2 ПОГРЕШНОСТИ ИЗМЕРЕНИЙ

Принято различать три типа погрешностей измерений: систематические, случайные и грубые [1].

2.1 Систематические погрешности

Систематические погрешности связаны с точностью измерительного прибора или метода измерения. То есть причина их появления проявляется при каждом измерении и величина их не меняется при измерениях. Например, если измерение производится масштабной линейкой, деления которой нанесены с ошибкой, то величина этой ошибки будет повторятся в каждом измерении.

Обычно систематическая погрешность измерения составляет одну десятую цены деления измерительного прибора. Систематические погрешности, природа и величина которых нам известна, устраняются введением поправок, которые устанавливаются при измерении эталонными или другими более точными средствами. Например, можно ввести поправки, учитывающие температурные удлинения измерительной линейки.

2.2 Случайные погрешности и их числовые оценки

Случайные погрешности обусловлены большим числом причин (факторов), о существовании и влиянии которых нам заранее неизвестно, и влияние каждого из них на результат измерения незначительно (не является определяющим), но в совокупности их влияние при повторении измерения приводит к различным результатам.

Источником этих погрешностей могут служить изменяющиеся условия внешней среды, ошибки отсчета на измерительной шкале прибора и другие.

Рассмотрим n прямых измерений одной и той же величины a. Результаты этих измерений x будут отличаться от измеряемой величины a на величину ошибки измерения Δ.

Оценкой математического ожидания (генеральной средней) и, следовательно, истинного значения измеряемой величины a является среднее арифметическое значение выборки – результатов измерений. Таким образом, если $x_1, x_2, x_3, \ldots, x_n$ результаты n измерений одной и той же величины, то истинное значение a может быть приближенно оценено выборочной средней арифметической величиной (средним значением выборки)

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i x_i}{n},$$

где n_i – частота появления одного и того же результата измерения x_i;

x_i – результат i -го независимого измерения;

k – число групп, имеющих одинаковые результаты независимых измерений;

$n = \sum_{i=1}^{k} n_i$ – объем выборки, то есть число независимых измерений.

Величина рассеивания результатов измерения может быть оценена дисперсией выборки (выборочной дисперсией):

$$D_x = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{k} n_i x_i^2}{n} - (\bar{x})^2 = \overline{x^2} - (\bar{x})^2,$$

где $\overline{x^2} = \frac{\sum_{i=1}^{k} n_i x_i^2}{n}$.

Учитывая, что выборочная средняя арифметическая величина сама является случайной величиной, то ее рассеивание можно оценить также выборочной дисперсией среднего арифметического:

$$D_{\bar{x}} = \frac{D_x}{n}.$$

Соответственно, среднее квадратическое отклонение среднего арифметического (стандартная ошибка):

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}},$$

где σ_x – среднее квадратическое отклонение, равное $\sigma_x = \sqrt{D_x}$.

Однако эти оценки являются смещенными, то есть при малых объемах наблюдений они дают заниженные значения. Поэтому для оценки истинной дисперсии используют исправленную дисперсию выборки:

$$S_x^2 = D_x \cdot \frac{n}{n-1}.$$

Другой характеристикой рассеивания результатов наблюдения является коэффициент вариации, представляющий собой отношение среднего квадратического отклонения к среднему арифметическому:

$$K_v = \frac{S_x}{\bar{x}}.$$

ŋqdqɐəɤ bɆ dədɆbɤзə d ɐɔɑbɔɐ Ɉqddɤqɐè

2.3 Интервальная оценка истинного значения измеряемой величины

Оценка истинного значения выборочным средним арифметическим является точечной, которая при малом объеме выборки может значительно отличаться от истинного значения, то есть приводить к большим ошибкам. Поэтому для оценки истинного значения измеряемой величины при малом числе измерений используют интервальные оценки, которые определяются двумя числами – концами интервала. Интервальная оценка устанавливает точность и надежность при оценивании результата измерения.

Если рассматривать результаты измерений как независимые нормально распределенные случайные величины с математическими ожиданиями равными измеряемой величине a, то среднее арифметическое результатов измерения будет представлять собой тоже нормально распределенную случайную величину с математическим ожиданием, равным измеряемой величине a и средним квадратическим отклонением $S_{\bar{x}}$.

Учитывая, что центрированная нормированная случайная величина

$$T = \frac{\bar{x} - a}{S_{\bar{x}}},$$

распределена по закону Стьюдента, то доверительный интервал может быть определен из уравнения

$$P\left(\bar{x} - t_{\gamma}\frac{S_x}{\sqrt{n}} < a < \bar{x} + t_{\gamma}\frac{S_x}{\sqrt{n}}\right) = \gamma,$$

где P – вероятность события, указанного в скобках;

γ – заданная надежность (вероятность);

t_{γ} – критическое значение случайной величины T, распределенной по закону Стьюдента.

Величина t_{γ} берётся из таблицы для заданной надежности γ и числа измерений n (или числа степеней свободы $k = n - 1$).

Критическое значение критерия Стьюдента можно получить с помощью функции Листа книги Excel.

$$t_{\gamma} = СТЬЮДРАСПОБР\,(вероятность; степени\ \ свободы)$$

Здесь вероятность $\alpha = 1 - \gamma$ – вероятность, соответствующая двустороннему распределению Стьюдента, а степени свободы – число степеней свободы, характеризующее распределение.

Таким образом, можно с надежностью γ утверждать, что истинное значение измеряемой величины a не выйдет за пределы интервала, указанного в скобках, то есть:

$$a = \overline{x} \pm t_\gamma \frac{S_x}{\sqrt{n}}.$$

Величина $\Delta = t_\gamma \frac{S_x}{\sqrt{n}}$, является предельной ошибкой при оценке измеряемой величины a с надежностью γ. Из этого следует, что для того, чтобы случайная ошибка при измерении величины a с заданной надежностью γ не превысила заданную ошибку Δ, необходимо произвести число измерений

$$n \geq \left(\frac{t_\gamma S_x}{\Delta} \right)^2 .$$

Если число наблюдений достаточно велико (более 30), то исправленная дисперсия выборки равна неисправленной, то есть $S_x^2 \approx D_x = \sigma_x^2$. Предельная ошибка для оценки генеральной средней определяется из условия

$$P\left(\overline{x} - t \cdot \frac{\sigma_x}{\sqrt{n}} < a < \overline{x} + t \cdot \frac{\sigma_x}{\sqrt{n}} \right) = 2 \cdot \Phi(t) = \gamma,$$

где P – вероятность события, указанного в скобках;

γ – заданная надежность (вероятность);

t – критическое значение центрированной нормированной случайной величины $Z = \frac{x-a}{\sigma}$, распределенной по нормальному закону, у которой математическое ожидание равно нулю, а дисперсия единице;

$\Phi(t)$ – функция Лапласа, удвоенное значение которой для аргумента t равно γ.

Функция Лапласа имеет вид: $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{z^2}{2}} dz$.

Величина t берётся из таблицы функции Лапласа для надежности $\frac{\gamma}{2}$. Величину t можно получить с помощью функции Листа книги Excel

$$t = НОРМСТОБР\left(0{,}5 + \frac{\gamma}{2}\right).$$

Например, для гамма $\gamma = 0{,}95$ (надежности 0,95) $t = НОРМСТОБР\left(0{,}5 + \frac{\gamma}{2}\right) = НОРМСРОБР(0{,}975) = 1{,}95996 \approx 1{,}96$.

Предельная ошибка для оценки математического ожидания нормально распределенного признака по выборочной средней равна $\Delta = t \cdot \frac{\sigma_x}{\sqrt{n}}$. Следовательно, для того, чтобы случайная ошибка при измерении величины a с заданной надежностью γ не превысила заданную ошибку Δ, необходимо произвести число измерений $n \geq \left(\frac{t \cdot \sigma_x}{\Delta}\right)^2$.

Если считать допустимой десятипроцентную предельную относительную ошибку при оценке генеральной средней по выборочной средней $\delta_{\bar{x}} = \frac{\Delta}{\bar{x}} = 0{,}1$, то

$$n \geq \left(\frac{t \cdot \sigma_x}{\delta_{\bar{x}} \cdot \bar{x}}\right)^2 = \left(\frac{t \cdot k_x}{\delta_{\bar{x}}}\right)^2 = \left(\frac{t \cdot k_x}{0{,}1}\right)^2 .$$

Так, для надежности 95% можно приблизительно принять $t = 2$. Тогда при коэффициенте вариации $k_x = 0{,}1$ получим минимально необходимое число наблюдений, оно должно быть:

$$n \geq \left(\frac{2 \cdot k_x}{0{,}1}\right)^2 = \left(\frac{2 \cdot 0{,}1}{0{,}1}\right)^2 = 4 .$$

Соответственно, при x_i x_j и при $r_{kj} = \frac{k_{kj}}{\sigma_k \cdot \sigma_j}$

$$r_{x_k x_j} = \frac{\overline{x_{kj} \cdot x_j} - \overline{x_k} \cdot \overline{x_j}}{\sigma_k \cdot \sigma_j} .$$

Для более точной оценки $k_{kj} = \frac{\sum_{i=1}^{n}(x_{ki} - \overline{x_k}) \cdot (x_{ji} - \overline{x_j})}{n} = \overline{x_k \cdot x_j} - \overline{x_k} \cdot \overline{x_j}$ необходимое число наблюдений будет резко возрастать.