскольку для этого случая $F = 2,62 / 1,39 = 1,89$, что меньше $F_{крит} = 2,21$, то согласно критерию Фишера полученные выборочные оценки истинного значения стандартного отклонения, равные соответственно 1,62 и 1,18 ммоль/л, статистически значимо друг от друга не отличаются.

Теперь проверим наличие грубой ошибки с 95% уровнем доверия ($\alpha=0,05$) среди оставшихся 19 результатов измерений из Таблицы 3.4. По формулам (3.4) и (3.5) вычисляем значения G_1 и G_2 : $G_1 = (145-143) / 1,18 = 1,695 < G_T = 2,681$ и соответственно $G_2 = (147-145) / 1,18 = 1,695 < G_T = 2,681$. Таким образом, по критерию Граббса среди оставшихся 19 результатов, приведенных в Таблице 3.4, грубых ошибок более нет.

Опыт эксплуатации показывает, что в среднем стабильность калибровки аналитических систем не превышает 48 часов, что может вызывать флуктуации или дрейф систематического смещения АС.

Это, в свою очередь, может ухудшать прецизионность результатов повторных измерений уровня аналита в одном и том же биоматериале, в том числе и при проведении установочных серий. Чтобы выяснить насколько сильно ухудшается прецизионность АС из-за нестабильности ее калибровки можно использовать так называемый метод дубликатов, или иначе метод двойных измерений уровня аналита или в образцах КМ или непосредственно пробах пациентов. Суть метода состоит в том, что в течение n рабочих дней, где $n \geq 20$, ежедневно в одной из аналитических серий измеряют уровень аналита в одном и том же биоматериале дважды, то есть в дубликатах, и затем вычисляют их разности $d_i = X_{1i} - X_{2i}$, где X_{1i} и X_{2i} – первый и второй результаты измерений, полученные в течение рабочего дня с порядковым номером $i = 1, 2, 3, \dots, n$. Вспомним, что из свойств математического ожидания и дисперсии следует, что математическое ожидание μ_d разности d случайных величин X_1 и X_2 равно разности их математических ожиданий $\mu_d = \mu_1 - \mu_2$, а дисперсия σ_d^2 разности двух случайных независимых величин X_1 и X_2 равна сумме их дисперсий $\sigma_d^2 = \sigma_1^2 + \sigma_2^2$. Воспользуемся этими свойствами и тогда получим для случайной величины d_i , что ее математическое ожидание μ_d равно нулю, поскольку $\mu_1 = \mu_2$, а ее дисперсия равна удвоенной дисперсии результатов измерений, поскольку $\sigma_1^2 = \sigma_2^2 = \sigma^2$. И тогда, соответственно, стандартные отклонения σ_d и σ будут связаны формулой: $\sigma_d = \sqrt{2} \cdot \sigma$. Воспользуемся формулой (3.2а) для определения показателя σ и тогда получим формулу для вычисления оценочных значений стандартного отклонения S_n по разностям результатов, полученным методом двойных измерений:

$$S_n = \sqrt{\frac{\sum_{i=1}^n (d_i)^2}{2 \cdot n}} \quad (3.7)$$

где $d_i = X_{1i} - X_{2i}$ - разность между дубликатами, полученными в i -тый рабочий день, n – число биопроб, в которых уровень аналита был измерен в дубликатах. Следует иметь ввиду, что уровень аналита в используемых биопробах должен находиться в пределах диапазона измерений АС, в котором ее прецизионность, характеризуемая стандартным отклонением, практически не меняется. Конкретный пример вычисления для натрия значения S_n при измерении уровня аналита в дубликатах приведен в Таблице 3.5.

Таблица 3.5. Пример вычисления значения S_n по формуле (3.7) методом двойных измерений уровня натрия в биопробах.

Номер пробы, i	Первый результат, X_{1i}	Второй результат, X_{2i}	Модуль разницы $ d_i $	Квадрат разницы d_i^2
1	143	145	2	4
2	144	145	1	1
3	146	147	1	1
4	142	144	2	4
5	145	143	2	4
6	147	146	1	1
7	145	146	1	1
8	146	145	1	1
9	145	146	1	1
10	145	144	1	1
11	145	144	1	1
12	146	148	2	4
13	145	147	2	4
14	146	145	1	1
15	143	144	1	1
16	145	147	2	4
17	146	149	3	9
18	143	144	1	1
19	144	141	3	9
20	145	143	2	4
				$\Sigma d_i^2 = 57$
$S_n = \sqrt{57/40} = \sqrt{1,43} = 1,19$ ммоль/л				

Теперь, используя критерий Фишера, проверим, отличаются статистически значимо друг от друга выборочные оценки S_n , вычисленные в Таблицах 3.1 и 3.5. При уровне доверия, равном 95%, $S_1 = 1,62$ ммоль/л, $S_2 = 1,19$ ммоль/л, $n_1 = 20$ и $n_2 = 20$ из таблиц критических значений для статистики Фишера находим, что $F_{крит} = 2,20$. И поскольку для этого случая $F = 2,62 / 1,19 = 2,20$, что не больше, чем $F_{крит} = 2,20$, то согласно критерию Фишера полученные выборочные оценки истинного значения стандартного отклонения, равные соответственно 1,62 и 1,19 ммоль/л, не отличаются статистически значимо друг от друга. Это говорит о том, что эти выборочные значения стандартного отклонения характеризуют одно и тоже генеральное значение σ . Что, скорее всего, объясняется отсутствием значимых флуктуации или дрейфа калибровки АС за время проведения измерений.

Если установлено, что выбросов среди результатов установочной серии более нет, то тогда считают полученные выборочные значения для среднего $(X_{ср})_n$ и стандартного отклонения S_n наилучшими оценками их истинных значений μ и σ . На основании полученного выборочного значения $(X_{ср})_n$ вычисляют по нижеприведенной формуле оценочное эксплуатационное значение для систематического смещения B_n , которое является выборочной оценкой его истинного значения Δ :

$$\Delta = \mu - A_{\text{КМ}} \quad (3.8)$$

$$B_n = (X_{\text{ср}})_n - A_{\text{КМ}} \quad (3.9)$$

где $A_{\text{КМ}}$ - паспортное значение уровня анализа в используемом контрольном материале. Как видно из формулы (3.9) показатель B_n , в отличие от показателя S_n , может иметь и положительные, и отрицательные значения.

Точность оценки генеральных значений математического ожидания μ и стандартного отклонения σ с использованием соответственно выборочных значений среднеарифметического $(X_{\text{ср}})_n$ и стандартного отклонения S_n логично определять по ширине их доверительных интервалов (далее – ДИ). Доверительным называют интервал, который содержит генеральное значение искомого параметра с заданной надежностью. Можно также сказать, что ДИ содержит с заданной надежностью все выборочные значения искомого параметра. При описании надежности обычно используют два термина: уровень значимости и уровень доверия. Уровень значимости, который обозначают обычно буквой α , - это вероятность, с которой генеральное значение параметра не попадает в доверительный интервал. Уровень доверия, который обычно обозначают буквой p , - это вероятность того, что доверительный интервал содержит генеральное значение искомого параметра. Таким образом $p = 1 - \alpha$. Как правило, уровень значимости выбирают равным 0,01, 0,05 и 0,10, что соответствует уровню доверия 0,99, 0,95 и 0,90. Собственно поэтому ДИ часто называют интервальной оценкой искомого параметра. Очевидно, что чем меньше ширина ДИ для генерального значения искомого параметра, тем точнее его оценка. Из дальнейшего станет видно, что ширина ДИ уменьшается с ростом объема выборки n , поэтому чем больше объем выборки, тем меньше ширина ДИ. ДИ для генеральных значений μ и σ обычно обозначают так: $\mu_{\text{нг}} < \mu < \mu_{\text{вг}}$ и $\sigma_{\text{нг}} < \sigma < \sigma_{\text{вг}}$, где $\mu_{\text{нг}}$ и $\sigma_{\text{нг}}$ являются нижними, а $\mu_{\text{вг}}$ и $\sigma_{\text{вг}}$ - верхними границами ДИ для генеральных значений соответственно μ и σ .

Заметим, что размеры ДИ для генерального значения Δ автоматически вытекают из формулы (3.8). Воспользуемся этой формулой и вышеупомянутым выражением ДИ для μ и тогда получим: $\mu_{\text{нг}} - A_{\text{КМ}} < \mu - A_{\text{КМ}} < \mu_{\text{вг}} - A_{\text{КМ}}$ или $\Delta_{\text{нг}} < \Delta < \Delta_{\text{вг}}$, где $\Delta_{\text{нг}}$ и $\Delta_{\text{вг}}$ являются нижней и соответственно верхней границами ДИ для генерального значения Δ .

Сначала определим степень точности оценки генерального значения μ его выборочными значениями $(X_{\text{ср}})_n$. Из математической статистики известно, что расчет нижней $\mu_{\text{нг}}(n)$ и верхней $\mu_{\text{вг}}(n)$ границ двусторонних ДИ для генерального значения μ проводится по формулам:

$$\mu_{\text{нг}}(n) = \left[(X_{\text{ср}})_n - (t_{\frac{\alpha}{2}, n-1}) \cdot \frac{S_n}{\sqrt{n}} \right] \quad (3.10)$$

$$\mu_{\text{вг}}(n) = \left[(X_{\text{ср}})_n + (t_{\frac{\alpha}{2}, n-1}) \cdot \frac{S_n}{\sqrt{n}} \right] \quad (3.11)$$

где $t_{\alpha/2, n-1}$ - [квантили распределения Стьюдента](#), значения для которых можно найти в интернете или в справочниках по математической статистике; α – уровень значимости или, что тоже самое, значение вероятности выхода значения математического ожидания μ за пределы доверительного интервала; $p=(1-\alpha)$ – выбираемый уровень доверия, то есть доверительная вероятность попадания значения математического ожидания μ в пределы доверительного интервала; n - объем выборки; $n-1$ – число степеней свободы.

Двумя параметрами, задающими выбор значений квантилей из таблиц, являются обычно уровень значимости или доверительная вероятность и число степеней свободы. При выборе нужного значения из таблицы квантилей Стьюдента с уровнями значимости важно помнить, что в данной задаче рассматривается случай двустороннего доверительного интервала или, как иногда говорят, случай двусторонней гипотезы. Поэтому при выборе уровня доверия в 95%, то есть при выборе доверительной вероятности $p = 1 - \alpha = 0,95$, что соответствует уровню значимости $\alpha=0,05$, надо выбирать табличное значение квантиля Стьюдента из колонки для $t_{\alpha/2, n-1}$, то есть из колонки $t_{0,025; n-1}$. Для удобства проведения расчетов в Таблице 3.6 приведены некоторые значения $t_{\alpha/2, n-1}$ для квантилей Стьюдента для двусторонних доверительных интервалов для разных доверительных вероятностей p и разного числа степеней свободы $n-1$.

Степень точности оценки генерального значения μ его выборочными значениями $(X_{ср})_n$ будем определять по ширине его ДИ, а максимально возможную ошибку такой оценки будем определять отношением верхней границы ДИ к его нижней границе.

Покажем на примере типовой расчет границ ДИ для генерального значения μ , используя данные Таблицы 3.4, в которой количество n результатов повторных измерений, для вычислений $(X_{ср})_n$ и S_n , равно 19. Это соответствует числу степеней свободы $n - 1 = 19 - 1 = 18$. Выбирая в Таблице 3.6 число степеней свободы равное 19, и выбирая доверительную вероятность равную 0,95, находим значение квантиля, которое будет равно $t_{0,025; 18} = 2,101$. Соответственно нижняя граница ДИ для генерального значения μ будет равна $144,70 - 2,101 \cdot 1,18 / 4,36 = 144,7 - 0,57 = 144,13$ ммоль/л, а верхняя – $144,7 + 0,57 = 145,27$ ммоль/л. Таким образом, можно утверждать, что с 95% доверительной вероятностью генеральное значение μ будет находиться в диапазоне от 144,13 до 145,27 ммоль/л.

Теперь, вычитая из верхней границы доверительного интервала его нижнюю границу, получим формулу для определения ширины доверительного интервала $\Delta\mu_n$, которая, в свою очередь, при выбранном уровне доверия определяет точность оценки генерального значения математического ожидания μ по выборочному среднему $(X_{ср})_n$:

$$\Delta\mu_n = \left[(X_{ср})_n + (t_{\frac{\alpha}{2}, n-1}) \cdot \frac{S_n}{\sqrt{n}} \right] - \left[(X_{ср})_n - (t_{\frac{\alpha}{2}, n-1}) \cdot \frac{S_n}{\sqrt{n}} \right] = 2 \cdot (t_{\frac{\alpha}{2}, n-1}) \cdot \frac{S_n}{\sqrt{n}} \quad (3.12)$$

Как видно из формулы (3.12), ширина двустороннего доверительного интервала для генерального значения μ зависит только от выбранной доверительной вероятности, объема выборки n и выборочного значения стандартного отклонения S_n . Из математической статистики известно, что при значениях $n > 20$ выборочные значения $(X_{ср})_n$ и S_n также, как и значения квантилей Стьюдента, не изменяются существенно. Поэтому, исходя из формулы (3.12), можно сделать вывод, что точность оценки генерального значения μ с использованием выборочного значения среднего $(X_{ср})_n$ возрастает пропорционально \sqrt{n} , то есть корню квадратному из объема выборки n .

Таблица 3.6. Таблица квантилей $t_{\alpha/2, n-1}$ распределения Стьюдента для случая двусторонней гипотезы.

Число степеней свободы, $k=n-1$	Доверительные вероятности, p		
	0,9	0,95	0,99
1	6,31375	12,70620	63,65590
2	2,91999	4,30266	9,92499
3	2,35336	3,18245	5,84085
4	2,13185	2,77645	4,60408
5	2,01505	2,57058	4,03212
6	1,94318	2,44691	3,70743
7	1,89458	2,36462	3,49948
8	1,85955	2,30601	3,35538
9	1,83311	2,26216	3,24984
10	1,81246	2,22814	3,16926
11	1,79588	2,20099	3,10582
12	1,78229	2,17881	3,05454
13	1,77093	2,16037	3,01228
14	1,76131	2,14479	2,97685
15	1,75305	2,13145	2,94673
16	1,74588	2,11990	2,92079
17	1,73961	2,10982	2,89823
18	1,73406	2,10092	2,87844
19	1,72913	2,09302	2,86094
20	1,72472	2,08596	2,84534
21	1,72074	2,07961	2,83137
22	1,71714	2,07388	2,81876
23	1,71387	2,06865	2,80734
24	1,71088	2,06390	2,79695
25	1,70814	2,05954	2,78744
26	1,70562	2,05553	2,77873
27	1,70329	2,05183	2,77069
28	1,70113	2,04841	2,76326
29	1,69913	2,04523	2,75639
30	1,69726	2,04227	2,74999
60	1,67065	2,00030	2,66027
120	1,65765	1,97993	2,61742

Ширину доверительного интервала $\Delta\mu_n$ для генерального значения μ , рассчитываемую по формуле (3.11), удобнее представлять в относительных единицах, выражая ее в долях выборочного значения стандартного отклонения S_n по формуле:

$$\Delta\mu_n (S_n) = \frac{\Delta\mu_n}{S_n} = 2 \cdot \frac{(t_{\alpha/2, n-1})}{\sqrt{n}} \quad (3.13)$$

Из формулы (3.13) следует, что ширина доверительного интервала для μ в долях выборочного значения S_n зависит только от величины квантиля, практически не изменяющейся при $n > 20$, и объема выборки, что очень удобно для визуализации степени точности интервальной оценки μ в зависимости от объема выборки n . Изменение ширины 95% доверительного интервала $\Delta\mu_n(S_n)$ по мере увеличения объема выборки n приведена на Рис 3.1.

В качестве типового практического примера, воспользуемся данными Таблиц 3.4 и 3.6 и вычислим для них по формулам (3.12) и (3.13) в абсолютных и относительных единицах ширину 95% доверительного интервала $\Delta\mu_n$ и $\Delta\mu_n(S_n)$ для генерального значения μ : $\Delta\mu_{19} = 2 * 2,1 * 1,18 / \sqrt{19} = 4,96 / 4,36 = 1,14$ ммоль/л. Соответственно $\Delta\mu_{19}(S_{19}) = 2 * 2,1 / \sqrt{19} = 4,2 / 4,36 = 0,96$. При объемах выборки $n=10, n=15, n=25$ и $n=30$ для показателей точности интервальной оценки в долях S_n будем иметь: $\Delta\mu_{10}(S_{10}) = 2 * 2,26 / \sqrt{10} = 4,52 / 3,16 = 1,43$; $\Delta\mu_{15}(S_{15}) = 2 * 2,13 / \sqrt{15} = 4,26 / 3,87 = 1,10$; $\Delta\mu_{25}(S_{25}) = 2 * 2,06 / \sqrt{25} = 4,12 / 5 = 0,82$ и $\Delta\mu_{30}(S_{30}) = 2 * 2,04 / \sqrt{30} = 4,08 / 5,47 = 0,75$. То есть точность интервальной оценки генерального значения μ с помощью выборочной средней $(X_{ср})_n$ при объеме выборки $n=30$ будет почти в 2

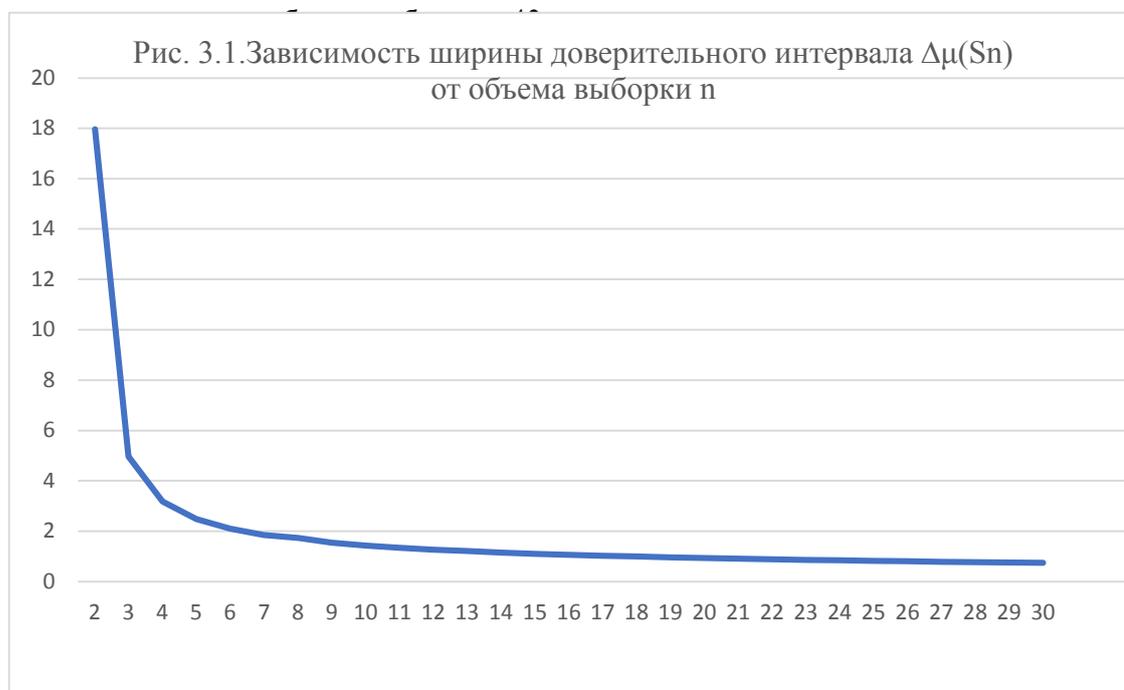


Рис. 3.1. Зависимость ширины доверительного интервала для генерального значения μ от объема выборки n в долях выборочного значения S_n стандартного отклонения σ .

Теперь перейдем к определению ширины ДИ для дисперсии σ^2 , а затем и для стандартного отклонения σ . Из математической статистики известно, что нижняя $\sigma^2_{нг}(n)$ и верхняя $\sigma^2_{вр}(n)$ границы двустороннего доверительного интервала для дисперсии σ^2 при неизвестном генеральном значении μ вычисляются по формулам (3.14) и (3.15):

$$\sigma^2_{нг}(n) = \left[(n - 1) \cdot \frac{S_n^2}{\chi^2_{\frac{1-\alpha}{2}, n-1}} \right] \quad (3.14)$$

$$\sigma^2_{\text{ВГ}}(n) = \left[(n - 1) \cdot \frac{S_n^2}{\chi^2_{\frac{1+\alpha}{2}, n-1}} \right] \quad (3.15)$$

где α - доверительная вероятность, $\chi^2_{(1-\alpha)/2, n-1}$ и $\chi^2_{(1+\alpha)/2, n-1}$ - [квантили распределения](#) χ^2 (Хи-квадрат). Для удобства проведения расчетов некоторые значения квантилей χ^2 для разных доверительных вероятностей α и разного числа степеней свободы $n-1$ приведены в Таблице 3.7.

Исходя из своего определения, ДИ для дисперсии σ^2 и для стандартного отклонения σ будут выглядеть следующим образом:

$$\frac{(n - 1)S_n^2}{\chi^2_{\frac{1-\alpha}{2}, n-1}} < \sigma^2 < \frac{(n - 1)S_n^2}{\chi^2_{\frac{1+\alpha}{2}, n-1}} \quad (3.16)$$

$$S_n \cdot \sqrt{\frac{n - 1}{\chi^2_{(1-\alpha)/2, n-1}}} < \sigma < S_n \cdot \sqrt{\frac{n - 1}{\chi^2_{(1+\alpha)/2, n-1}}} \quad (3.17)$$

Степень точности оценки генерального значения σ его выборочными значениями S_n будем определять по ширине его ДИ, а максимально возможную ошибку такой оценки будем определять отношением верхней границы ДИ к его нижней границе.

Таблица 3.7. Таблица квантилей (критических значений) распределения Хи-квадрат.

Число степеней свободы, $k = n - 1$	Уровни значимости для критических значений			
	0,025	0,05	0,95	0,975
1	5,0	3,8	0,0039	0,00098
2	7,4	6,0	0,103	0,051
3	9,4	7,8	0,352	0,216
4	11,1	9,5	0,711	0,484
5	12,8	11,1	1,15	0,831
6	14,4	12,6	1,64	1,24
7	16,0	14,1	2,17	1,69
8	17,5	15,5	2,73	2,18
9	19,0	16,9	3,33	2,70
10	20,5	18,3	3,94	3,25
11	21,9	19,7	4,57	3,82
12	23,3	21,0	5,23	4,40
13	24,7	22,4	5,89	5,01
14	26,1	23,7	6,57	5,63
15	27,5	25,0	7,26	6,26
16	28,8	26,3	7,96	6,91
17	30,2	27,6	8,67	7,56
18	31,5	28,9	9,39	8,23
19	32,9	30,1	10,1	8,91
20	34,2	31,4	10,9	9,59
21	35,5	32,7	11,6	10,3
22	36,8	33,9	12,3	11,0
23	38,1	35,2	13,1	11,7
24	39,4	36,4	13,8	12,4
25	40,6	37,7	14,6	13,1
26	41,9	38,9	15,4	13,8
27	43,2	40,1	16,2	14,6
28	44,5	41,3	16,9	15,3
29	45,7	42,6	17,7	16,0
30	47,0	43,8	18,5	16,8
60	83,3	79,1	43,2	40,48
120	152,2	146,6	95,7	91,6

Теперь покажем на примере типовой расчет границ ДИ для генеральных значений σ^2 и σ , используя данные, приведенные в Таблице 3.4, в которой количество n результатов повторных измерений, для вычисления выборочного значения S_n , равно 19. Это соответствует числу степеней свободы $n - 1 = 19 - 1 = 18$. Выбирая в Таблице 3.7 число степеней свободы равное 18, и используя доверительную вероятность α равную 0,95, находим в ней нужные значения квантилей, которые будут равны $\chi^2_{(1-\alpha)/2, n-1} = \chi^2_{[(1-0,95)/2, (19-1)]} = \chi^2_{[(0,025), 18]} = 31,53$ и соответственно $\chi^2_{(1+\alpha)/2, n-1} = \chi^2_{[(1+0,95)/2,$

$(19-1) = \chi^2 [(0,975), 18] = 8,23$. Таким образом, рассчитываемые по формулам (3.14) и (3.15) нижняя и верхняя границы ДИ для генерального значения σ^2 будут равны $\sigma_{\text{нг}}^2(19) = 18 * 1,18 * 1,18 / 31,53 = 0,795$ и соответственно $\sigma_{\text{вр}}^2(19) = 18 * 1,18 * 1,18 / 8,23 = 3,045$. И в конечном итоге, получим ДИ для генерального значения стандартного отклонения σ в ммоль/л: $0,89 < \sigma < 1,75$. Таким образом можно утверждать, что с 95% доверительной вероятностью генеральное значение σ будет находиться в диапазоне от 0,89 до 1,75 ммоль/л, хотя для выборочной оценки было получено значение $S_{(19)} = 1,18$ ммоль/л.

Теперь воспользуемся формулой (3.18) и на основе ширины соответствующего ДИ определим точность выборочной оценки генерального значения дисперсии σ^2 и стандартного отклонения σ . Ширина ДИ $\Delta\sigma_n^2$ для дисперсии σ^2 будет равна:

$$\Delta\sigma_n^2 = (n - 1) \cdot S_n^2 \left[\frac{1}{\chi_{\frac{1+\alpha}{2}, n-1}^2} - \frac{1}{\chi_{\frac{1-\alpha}{2}, n-1}^2} \right] \quad (3.18)$$

Соответственно ширина ДИ $\Delta\sigma_n$ для показателя σ будет равна разнице корней квадратных из верхней и нижней границ ДИ для дисперсии σ^2 , а именно:

$$\Delta\sigma_n = \sqrt{n - 1} \cdot S_n \cdot \left\{ \sqrt{\frac{1}{\chi_{(1+\alpha)/2, n-1}^2}} - \sqrt{\frac{1}{\chi_{(1-\alpha)/2, n-1}^2}} \right\} \quad (3.19)$$

И, наконец, формула для расчета ширины ДИ $\Delta\sigma_n(S_n)$ для генерального значения σ в долях выборочного значения S_n будет иметь вид:

$$\Delta\sigma_n(S_n) = \frac{\Delta\sigma_n}{S_n} = \sqrt{n - 1} \cdot \left\{ \sqrt{\frac{1}{\chi_{(1+\alpha)/2, n-1}^2}} - \sqrt{\frac{1}{\chi_{(1-\alpha)/2, n-1}^2}} \right\} \quad (3.20)$$

Динамика изменения ширины 95% доверительного интервала $\Delta\sigma_n(S_n)$ по мере увеличения количества n результатов повторных измерений представлена на Рис. 3.2.

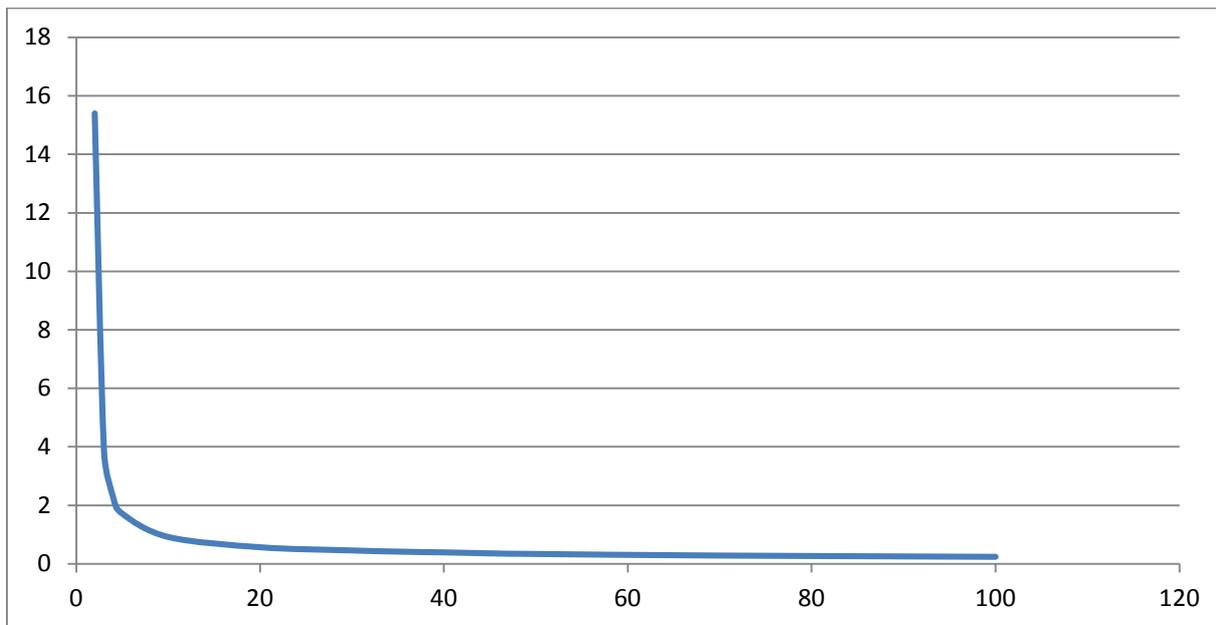


Рис. 3.2. Динамика изменения ширины доверительного интервала $\Delta\sigma_n(S_n)$ от объема выборки n в единицах стандартной ошибки S_n .

Как видно из Рис. 3.2, точность оценки генерального значения по его выборочным значениям изменяется просто драматически при уменьшении объема выборки с 20 до нескольких единиц. Так при 95% уровне доверия и при наличии выборки только из двух ($n=2$) результатов повторных измерений ширина ДИ для генерального значения σ , оцениваемого в долях выборочной дисперсии S_2 , оказывается близкой к $100 \cdot S_2$, а генеральное значение σ , характеризующее случайную составляющую ошибки результатов, получаемых при использовании исследуемой аналитической системы, может реально иметь любое значение от $0,44 \cdot S_2$ до $100 \cdot S_2$. Таким образом, при наличии только двух результатов повторных измерений ошибка в оценке генерального значения σ по выборочному значению S_2 может достигать $100 / 0,44 = 230$ раз. Соответственно, при том же 95% уровне доверия в случае выборки объемом $n=3$ ширина ДИ для σ оказывается равной $5,76 \cdot S_3$, а генеральное значение σ реально может иметь любое значение от $0,52 \cdot S_3$ до $6,28 \cdot S_3$. И таким образом, при наличии трех результатов повторных измерений ошибка в оценке генерального значения σ по выборочному значению S_3 может достигать $6,28 / 0,52 = 12$ раз. При том же 95% уровне доверия и при выборке объемом $n=10$ ширина ДИ для σ оказывается равной $1,14 \cdot S_{10}$, границы ДИ оказываются равными $0,69 \cdot S_{10}$ и $1,83 \cdot S_{10}$, а ошибка в оценке генерального значения σ по выборочному значению S_{10} может достигать $1,83 / 0,69 = 2,65$ раз. При 95% уровне доверия и при выборках объемом $n=20$, которые обычно получают в установочных сериях, ширина ДИ для σ оказывается равной $0,70 \cdot S_{20}$, границы ДИ оказываются равными $0,76 \cdot S_{20}$ и $1,46 \cdot S_{20}$, а ошибка в оценке генерального значения σ по выборочному значению S_{20} может достигать $1,46 / 0,76 = 1,9$ раз. Иными словами, используя полученное выборочное значение S_{20} для оценки генерального значения σ , можно ошибиться почти в 2 раза. И, наконец, при 95% уровне доверия и при выборках объемом $n=25$, которые рекомендуется использовать руководством CLSI EP15-A3 для проведения выборочных оценок аналитических характеристик, ширина ДИ для σ оказывается равной $0,61 \cdot S_{25}$, границы ДИ оказываются равными $0,78 \cdot S_{25}$ и $1,39 \cdot S_{25}$, а ошибка в оценке генерального значения σ по выборочному значению S_{25} может достигать $1,39 / 0,78 = 1,8$ раз.

Отдельно отметим, что при том же 95% уровне доверия и объемах выборки $n=50$; $n=100$; $n=150$; $n=200$; $n=250$ и $n=300$, максимальные ошибки в оценке генерального значения σ по их выборочным значениям могут составить соответственно 1,49; 1,32; 1,26; 1,22; 1,19 и 1,17 раз. Именно потому, что точность оценки генеральных значений μ и σ с использованием их выборочных значений $(\bar{X})_n$ и S_n постоянно возрастает с ростом объема выборки n , все руководства по ведению внутрилабораторного контроля качества рекомендуют периодически (по мере накопления результатов повторных контрольных измерений) пересчитывать выборочные значения для μ и σ и использовать эти новые более точные оценки в качестве обновленных параметров контрольной карты. Обычно рекомендуют переопределять параметры контрольной карты через каждые последующие 20-30 контрольных результатов.

Еще раз напомним, что выборочные значения V_n и S_n являются только приблизительными оценками генеральных значений смещения Δ и стандартного отклонения σ . И если повторить еще раз аналогичную серию повторных измерений количеством n , то скорее всего, для показателей V_n и S_n будут получены значения, отличные от полученных в первой серии, хотя, возможно, и близкие к ним. Причина такого отличия выборочных значений между собой заключается главным образом в том, что обе выборки имеют хотя и статистически значимый, но, тем не менее, ограниченный объем. Следует отметить, что при многократном повторении серий повторных измерений с одними тем же количеством n , получаемые на их основе выборочные значения будут с выбранной доверительной вероятностью попадать в ДИ для генерального значения соответствующего показателя. В этой связи, получаемые выборочные оценки V_n и S_n генеральных значений Δ и σ следует использовать с учетом того, что они не абсолютно точно описывают последние и могут существенно отличаться от них, даже в разы.

В настоящей лекции описаны процедуры общепринятых технологий определения на практике эксплуатационных значений основных аналитических характеристик АС, а также рассмотрены вопросы точности определения этих значений.

Краткие выводы по Лекции 3.

1) Генеральные значения основных аналитических характеристик АС, к которым относятся систематическое смещение и стандартное отклонение, точно определять не представляется возможным. На практике эти значения только оценивают, используя для этого результаты повторных измерений ограниченного объема, которые, в свою очередь, получают в процессе проведения установочных серий.

2) При проведении установочной серии национальные руководства рекомендуют проводить ежедневно по 1 измерению уровня аналита в контрольном материале в течение 20 рабочих дней, а международные - по 5 измерений ежедневно в течение 5 рабочих дней.

3) Результаты, полученные в установочной серии, в руководстве CLSI EP15-A3 рекомендуется проверять на отсутствие среди них грубых ошибок.

4) Генеральное значение систематического смещения аналитической системы за время проведения установочной серии может изменяться и даже существенно, что обычно обуславливается дрейфом и/или флуктуацией ее калибровки. В этой связи желательно проводить методом дубликатов дополнительную проверку адекватности выборочной оценки стандартного отклонения по данным, полученным при проведении установочной серии.

5) Используя полученные выборочные значения и выбранный уровень доверия, определяют границы доверительных интервалов для генеральных значений основных аналитических характеристик.

Литература к Лекции 3.

1. ГОСТ ISO/IEC 17025-2019. Общие требования к компетентности испытательных и калибровочных лабораторий.
2. Приказ МЗ РФ от 26.05.2003 №220 «Об утверждении отраслевого стандарта «Правила проведения внутрилабораторного контроля качества количественных методов клинических лабораторных исследований с использованием контрольных материалов». Приложение. «ОСТ 91500.13.0001-2003».
3. ГОСТ Р 53133.2-2008. Национальный стандарт РФ. Технологии лабораторные клинические. Контроль качества клинических лабораторных исследований. Часть 2. Правила проведения внутрилабораторного контроля качества количественных методов клинических лабораторных исследований с использованием контрольных материалов.
4. Clinical and Laboratory Standards Institute «User Verification of Precision and Estimation of Bias». CLSI Guideline. Document EP15-A3. September, 2014.
5. ГОСТ Р 8.736-2011. Государственная система обеспечения единства измерений (ГСИ). Измерения прямые многократные. Методы обработки результатов измерений. Основные положения.
6. Вентцель Е.С. Теория вероятностей. - 4-е изд. - М.: Наука, 1969. - 576 с.