Некоторые мысли по поводу лабораторных работ

проф. Марк А. Зеликман
Кафедра экспериментальной физики, Санкт-Петербургский Политехнический Университет Петра Великого

Я давно собирался вынести на обсуждение коллег некоторые вопросы по лабораторным работам. На сайте кафедры есть брошюра П.А. Родного «Методы обработки результатов измерений. Дополнения и пояснения» (кстати, советую всем прочитать эту брошюру). Считаю нужным еще раз поговорить о некоторых моментах.

I.

На первое место, вне очереди, поставлю вопрос, в котором, на мой взгляд, чаще всего преподавателями допускается ошибка.

При приеме отчета по 1-ой работе «Исследование случайной величины» многие преподаватели требуют, чтобы полученное распределение было похоже на гауссовское, и считают неправильным, если полученная гистограмма отличается от нормального распределения. Это неверно. Какие у нас основания ожидать нормальности полученного распределения? Да, оно часто встречается на опыте. Да, оно удовлетворяет центральной предельной теореме, утверждающей, что сумма большого количества слабо зависимых случайных величин имеет распределение, близкое к нормальному. Проще говоря, если случайность определяется большим количеством факторов, близких по своему значению, то закон распределения будет нормальный. Следствием этого и является распространенность нормального распределения в природе. Но в данной работе совершенно иная ситуация. Какие факторы влияют на значения измеренного промежутка времени? Только один - реакция человека, нажимающего кнопку. Один нажимает, только когда слышит сигнал, другой - загодя, третий - то так, то так. Других равноважных факторов нет. Так что ситуация под центральную предельную теорему не подпадает, и распределение не обязано быть гауссовским. Оно скорее говорит о свойствах характера студента. Но некоторые преподаватели требуют, чтобы экспериментальная гистограмма совпадала с гауссовской, и заставляют студентов подгонять ее нужным образом. Этого нельзя делать, так как студентам будет казаться, что физика - это такая наука, где что надо, то и получишь.

II.

Я бы выделил 3 главные цели лабораторного практикума для студентов:
1. поработать с приборами
2. научиться писать научный отчет о работе
3. разобраться с погрешностями, округлениями и понятиями теории погрешностей.

Первый пункт ясен, о нем я говорить не буду.

По 2-му приведу такой показательный пример. При приеме отчетов по 1-й работе из механики («Исследование случайной величины») я спрашиваю студентов, какую случайную величину они исследуют. Обычно отвечают, что это длительности интервала времени 5 секунд. Я спрашиваю: если это интервал в 5 секунд, как он может равняться, например, $4,92$ секунды. Они отвечают обычно, что это в связи с неточностью измерения. Так у вас неточный хронометр? Нет, точный. Тогда в чем же дело? После долгих обсуждений приходим к верному ответу, который, на мой взгляд, звучит так: в качестве случайной величины были выбраны точно измеренные значения промежутка времени, грубо оцениваемого нами по ручным часам как 5 секунд.

Ситуация похожа на такую. Человек собрал на пляже камешки массой, на его взгляд, 50 граммов, а потом пришел домой и взвесил их на достаточно точных весах. Получился набор значений случайной величины, который можно анализировать. Эти значения - не массы 50−граммовых камней, т.к. они были бы все по 50, а точно измеренные массы камней, грубо оцененных как 50−граммовые.

Можно сказать, что это мелочи, но я так не считаю. Мы учим точным наукам. Все утверждения должны быть логичными, а введенные понятия должны быть правильно понятыми.

III.

Но самым важным пунктом из трех я считаю последний. Именно в лаборатории студенты могут получить базовые знания по теории погрешностей, которая так важна в инженерии, по теории вероятностей, с которой многие из них больше нигде не сталкиваются.

Приведу примеры. Студент измеряет диаметр стального стержня. Для увеличения точности он делает это несколько раз, например 7. Далее он ищет среднее, т.е. складывает все полученные значения и делит на 7. Калькулятор, скорее всего, даст какое-то восьмизначное число. Вопрос: сколько цифр надо записать в этом числе, т.е. в каком знаке округлить? Это серьезный вопрос, мало кто из студентов может на него обоснованно ответить. Некоторые скажут до сотых, другие - до 3-го знака, третьи выпишут все 8 цифр. Да-да, это очень часто. Они не знают, где округлять, поэтому, на всякий случай, пишут все, что видят на калькуляторе.

Считаю необходимым обратить внимание на этот момент. Я даже пример им такой привожу. Приходит муж пьяный домой, жена спрашивает: "Сколько выпил?", а он отвечает: "Вот я тут на калькуляторе посчитал – пол-литра на троих, по $166,66666$ на каждого". Жена говорит: "Откуда такая точность? На стенках осталось, в стаканах осталось, да и Серега неровно разливает, да и прольет еще, потому как руки дрожат. Вот и получается, наверно, грамм по 150 на человека". А муж отвечает: "Э нет, мы сегодня в аптеке выпивали, а разливали по весам и с пипеткой, так что за 166 я ручаюсь".

Иначе говоря, округлять надо с учетом погрешности измерения. Тут можно им объяснить и про то, что число знаков после запятой не показатель, нельзя раз и навсегда постановить округлять до 2-го знака после запятой. Ведь еще и единицы измерения вмешиваются. $0,254$ метра - это $25,4$ *см* или 254 *мм* и т.д. Так что положение запятой - вещь относительная.

Допустим, в результате измерений, усреднений и т.д. получили, что $f = (253,4567 \pm 2,314)$. Как округлять? Тут следует еще такой момент пояснить. Как, мол, вы, друзья-студенты, эту строку словами скажете? Они говорят чаще всего: величина f равна $253,4567$ с погрешностью $2,314$. И что это означает? Как это учесть? А если изменить доверительную вероятность, то и погрешность изменится, а как это трактовать? Да и где здесь о доверительной вероятности что-нибудь сказано? И вот привожу я их к мысли, что правильно говорить, что f лежит в интервале от... и до.... А вы это гарантируете? Вот так добираемся до доверительной вероятности, от которой, конечно, зависит и интервал.

Тогда становится понятным, что $253,4567$ - это не точное значение f, а центр интервала, а $2,314$ - не какая-то абстрактная погрешность, а полуширина интервала (которую мы называем абсолютной погрешностью).

Следующий вопрос об округлении. Итак, f с доверительной вероятностью лежит в указанном диапазоне. А если мы $2,314$ округлим до $2,31$, т.е. уменьшим интервал на $\frac{4}{2314}$, т.е. на $0,17\%$, это скажется на чем-то. Нет, не скажется. Интервал изменился мало, да это и не жесткие границы, в них f находится только с некоторой доверительной вероятностью. Ага, стало $2,31$. А если еще на один знак округлим, т.е. на $\frac{1}{231}$ уменьшим, скажется? Нет, не скажется. Стало $2,3$. Ну, а до 2 можем округлить, т.е. уменьшить интервал на $\frac{3}{23}$? Нет, тут уже заметно будет, это 13%, много. Значит, правило такое: округляем «погрешность» до двух значащих цифр. А значащие цифры - это не положение после запятой, это цифры от первого «ненуля» до последнего. Например, в числе $0,02053$ имеем 4 значащих цифры (от 2 до 3).

После округления погрешности до двух значащих цифр нет смысла задавать центр интервала с большей точностью, так что центральное значение тоже можно округлить, но не до двух значащих цифр, а до порядка погрешности. Например, в рассмотренном примере правильной будет запись $f = (253,5 \pm 2,3)$.

IV.

Считаю важным разъяснить различие между понятиями среднее квадратичное отклонение и среднее квадратичное отклонение среднего. Они обозначаются по-разному, задаются разными формулами и используются в разных случаях. Но мало кто задумывается над их смыслом и областью приложения.

Представим себе, что перед нами стоит задача экспериментального нахождения заряда электрона из какого-то опыта. Опытов конечное количество N. Спрашивается, как близко будет находиться найденное среднее от истинного значения? Представим мысленно, что может быть проведено очень много опытов, например, миллион. Среднее от всех опытов даст точное значение заряда электрона. Пусть этот миллион опытов распределен по большому числу лабораторий (наша лаборатория среди них), каждая из которых проводит N опытов, по результатам которых находит среднее $\bar{x}$ и σ. Как далеко будет полученное нами среднее $\bar{x}$ от истинного значения? Средние разных лабораторий представляют собой случайную величину, распределенную по нормальному закону. Хорошо известно, что в этом случае среднее значение всех средних равно среднему значению по всему миллиону, т.е. истинному, а среднее квадратичное отклонение среднего $\sigma_{\bar{x}}$ в $\sqrt{N}$ раз меньше среднего квадратичного отклонения (разброса) значений x, т.е. $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} = \sqrt{\frac{\Sigma(x_i-\bar{x})^2}{N(N-1)}}$. После этого остается умножить $\sigma_{\bar{x}}$ на коэффициент Стьюдента и мы найдем абсолютную погрешность, т.е. интервал в котором с выбранной доверительной вероятностью находится искомое значение заряда электрона.

Таким образом, при ИЗМЕРЕНИИ какой-то величины, т.е. поиске ее истинного значения, для вычисления погрешности надо пользоваться $\sigma_{\bar{x}}$.

Но при АНАЛИЗЕ истинно случайной величины, т.е. величины, для которой случайность является ее неотъемлемым свойством, среднее квадратичное отклонение вычисляется по формуле для σ. Например, надо узнать срок службы электрической лампочки. Срок службы конкретной лампочки узнать нельзя, так как для этого надо дождаться, когда она перегорит. Поэтому нужно провести серию опытов с лампами из той же партии и определить средний срок жизни лампочки и среднюю квадратичную погрешность этой величины. Этот срок службы не имеет точного значения, т.е. задача не является задачей ИЗМЕРЕНИЯ. Это АНАЛИЗ случайной величины. В этом случае среднее квадратичное отклонение вычисляется по формуле $\sigma = \sqrt{\frac{\Sigma(x_i-\bar{x})^2}{(N-1)}}$.

Итак, в задачах ИЗМЕРЕНИЯ и АНАЛИЗА следует использовать различные формулы. В наших лабораторных мы практически везде имеем дело с измерением. Исключением является работа №1, посвященная анализу случайной величины. Именно поэтому в ней используется σ. Использование же во 2-ой части этой работы $\sigma_{\bar{x}}$ является неправильным и бессмысленным. Поэтому, на мой взгляд, эту часть надо исключить.