

Нижегородский Государственный Технический университет имени Р.Е. Алексеева

Кафедра ФТОС

Статистическая обработка результатов измерений в лабораторном практикуме

Попов Е.А., Успенская Г.И.

Нижний Новгород
2015 год

Введение

Любой **физический объект** изучения характеризуется набором **физических величин**, отражающих его свойства.

Измерить какую-либо физическую величину – это значит сравнить ее с величиной, принятой за эталон.

Все измерения делятся на **прямые** и **косвенные**.

Прямые измерения – это непосредственные измерения, производимые с помощью приборов.

Косвенные измерения – расчет по формулам, в которые входят непосредственно измеренные величины и табличные значения.

Измеряя какую-либо физическую величину, мы не рассчитываем получить ее истинные значения, поэтому необходимо указать, насколько результат близок к истинному значению, т.е. указать **точность измерения**.

Для этого вместе с полученным результатом указывается приближенная **ошибка измерения**.

Пример: фокусное расстояние линзы $f=(256\pm 2)$ мм. Это означает, что фокусное расстояние лежит в пределах от 254 до 258 мм (имеется некоторая вероятность, что f лежит в данном интервале).

Ошибки (погрешности) измерений делятся на **систематические** и **случайные**.

Систематические ошибки - остаются постоянными на протяжении всей серии измерений. Источник этих ошибок – погрешности измерительных приборов.

Случайные ошибки - всегда присутствуют в эксперименте и служат причиной разброса измерений относительно истинного значения.

Случайные ошибки можно обнаружить путем повторных измерений. Увеличивая число измерений и находя среднеарифметическое значение, мы будем получать величину все ближе к ее истинному значению.

К случайным ошибкам можно отнести и **промахи**, которые резко отличаются от остальных измеряемых величин.

Их следует исключить либо в процессе эксперимента, либо при проведении расчетов.

Случайные ошибки выявляются методами математической статистики. Систематические ошибки не поддаются подобному анализу и их требуется устранять проверкой аппаратуры.

Результаты серии измерений одной величины можно наглядно представить, построив **гистограмму** (рис. 1) – диаграмму, которая показывает, как часто выпадают те или иные значения. Для ее построения разбиваем весь диапазон измерений на равные интервалы и считаем, сколько раз наша измеряемая величина попадает в каждый интервал. На каждом интервале строится прямоугольник с основанием, равным ширине интервала, а высотой равной числу попаданий в этот интервал.

Пример: Требуется определить основной возраст поступающих в НГТУ на примере одной группы из 25 человек. Возрастной диапазон (16-18 лет) делим на интервалы по 4 месяца и считаем число попаданий в каждый интервал. Строим гистограмму.

<i>Возрастные интервалы</i>	<i>Число попаданий (n)</i>
16 лет – 16 лет 4 месяца	2
16 лет 4 мес. - 16 лет 8 мес.	6
16 лет 8 мес. – 17 лет	10
17 лет – 17 лет 4 мес.	4
17 лет 4 мес. - 17 лет 8 мес.	2
17 лет 8 мес. – 18 лет	1

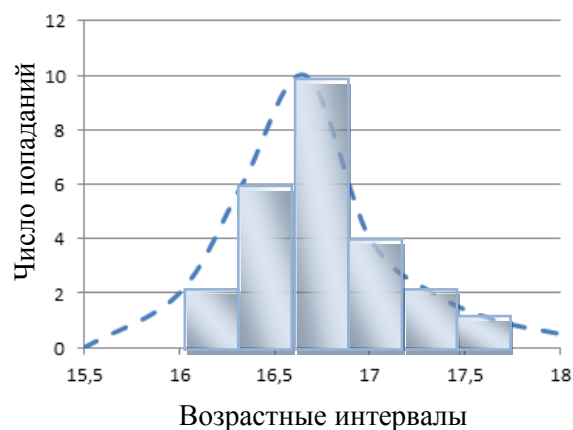


Рис. 1

При увеличении числа измерений ($n \rightarrow \infty$), например, при переходе к рассмотрению учащихся всего института, ширина интервала значительно сузится. Вместо $\Delta x = x_i - x_{\text{ист}}$ гистограммы логично построить уже график зависимости доли полного числа попаданий от интервала. Получится гладкая кривая – **кривая**

распределения (рис.1), максимум которой приходится на среднее арифметическое значение из всех измерений возраста

Исследование параметров подобных кривых распределения измеряемых величин и составляет основу обработки результатов измерений при проведении лабораторного эксперимента.

Пусть мы измеряем некоторую величину (x), (n) раз одним и тем же измерительным прибором.

В математической статистике совокупность всех результатов измерения $x_1, x_2, x_3 \dots x_n$ называется **генеральной совокупностью**, которая может быть сколько угодно велика. Конечное число (n) измерений называется **случайной выборкой** из генеральной совокупности.

Отклонение отдельного результата измерения Δx_i от истинного значения $X_{\text{ист}}$ называется **погрешностью отдельного измерения**:

$$\Delta x = x_i - X_{\text{ист}}$$

Опыт показывает, что для большого числа измерений ($n \rightarrow \infty$), положительные и отрицательные отклонения равновероятны и ближе всего к истинному значению $X_{\text{ист}}$ подойдет среднеарифметическое $\langle x_n \rangle$:

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

Установлено, что малое отклонение x_i от $X_{\text{ист}}$ встречается чаще, чем сильное, что соответствует **нормальному закону распределения Гаусса** для измеряемых величин. Зависимость плотности вероятности распределения результатов измерений $P(x)$ от значений измеряемой величины x изображена на рис. 2.

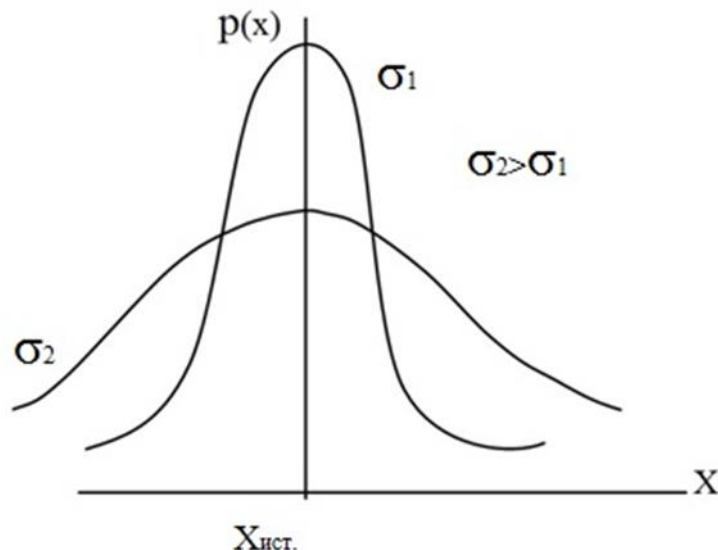


Рис. 2

Вводится понятие дисперсии σ^2 , характеризующее разброс результатов относительно истинного значения $X_{\text{ист}}$

$$\sigma = \lim_{n \rightarrow \infty} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - X_{\text{ист}})^2}$$

На практике истинное значение $X_{\text{ист}}$ неизвестно, поэтому можно найти приближенное значение S_n при замене $X_{\text{ист}}$ на среднеарифметическое $\langle x_n \rangle$

$$S_n = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \langle x_n \rangle)^2}$$

S_n называется **выборочным среднеквадратичным отклонением**, а число измерений n – выборка из генеральной совокупности. Средние значения для различных выборок одинакового объема n отличаются друг от друга и от $X_{\text{ист}}$

Тогда выборочное среднеквадратичное отклонение определяется

$$S_{\langle x_n \rangle} = \frac{S_n}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \langle x_n \rangle)^2}$$

это характеристика разброса средних значений $\langle x_n \rangle$ относительно $X_{\text{ист}}$ (в пределах ряда выборок).

При конечном числе измерений распределение подчиняется закону распределения Стьюдента¹, а абсолютная погрешность рассчитывается по формуле:

$$\Delta x = S_{<x_n>} \cdot t_{p,n},$$

где $S_{<x_n>}$ - выборочное среднеквадратичное отклонение, а $t_{p,n}$ – коэффициент Стьюдента, учитывающий конечное число измерений n и доверительную вероятность p , показывающую что $<x_n>$ отличается от $X_{\text{ист}}$ на величину не большую чем Δx .

Пример: $p=0,9$ означает, что из 100 измерений величины x_i она 90 раз окажется в интервале $(x \pm \Delta x)$, а 10 значений выйдут из этого интервала.

- $p=0,9$ (90%) используется в обычных измерениях
- $p=0,95$ (95%) используется при точных измерениях
- $p=0,99$ (99%) используется при особо ответственных измерениях

¹ Распределение Стьюдента введено в 1908 г. В.С. Госсетом, носящим псевдоним «Стьюдент»

Таблица коэффициентов Стьюдента для разных значений
выборок из генеральной совокупности.

$n \backslash p, \%$	90	95
2	6,31	12,71
3	2,92	4,30
4	2,35	3,18
5	2,13	2,78
6	2,02	2,57
7	1,94	2,45
8	1,89	2,36
9	1,86	2,31
10	1,83	2,26
11	1,81	2,23
12	1,80	2,20
13	1,78	2,18
14	1,77	2,16
15	1,76	2,14
16	1,75	2,13
17	1,75	2,12
18	1,74	2,11
19	1,73	2,10
20	1,73	2,09
60	1,67	2,00
120	1,66	1,98
∞	1,65	1,96

Алгоритм определения погрешности прямых измерений

1. Определяется выборка из n измеряемых величин

Пусть $n = 10$: $a_1, a_2, a_3, \dots, a_{10}$

2. Производится оценка погрешности приборов и табличных значений

$\delta x_{\text{пр}}$	по паспорту прибора	
	по шкале прибора	$\frac{1}{2}$ цены наименьшего деления
		линейка – 0,5 см
		термометр 0,1 °С, – 0,05 °С
		$\frac{1}{2}$ массы наименьшей гири
	по классу точности (электроизмерительные приборы)	

3. Вычисляется среднеарифметическое измеряемой величины

$$\langle a \rangle = \frac{a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9 + a_{10}}{10}$$

4. Определяется отклонение каждого измеренного значения от среднего

$$\begin{aligned} \Delta a_1 &= |a_1 - \langle a \rangle| \\ \Delta a_2 &= |a_2 - \langle a \rangle| \\ &\vdots \\ \Delta a_{10} &= |a_{10} - \langle a \rangle| \end{aligned}$$

5. Определяется среднеквадратичное отклонение отдельных измерений

$$S_n = \sqrt{\frac{\Delta a_1^2 + \Delta a_2^2 + \dots + \Delta a_{10}^2}{10}}$$

Результат считается промахом, если $|x_i - \langle x \rangle| \geq \alpha S_n$, где коэффициент α определяется из таблицы по размеру выборки.

n	3	4	5	6	7	8	9	10	11	12	15	20	30	100
α	1,41	1,69	1,87	2,00	2,09	2,17	2,24	2,29	2,34	2,39	2,49	2,62	2,80	3,20

$$\begin{aligned} |a_1 - \langle a \rangle| &\geq \alpha S_N \\ |a_2 - \langle a \rangle| &\geq \alpha S_N \\ &\vdots \\ |a_{10} - \langle a \rangle| &\geq \alpha S_N \end{aligned}$$

7. После исключения промахов снова вычисляется среднеарифметическое с измененным n' .

Пусть было 2 промаха - a_2 и a_5 , тогда $n' = 10 - 2 = 8$

$$\langle a \rangle = \frac{a_1 + a_3 + a_4 + a_6 + a_7 + a_8 + a_9 + a_{10}}{8}$$

$$S_{\langle a_{n'} \rangle} = \sqrt{\frac{\Delta a_1^2 + \Delta a_3^2 + \Delta a_4^2 + \Delta a_6^2 + \Delta a_7^2 + \Delta a_8^2 + \Delta a_9^2 + \Delta a_{10}^2}{7 \cdot 8}}$$

8. Задав доверительную вероятность, по таблице определяется коэффициент Стьюдента, а затем вычисляется **абсолютная погрешность результатов измерения**.

$$\Delta x = \sqrt{(t_{p,N} \cdot S_{\langle x_n \rangle})^2 + \left(\frac{1}{3} t_{p,\infty} \cdot \delta x_{\text{пр}}\right)^2}$$

В предельных случаях:

- если $\frac{\delta x_{\text{пр}}}{S_{\langle a_N \rangle}} < 0.8$, то пренебрегают приборной погрешностью $\delta x_{\text{пр}}$:

$$\Delta a = t_{p,N} \cdot S_{\langle a_N \rangle};$$

- если $\frac{\delta x_{\text{пр}}}{S_{\langle a_N \rangle}} > 8$, то $\Delta a = \delta x_{\text{пр}}$.

9. Рассчитывается относительная погрешность (в %)

$$\varepsilon = \frac{\Delta x}{\langle x \rangle} \cdot 100\%.$$

10. Записывается окончательный результат $x = \langle x \rangle \pm \Delta x$ с округлением по следующим правилам:

- сначала округляется погрешность до первой значащей цифры в сторону увеличения.
- затем округляется сама величина до совпадения ее десятичного разряда с погрешностью.

Пример: до округления: $F = 923,46 \pm 0,253$ Н

после округления: $F = 923,5 \pm 0,3$ Н

Для физических величин, взятых из справочников, погрешность принимается равной половине единицы последнего десятичного разряда.

Пример: $\pi = 3,14 \pm 0,005$

$$g=9,81\pm0,005 \text{ м/с}^2$$

Точность измеряемых физических величин **зависит** не от количества значащих цифр, а от количества верных значащих цифр, **определяемых погрешностью**.

Алгоритм определения погрешности косвенных измерений.

При косвенных измерениях результаты получаются путем вычисления по формулам, в которые входят:

- результаты прямых измерений
- табличные величины
- справочные данные и т.д.

определенные с конечной точностью (см. выше). Пусть есть функция $\Phi(a, b, c)$, зависящая от непосредственно измеряемых независимых величин (a, b, c) , справочных и табличных данных.

Пользуясь конкретным видом функции $\Phi(a, b, c)$ и имеющимся значениями

$$\begin{aligned} a &= \langle a \rangle \pm \Delta a, \\ b &= \langle b \rangle \pm \Delta b, \\ c &= \langle c \rangle \pm \Delta c, \end{aligned}$$

необходимо оценить среднее значение этой функции $\langle \Phi \rangle$ и величину погрешности $\Delta \Phi$, характеризующую разброс значений $\Phi_i = \Phi(a_i, b_i, c_i)$ относительно $\langle \Phi \rangle$.

Для определения этих оценок удобно использовать два тождественных по результатам метода.

I Метод выборок.

Он применяется в тех случаях, когда значения a_i, b_i, c_i определяются **из различных опытов**. Статистическое **усреднение** отдельных величин a, b, c в этом случае **не имеет смысла**.

Пример: лабораторная работа по определению $\frac{c_p}{c_v}$ воздуха методом

Клемана – Дезорма и лабораторная работа по определению модуля Юнга.

В этом случае вычисление погрешностей осуществляется по алгоритму аналогичному оценке погрешностей прямых измерений:

- 1) Определяется $\Phi_i(a_i, b_i, c_i)$ для каждого опыта
- 2) Находится среднее значение $\langle \Phi \rangle = \frac{1}{n} \sum_{i=1}^n \Phi_i$
- 3) Вычисляются промахи и среднеквадратичное отклонение

$$S_{\langle \Phi_n \rangle} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^{n'} (\Phi_i - \langle \Phi \rangle)^2}$$

- 4) Определяются абсолютная и относительная погрешности

$$\Delta \Phi = t_{p,n} \cdot S_{\langle \Phi_n \rangle},$$

$$\varepsilon_{\Phi} = \frac{\Delta \Phi}{\langle \Phi \rangle} \cdot 100\%.$$

II Метод, основанный на применении формулы полного дифференциала.

Этот метод используется в тех случаях, когда величины a_i, b_i, c_i определяются из одного эксперимента и могут быть статистически обработаны до подстановки в формулу $\Phi_i = \Phi(a_i, b_i, c_i)$.

Тогда среднее значение вычисляется по формуле $\langle \Phi \rangle = \Phi(\langle a \rangle, \langle b \rangle, \langle c \rangle)$

Используется два варианта расчета.

1 вариант используется, если функция $\Phi(a, b, c)$ – многочлен. В этом случае на основании формулы полного дифференциала для $\Phi(a, b, c)$

$$\Phi = \frac{\partial \Phi}{\partial a} da + \frac{\partial \Phi}{\partial b} db + \frac{\partial \Phi}{\partial c} dc.$$

С учетом квадратичного усреднения можно получить для **абсолютной погрешности** следующее выражение:

$$\Delta \Phi = \sqrt{\left(\frac{\partial \Phi}{\partial a} \Delta a\right)^2 + \left(\frac{\partial \Phi}{\partial b} \Delta b\right)^2 + \left(\frac{\partial \Phi}{\partial c} \Delta c\right)^2},$$

где Δa , Δb , Δc – погрешности прямых измерений, а численные значения производных берутся для средних значений $\langle a \rangle, \langle b \rangle, \langle c \rangle$. Затем вычисляется относительная погрешность:

$$\varepsilon_{\Phi} = \frac{\Delta \Phi}{\langle \Phi \rangle} \cdot 100\%$$

Пример расчета погрешности по первому варианту:

Формула $\Phi(a, b, c, f)$ – многочлен

$$\Phi(a, b, c, f) = a^3 - \cos b^2 + \sqrt{c - f^2}$$

1) Находятся частные производные:

$$\frac{\partial \Phi}{\partial a} = 3a^2$$

$$\frac{\partial \Phi}{\partial b} = 2b \sin b^2$$

$$\frac{\partial \Phi}{\partial c} = -\frac{1}{2\sqrt{c - f^2}}$$

$$\frac{\partial \Phi}{\partial f} = -\frac{f}{\sqrt{c - f^2}}$$

2) Вычисляется абсолютная погрешность

$$\begin{aligned} \Delta \Phi &= \sqrt{(3a^2 \Delta a)^2 + (2b \sin b^2 \Delta b)^2 + \left(-\frac{1}{2\sqrt{c - f^2}} \Delta c\right)^2 + \left(-\frac{f}{\sqrt{c - f^2}} \Delta f\right)^2} \\ &= \sqrt{9a^4 \Delta a^2 + 4b^2 (\sin b^2)^2 \Delta b^2 + \frac{\Delta c^2}{4(c - f^2)} + \frac{f \Delta f^2}{(c - f^2)}} \end{aligned}$$

3) Определяется относительная погрешность

$$\varepsilon_{\Phi} = \frac{\Delta \Phi}{\langle \Phi \rangle} \cdot 100\%,$$

$$\text{где } \langle \Phi \rangle = \langle a \rangle^3 - \cos \langle b \rangle^2 + \sqrt{\langle c \rangle - \langle f \rangle^2}$$

2 вариант используется расчета используется, если функция $\Phi(a, b, c)$ имеет логарифмируемый вид (т.е. независимые переменные входят как сомножители).

Если разделить абсолютную погрешность $\Delta \Phi$ на $\langle \Phi \rangle$, то можно сразу получить выражение для относительной погрешности ε_{Φ}

$$\varepsilon_{\Phi} = \frac{\Delta\Phi}{\langle \Phi \rangle} = \sqrt{\left(\frac{1}{\langle \Phi \rangle} \frac{\partial \Phi}{\partial a} \Delta a\right)^2 + \left(\frac{1}{\langle \Phi \rangle} \frac{\partial \Phi}{\partial b} \Delta b\right)^2 + \left(\frac{1}{\langle \Phi \rangle} \frac{\partial \Phi}{\partial c} \Delta c\right)^2}.$$

Учитывая, что $\frac{1}{\langle \Phi \rangle} \frac{\partial \Phi}{\partial a} = \frac{\partial \ln \Phi}{\partial a}$, $\langle \Phi \rangle = \Phi(\langle a \rangle, \langle b \rangle, \langle c \rangle)$,

$$\varepsilon_{\Phi} = \sqrt{\left(\frac{\partial \ln \Phi}{\partial a}\right)^2 \Delta a^2 + \left(\frac{\partial \ln \Phi}{\partial b}\right)^2 \Delta b^2 + \left(\frac{\partial \ln \Phi}{\partial c}\right)^2 \Delta c^2}.$$

Затем можно получить формулу абсолютной погрешности

$$\Delta\Phi = \varepsilon_{\Phi} \langle \Phi \rangle.$$

Пример расчета погрешности косвенных измерений 2-м способом:

Формула $\Phi(a, b, c, f)$ – одночлен

$$\Phi(a, b, c, f) = \frac{a^3 \cos b}{\sqrt{c - f^2}}$$

1) Осуществляется логарифмирование формулы для разделения переменных:

$$\ln \Phi(a, b, c, f) = 3 \ln a + \ln \cos b - \frac{1}{2} \ln(c - f^2)$$

Здесь

$$\begin{aligned} a &= \langle a \rangle \pm \Delta a \\ b &= \langle b \rangle \pm \Delta b \\ c &= \langle c \rangle \pm \Delta c \\ f &= \langle f \rangle \pm \Delta f \end{aligned}$$

2) Находятся частные производные $\ln \Phi$:

$$\begin{aligned} \frac{\partial \ln \Phi}{\partial a} &= \frac{3}{a} \\ \frac{\partial \ln \Phi}{\partial b} &= \frac{-\sin b}{\cos b} = -\operatorname{tg} b \\ \frac{\partial \ln \Phi}{\partial c} &= -\frac{1}{2(c - f^2)} \\ \frac{\partial \ln \Phi}{\partial f} &= \frac{f}{(c - f^2)} \end{aligned}$$

При численном вычислении частных производных используются средние значения $\langle a \rangle, \langle b \rangle, \langle c \rangle, \langle f \rangle$.

3) Определяются относительная и абсолютная погрешности

$$\varepsilon_{\Phi} = \frac{\Delta\Phi}{\langle \Phi \rangle} = \sqrt{\frac{9}{a^2} \Delta a^2 + (\operatorname{tg} b)^2 \Delta b^2 + \frac{\Delta c^2}{4(c - f^2)^2} + \frac{f^2 \Delta f^2}{(c - f^2)^2}}$$

$$\Delta\Phi = \varepsilon_{\Phi} \langle \Phi \rangle$$

Построение графиков и графические методы обработки результатов

В экспериментальной физике графиками пользуются:

- а) для **наглядности**, когда таблицы не выявляют характерных особенностей изучаемых процессов и требуется сравнить полученные результаты с теорией;
- б) для **графической интерполяции**, т.е. определения промежуточных значений аргумента, при которых измерения не проводились;
- в) для **графической экстраполяции**, когда значение искомой величины, меняющейся по линейному закону, лежит за пределами измерений.

Для построения графиков используется линейный масштаб (для функции вида: $y=kx+b$), полулогарифмический ($y=a^x$) и логарифмический (для $y=x^p$).

1) Масштаб должен быть простым, т.е. одно деление шкалы должно соответствовать 1, 2, 5, 10, 10^2 , 10^{-1} , 10^{-2} и т.д. единиц измеряемой величины.

2) Масштаб выбирается так, чтобы точки не сливались друг с другом.

3) График должен занимать все координатное поле. На осях координат указывается название или символ величины, а десятичный множитель относить к единице измерения. Деления на графике обозначаются цифрами 1, 2, 3...10, 20, 30 и т.д. по оси абсцисс откладывается независимая переменная – аргумент, а по оси ординат – зависимая переменная – функция.

4) Изображение экспериментальных результатов зависит от определения абсолютных погрешностей измерения. Если они неизвестны, то результаты изображаются точками, а если известны, то погрешности указываются в виде отрезков, полуразмеры которых соответствуют величинам абсолютных погрешностей.

Через экспериментальные значения проводятся кривые так, чтобы возможно больше точек легло на эту линию, а остальные распределились равномерно выше и ниже ее. Если какая-либо точка выпадает из графика – ее отбрасывают, а измерения проверяют на промах.

Недопустимо наносить на координатные оси точки, отражающие значения самих измеряемых величин!

Библиографический список

1. Голубев В.И., Методические указания по статистической обработке результатов измерений в лабораториях физического практикума / В.И. Голубев, Г.Д. Павлова, Г.И. Успенская, Т.Ф. Фёдорова - Нижний Новгород: НГПИ, 1991
2. Зейдель А.Н. Ошибки измерений физических величин / А.Н. Зейдель. - М.: Наука, 1974