Статистический анализ результатов мониторинга

Или что необходимо знать и уметь первокурснику… и магистранту

$$S_{\overline{X}} = \sqrt{\frac{d_1^2 + d_2^2 + d_3^2 + \cdots + d_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}}$$

- Каждому физическому измерению присуща некоторая погрешность, которая в лучшем случае может быть снижена лишь до какого-то приемлемого уровня. Определение величины этой погрешности нередко представляет сложную задачу, требующую от исследователя дополнительных усилий, изобретательности и интуиции. Тем не менее, этой работой нельзя пренебрегать, так как результаты *анализа*, выполненного с неизвестной степенью надежности, не имеют научной ценности. Напротив, не очень точный результат может оказаться весьма важным, если с высокой степенью надежности можно установить пределы возможных ошибок. К сожалению, не существует простого общего приема абсолютно точной оценки качества экспериментальных результатов. Поэтому нет ничего удивительного в том, что обработка результатов нередко представляет задачу не меньшей сложности, чем их получение. Эта работа включает изучение литературы, калибровку прибора, дополнительные эксперименты, специально разработанные для выявления причин возможных ошибок, и статистический анализ данных. Следует признать, что на каждом этапе также возможны ошибки. В конечном счете, исследователь может лишь *оценить* возможную достоверность измерения: чем опытнее исследователь, тем более строгими и менее оптимистичными становятся подобного рода суждения.
- Достоверность аналитических измерений прямо зависит от времени и усилий, затраченных на их получение. Чтобы добиться десятикратного увеличения точности может понадобиться дополнительная работа в течение многих часов, дней или даже недель. Поэтому опытный исследователь в первую очередь устанавливает желаемую степень достоверности результата, так как это определит затраты времени и труда на выполнение анализа. Тщательное продумывание исследования в самом начале часто обеспечивает большую экономию времени и труда. *Не следует тратить много времени в погоне за высокой точностью там, где она не нужна.*

- Далее мы обсуждаем типы ошибок, возникающих при проведении анализа, методы их выявления, а также способы оценки и представления их величин.
- Обычно исследователь повторяет анализ от 2 до 5 раз. Два раза – минимум, т.к. необходимо, чтобы результаты были достоверными.
- Чем больше выборка, тем достовернее результаты, но по двум значениям можно сделать выводы.
- <u>Некоторые определения</u>
- **Сходимость** – близость результатов, выполненных одним аналитиком (исследователем) за короткий период времени.
- **Воспроизводимость** – близость результатов, выполненных разными исследователями или за более длительный промежуток времени.
-
- Согласно ГОСТ Р ИСО 5725-1-2002
- **Прецизионность** – степень близости друг к другу независимых результатов измерений, полученных в конкретных установленных условиях.
- **Показатели** прецизионности: **сходимость**, **воспроизводимость**.
- **Сходимость** – прецизионность в условиях повторяемости.
- **Воспроизводимость** – прецизионность в условиях воспроизводимости.
- **Правильность** – степень близости полученного значения к значению, принятому за действительное, и выражается ошибкой (погрешностью).

Правильность

Абсолютная ошибка

$\Delta = Xi - X\text{ст}$

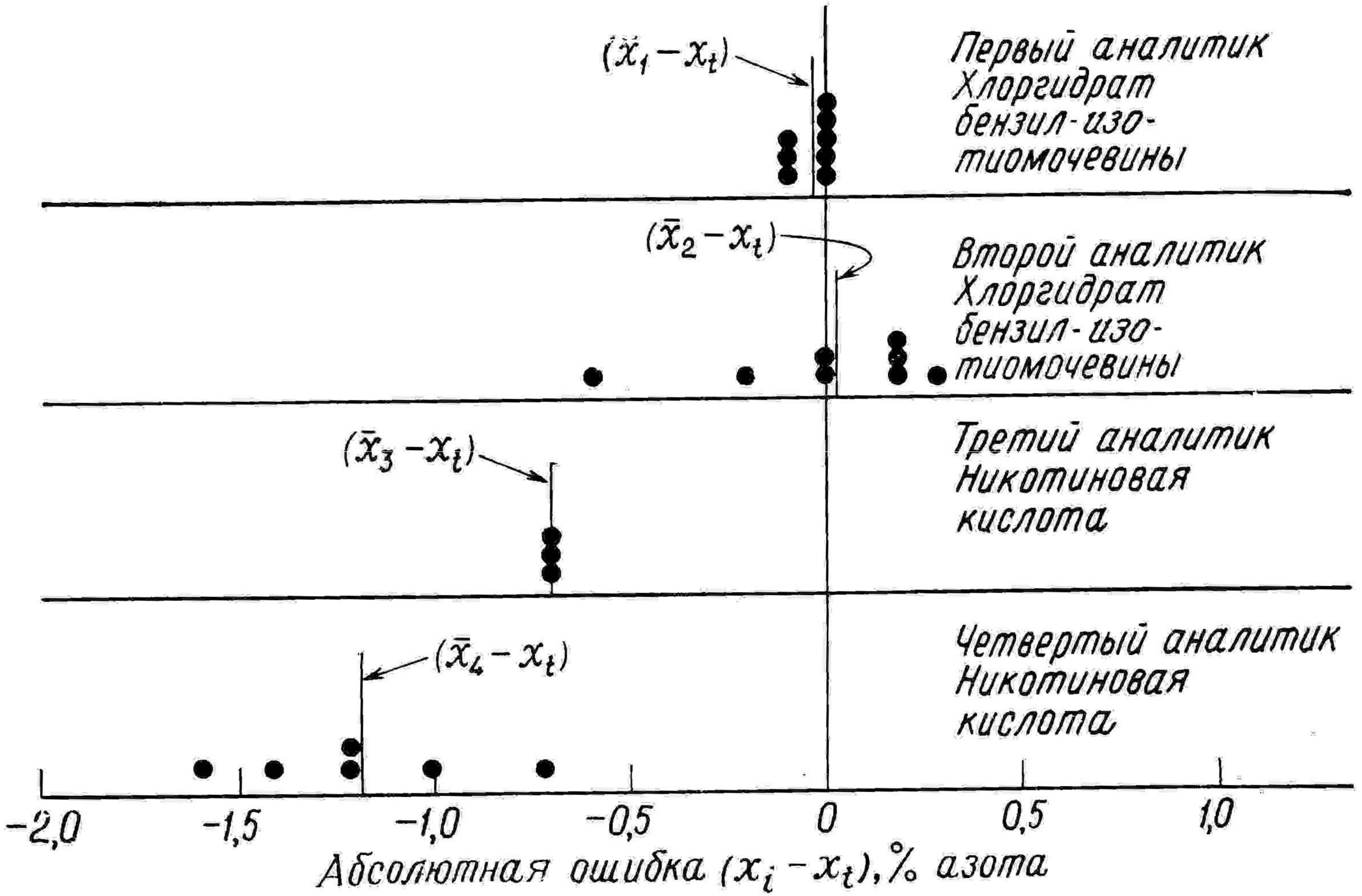
Относительная ошибка

$\delta = Xi - X\text{ст}/X\text{ст} \cdot 100\%$

X_i – измеренное значение,

$X_{\text{ст}}$ – опорное значение.

Нужно понимать, что хорошая сходимость не означает, что высока правильность. Например, на рисунке представлены результаты определения азота, полученные четырьмя аналитиками. Точки, нанесенные на диаграмму, означают абсолютные ошибки параллельных измерений в каждом образце, допущенные каждым аналитиком.



- Обратите внимание на то, что аналитик 1 получил относительно высокую *воспроизводимость* и высокую *правильность*. Аналитик 2, напротив, получил плохую *воспроизводимость*, но хорошую *правильность*. Результаты аналитика 3 нельзя признать хорошими; он добился исключительно высокой *воспроизводимости*, но в среднем значении результатов им допущена заметная ошибка. Исследователь сталкивается также с ситуацией, подобной той, с которой столкнулся аналитик 4, когда и *воспроизводимость* и *правильность* плохие.
- Наблюдаемую на рисунке картину можно объяснить, предположив, что при проведении эксперимента допущены ошибки двух основных типов, причем ошибки одного типа не связаны с воспроизводимостью измерений.

Классификация ошибок

Ошибки

Систематические

Случайные

Ошибки, величину которых, если не на практике, то в принципе можно измерить и учесть. Разность между средним и действительным значением (рисунок 3, 4) обусловлена одной или несколькими систематическими ошибками.

Ошибки, появляющиеся в результате многократных повторных измерений. Происхождение их неизвестно, а величина колеблется произвольно и не может быть измерена. Рассеяние единичных результатов около среднего (рисунок 2, 4) является прямым следствием случайных ошибок.

Типы систематических ошибок

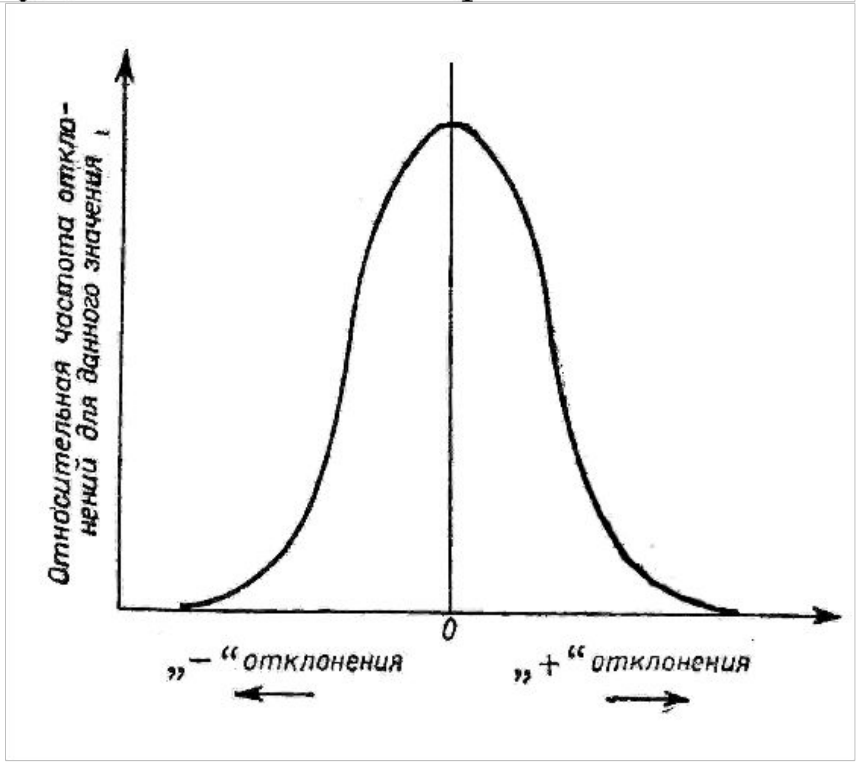
- Индивидуальные – ошибки, возникающие в результате незнания, небрежности, предвзятости или физических недостатков.
- Инструментальные – ошибки приборов (измерительного оборудования).
- Ошибки метода анализа.
-
- Постоянная – величина постоянной ошибки не зависит от измеряемого количества.
- Изменяющаяся – линейно изменяющаяся ошибка, наоборот, уменьшается или возрастает по абсолютной величине пропорционально размеру пробы, взятой для анализа.

Статистическая обработка полученных результатов

- При обсуждении общих принципов количественного анализа уже были затронуты вопросы воспроизводимости и точности определения. Показано, что в то время как систематические ошибки отражаются на точности определения, случайные ошибки определяют их воспроизводимость. Вероятность появления положительных и отрицательных случайных ошибок одинакова, в силу чего благодаря применению статистических методов можно получить достоверную оценку воспроизводимости для данного определения. Нужно подчеркнуть, однако, что статистической обработке поддаются только случайные ошибки.
- Случайные ошибки обладают нормальным распределением, которое графически изображается так называемой Гауссовой кривой, представленной на рисунке. Эта кривая отражает две основные зависимости, которым подчиняются случайные ошибки:
- а) малые отклонения от действительного значения более часты, чем большие;
- б) частота, с которой встречаются положительные и отрицательные ошибки, одинакова.

Из кривой следует непосредственный вывод о том, что насколько больше будет число измерений, настолько ближе к нулю будет среднее отклонение, т. е. определенная величина будет все больше приближаться к действительной.

Отклонения измеренных значений от действительных могут характеризоваться несколькими путями. В основе любого из них лежит допущение, основанное на Гауссовом распределении, о том, что среднее значение большого числа измерений практически идентично действительному. Если измеренные значения для отдельных определений обозначим как X_1, X_2, X_3, … X_n, то *среднее арифметическое значение* представляет



$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- Под *отклонением* d_i величины, полученной при i-м измерении, понимается величина
- $d_i = X_i - \overline{X}$
- а *среднее отклонение* $\overline{d}$ для всех n исследований определяется как среднее арифметическое значение величин d_i без учета их знака, т. е.
- $\overline{d} = \frac{|d_1| + |d_2| + |d_3| + \cdots + |d_n|}{n} = \frac{\sum_{i=1}^{n}|d_i|}{n} = \frac{\sum_{i=1}^{n}|X_i - \overline{X}|}{n}$
- Среднее отклонение можно выразить и в процентах от среднего арифметического $\overline{X}$:
- $\overline{d} = \frac{\overline{d}}{\overline{X}} \cdot 100\%$
- *Стандартное отклонение* $S_{\overline{X}}$ для одной серии из *n* параллельных анализов определяется как
- $S_{\overline{X}} = \sqrt{\frac{d_1^2 + d_2^2 + d_3^2 + \cdots + d_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}}$
- Если число исследований *n* – меньше 10, для выражения $S_{\overline{X}}$ предпочитают зависимость
- $S_{\overline{X}} = \sqrt{\frac{d_1^2 + d_2^2 + d_3^2 + \cdots + d_n^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}$

Пример

- Определить среднее и стандартное отклонения для серии из четырех определений, которые дали следующие результаты: 18,50; 18,68; 18,43; 18,70 г.

	18,50	-0,08	0,0064
	18,68	0,10	0,0100
	18,43	-0,15	0,0225
	18,70	0,12	0,0144
Σ	**74,31**	**0,45**	**0,0533**

- $\overline{X} = \frac{74{,}31}{4} = 18{,}58;\ \ \overline{d} = \frac{0{,}45}{4} = 0{,}11;$
- $S_{\overline{X}} = \sqrt{\frac{0{,}0533}{3}} = 0{,}13;\ \ \overline{d} = \frac{0{,}11}{18{,}58} \cdot 100\% = 0{,}59\%$

- Стандартное отклонение можно выразить также в процентах от среднего арифметического $\bar{X}$. Полученные таким образом значения называются *коэффициентом вариации*. В рассмотренном примере коэффициент вариации равен:
- $S = \frac{0{,}13}{18{,}58} \cdot 100\% = 0{,}70\%$
- Математически можно показать, что при очень большом числе определений ($n \to \infty$), когда распределение случайных ошибок строго следует Гауссовой кривой, 68% отдельных определений отличаются от действительного значения на величину, меньшую стандартного отклонения, т. е. $X_i = \bar{X} \pm S_{\bar{X}}$. В интервал $\bar{X} \pm 2S_{\bar{X}}$ попадает 95% определений, а в интервал $\bar{X} \pm 2{,}5S_{\bar{X}}$ - 99% из них. На практике, однако, вероятность того, что определение находится в интервале $\bar{X} \pm 2S_{\bar{X}}$ меньше 95% по двум причинам. С одной стороны, аналитик проводит конечное, а то и очень ограниченное число определений, а, с другой стороны, Гауссово распределение справедливо только при отсутствии систематических ошибок, которых при анализе нельзя избежать полностью. Поэтому нельзя быть уверенным в том, что среднее арифметическое измерений истинно приближается к действительному значению. Тем не менее, благодаря статистической обработке можно определить, какова вероятность того, что действительное значение лежит в определенном интервале, т. е. найти *доверительный интервал*, в котором находится искомая величина с *определенной статистической вероятностью*. Доверительный интервал μ, можно определить по зависимости
- $\mu = \bar{X} \pm \frac{t \cdot S_{\bar{X}}}{\sqrt{n}}$

- которая показывает, что μ уменьшается обратно пропорционально корню квадратному из числа определении. Коэффициент *t* известный как *фактор (критерий) Стьюдента*, приводится в таблицах и имеет разное значение в зависимости от статистической вероятности и числа определений n. Значения *t* быстро уменьшаются с возрастанием n от 2 до 5, после чего изменяются сравнительно медленно. Данные таблицы *критерия Стьюдента* показывают, что увеличение числа определений более 10 связано с очень маленьким изменением искомого значения, которое практически не оправдывает израсходованного времени и реактивов, так как очень мало сужает доверительный интервал.

n	Статистическая вероятность			*n*	Статистическая вероятность		
	50%	95%	99%		50%	95%	99%
2	1,000	12,706	63,657	7	0,718	2,447	3,707
3	0,816	4,303	9,925	8	0,711	2,365	3,500
4	0,765	3,182	5,841	9	0,706	2,306	3,355
5	0,741	2,776	4,604	10	0,703	2,262	3,250
6	0,727	2,571	4,032	20	0,687	2,086	2,845
				∞	0,674	1,960	2,576

- Зависимость $\mu = \bar{X} \pm \frac{t \cdot S_{\bar{X}}}{\sqrt{n}}$ и данные таблицы *критерия Стьюдента* позволяют при проведении нескольких параллельных измерений прикинуть с определенным процентом вероятности, каков будет интервал, в котором заключено действительное значение искомой величины. Чаще всего в практике используют статистическую вероятность 95%.

Пример

- Определить доверительный интервал при 50, 95 и 99%-ном уровне статистической вероятности для искомой массы в предыдущем примере,
- $\bar{X} = 18{,}58$ г; $S_{\bar{X}} = 0{,}13$; $n = 4$
- --------------------------------------
- $t_{50\%} = 0{,}765$; $t_{95\%} = 3{,}182$; $t_{99\%} = 5{,}841$. Из уравнения следует:
- $\mu_{50\%} = 18{,}58 \pm 0{,}05$; $\mu_{95\%} = 18{,}58 \pm 0{,}21$; $\mu_{99\%} = 18{,}58 \pm 0{,}76$.

- ## *Ответ*

- Согласно полученному результату существует 50%-ная вероятность того, что действительная масса находится в интервале от 18,53 до 18,63 г, 95%-ная вероятность – в интервале от 18,37 до 18,79 г и 99%-ная вероятность – от 17,82 до 19,34 г. Видно, что с увеличением статистической вероятности границы доверительного интервала значительно расширяются.

- Полученные при непосредственном измерении величины неизбежно содержат ошибки, обусловленные самыми разными причинами. Среди этих ошибок следует различать систематические и случайные.
- Систематические ошибки обусловливаются причинами, действующими вполне определённым образом. Примером систематической ошибки при взвешивании может являться смещение стрелки ненагруженных весов относительно нулевой отметки на некоторую постоянную величину Δm . Зная это смещение (например, взвесив гирю, масса которой точно известна), можно, всякий раз измеряя массу на этих весах, вычитать Δm из показаний прибора. Таким образом, систематические ошибки могут быть устранены или достаточно точно учтены.
- Случайные ошибки вызываются большим числом отдельных причин, действующих в каждом отдельном измерении различными способами. В примере со взвешиванием это могут быть незаметные глазу колебания чаши весов, потоки воздуха, толчки фундамента здания, в котором стоят весы. Эти ошибки полностью исключить невозможно.
- Погрешность измерения не является суммой систематической и случайной ошибок!
- $\delta^2 = \delta^2_{\text{сист}} + \delta^2_{\text{случ}}$ или $\sigma = \sqrt{\delta^2_{\text{сист}} + \delta^2_{\text{случ}}}$
- Таким образом, погрешность измерения $\Delta = 2\delta$.

Пример

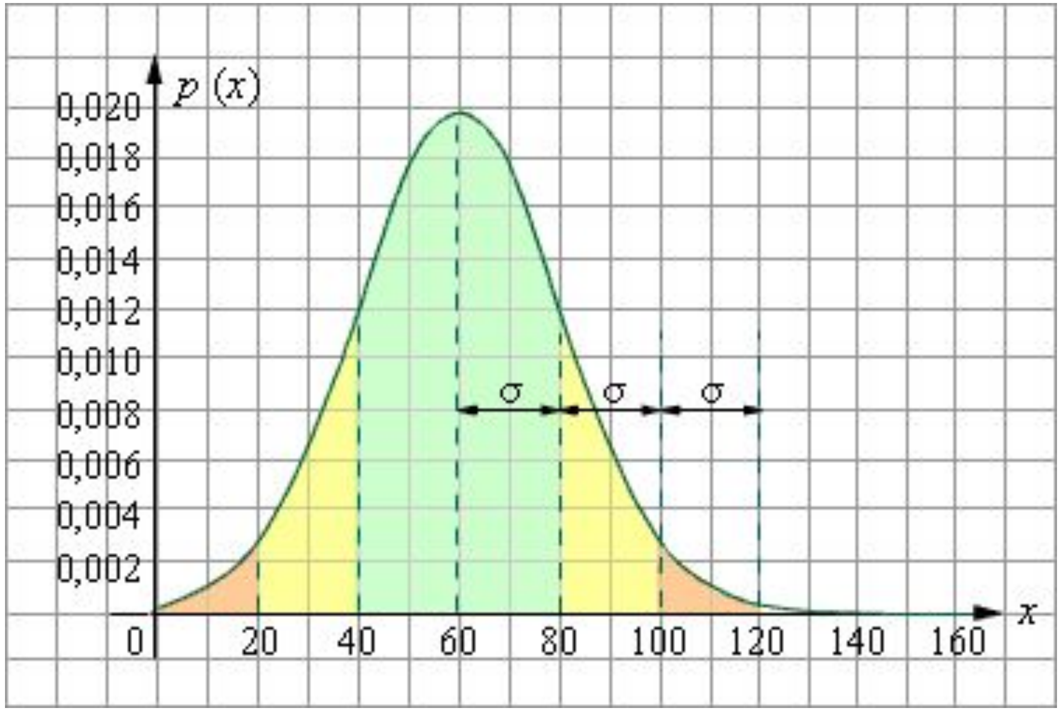
- Определите, какие ошибки из перечисленных являются случайными:
  1) ошибка при однократном измерении сопротивления проводника;
  2) отклонение значения сопротивления проводника от измеренного более точным прибором в процессе измерения сопротивления одного и того же проводника 100 раз в одну и ту же сторону;
  3) однократное измерение диаметра сосуда;
  4) отклонение значения внутреннего диаметра одного и того же сосуда при измерении 30 раз в разные стороны.

- ## Решение

- В случае 2) ошибка является суммой систематической и случайной ошибок, потому что отклонение каждый раз происходит в одну и ту же сторону. Если бы отклонение каждый раз происходило ещё и на одну и ту же величину, то ошибка была бы чисто систематической. В случае 4) отклонение зафиксировано в разные стороны – это признак того, что систематическая ошибка, если и есть, то меньше случайной. Про ошибки в 1) и 3) определённо ничего сказать нельзя, так как сделано всего одно измерение.

Случайная ошибка, возникающая при измерении некоторой величины, может теоретически принимать любые значения. Она является непрерывной случайной величиной, подчинённой определённому закону распределения вероятности.

Расчёты показывают, что в 68,27 % отклонения случайной величины, распределённой по нормальному закону, не превышают σ, в 95,45 % – 2σ. Наконец, вероятность того, что случайная величина, распределённая нормально, отклоняется от математического ожидания больше, чем на 3σ, пренебрежимо мала и составляет 0,27 % – правило трёх сигм.



Уильям Силей Госсет

- Уильям Силей Госсет родился 13 июня 1876 г. в Английском городе Кантербури, Англия, где он был самым старшим из пяти детей. Он умер в возрасте 61 года в Английском городе Биконсфилде 16 октября 1937 г. Он посещал Королевскую Военную Академию в Вулидже для того, чтобы стать инженером прежде, чем он был отклонен из-за плохого зрения. Уильям Госсет никогда не работал как статистик.
- Он пошел в школу в Уинчестере и был хорошо образован перед поступлением в Новый Колледж в Оксфорде. Здесь он завоевал первую степень по химии в 1899 году. После получения своей степени химика он получил работу в пивоваренном заводе Гуиннеса в Дублине в 1899, где он выполнял важную статистическую работу, но на которую никогда не нанимали статиста. Именно окружающая среда в Гуиннессе сделала его статистом. Пивоваренный завод был заинтересован в том, чтобы они могли делать лучшее пиво.
- В 1900 году была открыта Научно-исследовательская лаборатория Гуиннесса, которую возглавил наиболее выдающийся молодой химик Хорас Броун. Хорас Броун наряду с другими варевами задавался вопросом, как получить сырье для назревающего пива наиболее дешево, но получить при этом максимум. Было много факторов, которые они были должны принять во внимание типа множества видов ячменя и хмеля, какие условия изготовления, факторы культивирования и назревания.
После нескольких лет исследования, учитывая что им давали свободу в исследовании условий назревания. Это дало Госсету шанс, чтобы работать как статистик. Он был способен брать данные от различных примеров назревания, что помогало ему выяснить, который путь был лучшее. Поскольку молодые пивовары работают вместе, это казалось естественным для них, чтобы собирать данные для Госсета, чтобы решить числовые проблемы.
- Госсет в 1903 году мог вычислять стандартные ошибки. В 1904 он написал о назревании пива. Этот рапорт привел к Карлу Пирсону, консультирующему Госсета. Госсет встретил Пирсона в июле 1905, когда они долго вместе говорили. Пирсон за полтора часа заставил Госсета понимать теорию стандартных ошибок. Госсет возвратился к пивоваренному заводу и занимался тем методом в течение следующего года. Встреча была также успешна, в котором Пеарсон заставил Госсета взяться за изучение закона ошибок.
- Госсет написал работу в свое свободное время под псевдонимом "Student". Его работа рассматривала вероятности средних ошибок и коэффициент корреляции для публикации. Госсет даже сумел управлять совместными экспериментами с Хантером и Беннеттом в Баллинакурре, Баффином в Кембридже, и Бивином в Уарминстре в испытании одних семян против других. Госсет также работает с Р. А. Фишером.
- Короче говоря, Уильям Госсет родился в 1876 г. и умер в 1937 г. Он проводил математическое исследование для назревания пива, но у него возникла проблема, связанная с тем, что он работал только с малыми выборками. Он работал над концепцией вероятных средних ошибок. Он также изучал проблему вероятной ошибки коэффициента корреляции.