5. ОЦЕНИВАНИЕ ЧИСЛОВЫХ ХАРАКТЕРИСТИК И ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ОБЪЕКТОВ

Закон распределения в любой его форме является исчерпывающей характеристикой вероятностного поведения случайного объекта (величины, вектора, функции). Поскольку задача его определения достаточно сложна, на практике часто определяются лишь числовые характеристики, основными из которых являются математическое ожидание и дисперсия. Для случайных векторов характеристикой рассеяния является корреляционная матрица.

Следует отметить, что решение такой более упрощённой задачи имеет большую практическую ценность, так как во многих случаях знать закон распределения не требуется. Кроме того, часто на основе каких-то априорных предположений вид закона распределения исследуемого случайного объекта известен и для его полного описания достаточно определить лишь параметры.

Как уже отмечалось, задача оценивания параметров распределения, в частности числовых характеристик, сводится к отысканию таких статистик (функций случайной выборки), которые могут служить наилучшими в каком-то смысле оценками истинных значений параметров. Все рассматриваемые оценки должны проверяться на наличие у них основных необходимых свойств: несмещённости, состоятельности и эффективности.

5.1. Оценивание математического ожидания случайной величины

Пусть имеется случайная величина $\hat{x}$, математическое ожидание которой $M_{\hat{x}} = \bar{x}$ неизвестно. Над случайной величиной проведено n независимых опытов (наблюдений). По их результатам x_1, $x_2, \ldots, x_n$ требуется найти состоятельную, несмещённую и эффективную оценку $\widetilde{M}_{\hat{x}}$ параметра $M_{\hat{x}}$, т.е. найти функцию

$$\widetilde{M}_{\hat{x}} = M_{\hat{x}}(x_1, x_2, \ldots, x_n) = \widetilde{M}_{\hat{x}}(X_{<n>}).$$

В качестве оценки математического ожидания (среднего значения) случайной величины $\hat{x}$ могут приниматься различные характеристики случайной выборки. Все общие методы статистического оценивания дают в качестве наилучшей точечной оценки математического ожидания случайной величины её статистическое среднее. Однако при нахождении оценки математического ожидания, удовлетворяющей требованиям состоятельности, несмещённости и эффективности, следует различать равноточные и неравноточные наблюдения (однородные и неоднородные опыты).

5.1.1. Равноточные наблюдения

Статистическое (выборочное) среднее или ***статистическое математическое ожидание*** случайной величины находится по формуле

$$\tilde{M}_{\hat{x}} = \tilde{M}_{\hat{x}}(X_{<n>}) = M^*_{\hat{x}} = \bar{x}^* = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i\,. \qquad (5.1.1)$$

Поскольку наблюдения равноточны, то случайные величины $\hat{x}_1,\ \hat{x}_2,...,\hat{x}_n$ по существу представляют собой «экземпляры» одной и той же случайной величины $\hat{x}$ и, следовательно, имеют один и тот же закон распределения с числовыми характеристиками:

$$M_{\hat{x}_i} = M_{\hat{x}} = \bar{x}\,;\quad D_{\hat{x}_i} = D_{\hat{x}};\quad \sigma_{\hat{x}_i} = \sigma_{\hat{x}}\,,\quad i = \overline{1,n}\,.$$

Покажем, что оценка $M^*_{\hat{x}}$ удовлетворяет всем трём общим требованиям.

1. Из выражения (5.1.1) следует, что

$$M[M^*_{\hat{x}}] = M\left[\frac{1}{n}\sum_{i=1}^{n}\hat{x}_i\right] = \frac{1}{n}\sum_{i=1}^{n}M_{\hat{x}_i} = \frac{nM_{\hat{x}}}{n} = M_{\hat{x}}\,.$$

Таким образом, $M^*_{\hat{x}}$ является несмещённой оценкой параметра $M_{\hat{x}}$.

2. Согласно теореме Чебышева среднее арифметическое наблюдаемых значений случайной величины сходится по вероятности к её математическому ожиданию

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\hat{x}_i - M_{\hat{x}_i}\right| < \varepsilon\right) = 1.$$

Следовательно, статистическое среднее $M^*_{\hat{x}}$ есть состоятельная оценка параметра $M_{\hat{x}}$.

3. Согласно выражению (5.1.1) дисперсия статистического среднего

$$D[M^*_{\hat{x}}] = D\left[\frac{1}{n}\sum_{i=1}^{n}\hat{x}_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}D_{\hat{x}_i} = \frac{nD_{\hat{x}}}{n^2} = \frac{D_{\hat{x}}}{n} \qquad (5.1.2)$$

с ростом объёма n выборки неограниченно убывает и, следовательно, $M^*_{\hat{x}}$ асимптотически эффективная оценка $M_{\hat{x}}$. Доказано, что если случайная величина $\hat{x}$ подчинена нормальному закону распределения, то при любых n дисперсия (5.1.2) будет минимально возможной. В таком случае $M^*_{\hat{x}}$ является эффективной оценкой математического ожидания $M_{\hat{x}}$.

Следовательно, $M^*_{\hat{x}}$ – это подходящее значение $M_{\hat{x}}$:

$$M_{\hat{x}} \approx \widetilde{M}_{\hat{x}} = M_{\hat{x}}^* = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i \,. \qquad (5.1.3)$$

<u>П р и м е р 5.1</u>. В условиях примера 4.1 найти оценку математического ожидания случайной величины $\tilde{\tau}$.

▼ Согласно формуле (5.1.3)

$$\widetilde{M}_{\hat{\tau}} = M_{\hat{\tau}}^* = \frac{1}{12}\sum_{i=1}^{12} \tau_i = 127{,}6 \text{ ч.}$$

Случайная выборка, приведённая в табл.4.3, была получена при наблюдении случайной величины $\tilde{\tau}$, математическое ожидание которой $M_{\hat{\tau}} = \bar{\tau} = 100$ ч. Сравнительно невысокая точность полученной оценки обусловлена малым объёмом выборки, но ни в коей мере не способом её вычисления. ▲

Если объём n выборки достаточно велик, то вычисления по формуле (5.1.3) оказываются громоздкими. Задачу можно упростить, если использовать данные интервального вариационного ряда, т.е. полагать

$$\widetilde{\widetilde{M}}_{\hat{x}} = \sum_{l=1}^{r} \bar{x}_l P_l^* \,, \qquad (5.1.4)$$

где $\bar{x}_l = 0{,}5(x_l + x_{l+1})$ – представитель (середина) l-го разряда вариационного ряда; P_l^* – частота попадания вариантов x_i случайной величины $\hat{x}$ в l-й разряд.

Значение оценки $\widetilde{\widetilde{M}}_{\hat{x}}$, определяемое по формуле (5.1.4), оказывается приближённым, однако с ростом n (а, следовательно, и r) точность данной формулы возрастает.

<u>П р и м е р 5.2</u>. В условиях примера 4.2 найти приближённое значение оценки $\widetilde{\widetilde{M}}_{\hat{\tau}}$ математического ожидания $M_{\hat{\tau}}$ случайной величины $\hat{\tau}$.

▼ Используем данные табл.4.7. По формуле (5.1.4) получим

$$\widetilde{M}_{\hat{\tau}} \approx \widetilde{\widetilde{M}}_{\hat{\tau}} = \sum_{l=1}^{100} \bar{\tau}_l P_l^* = 102 \text{ ч.},$$

▲

Следует обратить внимание на то, что выборка, приведённая в табл.4.7, принадлежит той же генеральной совокупности, что и в табл.4.3. Однако, как видно из решений примеров 5.1 и 5.2, в последнем случае даже приближённое значение оценки $\widetilde{M}_{\hat{\tau}}$ меньше отличается от истинного значения параметра $M_{\hat{\tau}}$ = 100 ч. Это объясняется бóльшим объёмом n и, следовательно, большей информативностью выборки, приведённой в табл.4.7.

5.1.2. Неравноточные наблюдения

Пусть характеристики точности наблюдений от опыта к опыту изменяются так, что наблюдаемая в i-м опыте случайная величина $\hat{x}_i$ имеет дисперсию

$$D_{\hat{x}_i} = \sigma_i^2, \quad i = \overline{1,n}.$$

При этом среднее значение случайной величины $\hat{x}$ от опыта к опыту не изменяется:

$$M_{\hat{x}_i} = \bar{x} = M_{\hat{x}}, \quad i = \overline{1,n}.$$

В данном случае оценка $\tilde{M}_{\hat{x}}$ математического ожидания случайной величины $\hat{x}$ по-прежнему будет являться функцией случайной выборки:

$$\tilde{M}_{\hat{x}} = \tilde{M}_{\hat{x}}(X_{<n>}).$$

Необходимо так выбрать вид этой зависимости, чтобы оценка имела простое аналитическое выражение и была несмещённой, состоятельной и эффективной.

Так как наиболее простой функциональной зависимостью является линейная, то будем искать оценку $\tilde{M}_{\hat{x}}$ в классе линейных функций:

$$\tilde{M}_{\hat{x}} = \sum_{i=1}^{n} c_i x_i. \tag{5.1.5}$$

Очевидно, что теперь решение поставленной задачи состоит в отыскании значений коэффициентов c_i, $i = \overline{1,n}$ линейной формы (5.1.5), при которых оценка $\tilde{M}_{\hat{x}}$ будет удовлетворять всем трём указанным выше требованиям.

1. Чтобы оценка была несмещённой, должно выполняться равенство

$$M[\tilde{M}_{\hat{x}}] = M_{\hat{x}}.$$

Поскольку в этом случае

$$M[\tilde{M}_{\hat{x}}] = M\left[\sum_{i=1}^{n} c_i x_i\right] = \sum_{i=1}^{n} c_i M_{\hat{x}_i} = M_{\hat{x}} \sum_{i=1}^{n} c_i = M_{\hat{x}},$$

то коэффициенты c_i должны удовлетворять условию

$$\sum_{i=1}^{n} c_i = 1.$$

2. Для того чтобы оценка $\tilde{M}_{\hat{x}}$ была эффективной, её дисперсия

$$D[\tilde{M}_{\hat{x}}] = D\left[\sum_{i=1}^{n} c_i \hat{x}_i\right] = \sum_{i=1}^{n} c_i^2 D_{\hat{x}_i} = \sum_{i=1}^{n} c_i^2 D_i \tag{5.1.6}$$

должна быть минимальной при условии, что

$$1-\sum_{i=1}^{n} c_i = 0. \qquad (5.1.7)$$

Условный экстремум (минимум) функции (5.1.6) с переменными $c_1,\ c_2,\ldots,\ c_n$ отыскиваем методом неопределённых множителей Лагранжа. При этом учитываем, что должно выполняться равенство (5.1.7). Следовательно, исследуем на минимум вспомогательную функцию

$$D[\widetilde{M}_{\hat{x}}] = \sum_{i=1}^{n} c_i^2 D_i + 2\lambda\left(1-\sum_{i=1}^{n} c_i\right) = Q(c_1,\ c_2,\ldots,c_n),$$

где λ – неопределённый множитель Лагранжа.

Решаем систему n уравнений

$$\frac{\partial Q}{\partial c_i} = 2c_i D_i - 2\lambda = 0, \quad i=\overline{1,n}$$

относительно переменных $c_1,\ c_2,\ldots,\ c_n$ и получаем

$$c_i = \frac{\lambda}{D_i}.$$

Таким образом, вес c_i, с которым должен входить результат i-го наблюдения в формулу для оценки $\widetilde{M}_{\hat{x}}$, должен быть обратно пропорционален его дисперсии. Иными словами, чем точнее наблюдение, тем с бо́льшим весом необходимо учитывать его результат. Вывод, полученный формально, полностью согласуется с вербальными рассуждениями: чем точнее наблюдение, тем больше ему следует доверять.

Поскольку $\sum_{i=1}^{n} c_i = 1$, то $\lambda\sum_{i=1}^{n}\frac{1}{D_i} = 1$ и, следовательно,

$$\lambda = \frac{1}{\sum_{i=1}^{n}\frac{1}{D_i}}. \qquad (5.1.8)$$

Обозначим $1/D_i = d_i$, тогда (5.1.8) представляется как $\lambda = 1\Big/\sum_{i=1}^{n} d_i$ и

$$c_i = \frac{d_i}{\sum_{i=1}^{n} d_i}, \quad i=\overline{1,n}. \qquad (5.1.9)$$

Таким образом, выражение для оценки (5.1.5) будет иметь вид

$$\widetilde{M}_{\hat{x}} = \frac{\sum_{i=1}^{n} d_i x_i}{\sum_{i=1}^{n} d_i}. \qquad (5.1.10)$$

Оценка вида (5.1.10) является эффективной, так как получена на основе требования минимума дисперсии.

3. Минимальная дисперсия несмещённой оценки $\tilde{M}_{\hat{x}}$

$$D[\tilde{M}_{\hat{x}}] = \sum_{i=1}^{n} c_i^2 D_i = \frac{1}{\left(\sum_{i=1}^{n} d_i\right)^2} \sum_{i=1}^{n} d_i^2 \frac{1}{d_i} = \frac{1}{\sum_{i=1}^{n} d_i} = \lambda, \qquad (5.1.11)$$

а её среднее квадратическое отклонение

$$\sigma[\tilde{M}_{\hat{x}}] = \sqrt{\lambda} = \frac{1}{\sqrt{\sum_{i=1}^{n} d_i}}.$$

Поскольку $d_i = 1/D_i = \text{const}$, $i = \overline{1,n}$, то из выражения (5.1.11) вытекает, что при неограниченном возрастании количества наблюдений $\lambda \to 0$. Следовательно, $\tilde{M}_{\hat{x}}$ сходится по вероятности к $M_{\hat{x}}$, т.е. является состоятельной оценкой математического ожидания $M_{\hat{x}}$.

<u>Частный случай</u>. Предположим, что все наблюдения равноточны. Это означает, что $D_i = D_{\hat{x}}$, $d_i = 1/D_{\hat{x}} = d$, $c_i = 1/n$, $i = \overline{1,n}$ и тогда

$$\tilde{M}_{\hat{x}} = \frac{1}{n} \sum_{i=1}^{n} x_i = M_{\hat{x}}^*$$

Получим результат как и в пп.5.1.1 – оценкой математического ожидания случайной величины $\hat{x}$ является её статистическое среднее $M_{\hat{x}}^*$.

<u>П р и м е р 5.3</u>. Дальность $\hat{x}$ до центра масс ракеты измеряется тремя методами, точность которых характеризуется средними квадратическими отклонениями $\sigma_{\hat{x}_1}$ = 0,2 км, $\sigma_{\hat{x}_2}$ = 0,5 км, $\sigma_{\hat{x}_3}$ = 1 км. Измерения дальности $\hat{x}$ этими методами дали следующие результаты: x_1 = 10,0 км; x_2 = 9,5 км; x_3 = 10,8 км.

Найти оценку $\tilde{M}_{\hat{x}}$ математического ожидания $M_{\hat{x}}$ дальности $\hat{x}$ и среднее квадратическое отклонение $\sigma[\tilde{M}_{\hat{x}}]$ этой оценки.

▼ По условию задачи

$$d_1 = \frac{1}{\sigma_{\hat{x}_1}^2} = \frac{1}{0{,}04} = 25, \quad d_2 = \frac{1}{\sigma_{\hat{x}_2}^2} = \frac{1}{0{,}25} = 4, \quad d_3 = \frac{1}{\sigma_{\hat{x}_3}^2} = \frac{1}{1} = 1.$$

Далее согласно равенствам (5.1.9)

$$c_1 = \frac{25}{30}, \quad c_2 = \frac{4}{30}, \quad c_3 = \frac{1}{30}.$$

По формуле (5.1.5) получаем

$$\tilde{M}_{\hat{x}} = \frac{1}{30}(25 \cdot 10 + 4 \cdot 9{,}5 + 1 \cdot 10{,}8) = 9{,}9 \text{ км}.$$

В соответствии с выражением (5.1.11)

$$D[\tilde{M}_{\hat{x}}] = \frac{1}{\sum_{i=1}^{3} d_i} = \frac{1}{30} \text{ км}^2, \quad \sigma[\tilde{M}_{\hat{x}}] = \sqrt{\frac{1}{30}} = 0{,}183 \text{ км}.$$

▲

5.1.3. Качество оценивания математического ожидания

Качество статистического оценивания математического ожидания при заданном объёме n выборки определяется доверительным интервалом

$$I_{\beta,n} = [M'_{\hat{x}};\ M''_{\hat{x}}] \tag{5.1.12}$$

и доверительной вероятностью

$$\beta_{I,n} = \beta_{\varepsilon,n} = P[M'_{\hat{x}} \le M_{\hat{x}} \le M''_{\hat{x}}]. \tag{5.1.13}$$

Как указывалось в § 3.2, процедура построения доверительного интервала зависит, с одной стороны, от характера распределения наблюдаемого признака $\hat{x}$ и, как следствие, от распределения оценки $\tilde{M}_{\hat{x}}$, а с другой – от объёма n случайной выборки $\hat{X}_{<n>}$. Кроме того, и в первую очередь, она зависит от типа статистики $s(X_{<n>})$, используемой в качестве оценки $\tilde{M}_{\hat{x}}$ математического ожидания. В п.п. 5.1.1 и 5.1.2 применялись линейные статистики в виде средневзвешенного элементов выборки.

В случае равноточных независимых наблюдений наилучшей (по трём критериям, см. § 3.1) оценкой математического ожидания является статистическое математическое ожидание (5.1.1). При этом, если наблюдаемая случайная величина $\hat{x}$ подчинена нормальному закону распределения, то при любом n оценка $\tilde{M}_{\hat{x}}$ будет иметь нормальное распределение с параметрами

$$\left.\begin{aligned} M[\tilde{M}_{\hat{x}}] &= M_{\hat{x}}; \\ D[\tilde{M}_{\hat{x}}] &= \frac{D_{\hat{x}}}{n}; \\ \sigma[\tilde{M}_{\hat{x}}] &= \frac{\sigma_{\hat{x}}}{\sqrt{n}} \end{aligned}\right\}. \tag{5.1.14}$$

Наряду с этим оценка $\tilde{M}_{\hat{x}} = M^*_{\hat{x}}$ асимптотически нормальна, т.е. при $n \to \infty$ для любого распределения признака $\hat{x}$ распределение оценки $\tilde{M}_{\hat{x}}$ приближается к нормальному с параметрами, определяемыми равенствами (5.1.14). Данное утверждение вытекает из предельной теоремы Ляпунова.

Тогда доверительная вероятность (5.1.13) будет определяться отношением

$$\beta = \beta_{I,n} = \beta_{\varepsilon,n} = P(|\tilde{M}_{\hat{x}} - M_{\hat{x}}| \le \varepsilon) \approx 2\Phi_0\left(\frac{\varepsilon}{\sigma[M^*_{\hat{x}}]}\right) = \\ = 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sigma_{\hat{x}}}\right) \approx 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\tilde{\sigma}_{\hat{x}}}\right). \quad (5.1.15)$$

Первое приближённое равенство в (5.1.15) обусловлено отличием закона распределения признака $\hat{x}$ от нормального, а второе – заменой неизвестного $\sigma_{\hat{x}}$ его оценкой $\tilde{\sigma}_{\hat{x}}$. При нормальном распределении $\hat{x}$ и известном $\sigma_{\hat{x}}$ соотношение (5.1.15) становится точным.

Разрешив (5.1.15) относительно ε, получим

$$\varepsilon = \varepsilon_{\beta,n} = \frac{\sigma_{\hat{x}}}{\sqrt{n}}\Phi_0^{-1}\left(\frac{\beta}{2}\right) = \frac{\sigma_{\hat{x}}}{\sqrt{n}}t_\beta \approx \frac{\tilde{\sigma}_{\hat{x}}}{\sqrt{n}}t_\beta, \quad (5.1.16)$$

откуда находим интервал (5.1.12):

$$I_{\beta,n} = [M'_{\hat{x}};\ M''_{\hat{x}}] = [\tilde{M}_{\hat{x}} - \varepsilon_{\beta,n};\ \tilde{M}_{\hat{x}} + \varepsilon_{\beta,n}] = \\ = \left[\tilde{M}_{\hat{x}} - \frac{\sigma_{\hat{x}}}{\sqrt{n}}t_\beta;\ \tilde{M}_{\hat{x}} + \frac{\sigma_{\hat{x}}}{\sqrt{n}}t_\beta\right] \approx \left[\tilde{M}_{\hat{x}} - \frac{\tilde{\sigma}_{\hat{x}}}{\sqrt{n}}t_\beta;\ \tilde{M}_{\hat{x}} + \frac{\tilde{\sigma}_{\hat{x}}}{\sqrt{n}}t_\beta\right]. \quad (5.1.17)$$

Из соотношения (5.1.17) видно, что при большом объёме выборки доверительный интервал для $M_{\hat{x}}$ симметричен и полностью определяется его оценкой и максимальной с вероятностью β абсолютной погрешностью $\varepsilon_{\beta,n}$. На рис.5.1 дана геометрическая интерпретация соотношения (5.1.17).

Из уравнения (5.1.16) выражаем n, при этом будем иметь

$$n = n_{\beta,I} = n_{\beta,\varepsilon} \ge \left(\frac{\sigma_{\hat{x}}}{\varepsilon}t_\beta\right)^2 \approx \left(\frac{\tilde{\sigma}_{\hat{x}}}{\varepsilon}t_\beta\right)^2. \quad (5.1.18)$$

Формулой (5.1.18) определяется потребный объём выборки для оценивания математического ожидания случайной величины $\hat{x}$.

5.2. Оценивание дисперсии и среднего квадратического отклонения случайной величины

Над случайной величиной $\hat{x}$ производится n независимых равноточных наблюдений. Требуется по результатам эксперимента определить состоятельные и несмещённые оценки $\tilde{D}_{\hat{x}}$ и $\tilde{\sigma}_{\hat{x}}$ характеристик рассеяния $D_{\hat{x}}$ и $\sigma_{\hat{x}}$ случайной величины $\hat{x}$, т.е. найти

$$\tilde{D}_{\hat{x}} = \tilde{D}_{\hat{x}}(X_{<n>}) \ \text{ и } \ \tilde{\sigma}_{\hat{x}} = \tilde{\sigma}_{\hat{x}}(X_{<n>}).$$

Ограничимся рассмотрением наиболее важного для практики случая, когда случайная величина $\hat{x}$ подчинена нормальному закону распределения с параметрами $M_{\hat{x}}$ и $\sigma_{\hat{x}}$.

При решении поставленной задачи следует различать два случая – когда параметр $M_{\hat{x}}$ известен и когда он неизвестен.

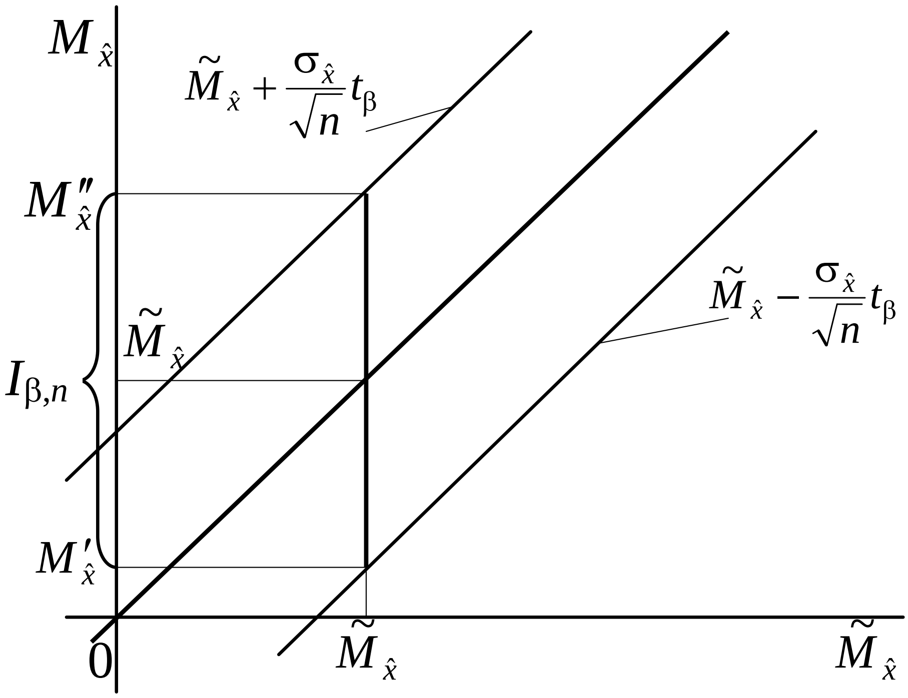

Рис.5.1. Доверительный интервал для математического ожидания

5.2.1. Оценивание дисперсии и среднего квадратического отклонения при известном математическом ожидании

Вводим случайную величину

$$\tilde{D}_{\hat{x}} = D^*_{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n}(\hat{x}_i - M_{\hat{x}})^2 , \tag{5.2.1}$$

которая называется ***дисперсией случайной выборки*** или ***статистической, выборочной дисперсией***. Установим некоторые из свойств случайной величины (5.2.1).

1. Преобразуем $D^*_{\hat{x}}$ к виду

$$D^*_{\hat{x}} = \frac{1}{n}\sigma^2_{\hat{x}}\sum_{i=1}^{n}\frac{(\hat{x}_i - M_{\hat{x}})^2}{\sigma^2_{\hat{x}}} = \frac{\sigma^2_{\hat{x}}}{n}\hat{\chi}^2_n,$$

т.е. $D^*_{\hat{x}}$ является линейной функцией от случайной величины $\hat{\chi}^2_n$, подчинённой хи-квадрат распределению (распределению К. Пирсона) с n степенями свободы. Следовательно

$$M[D^*_{\hat{x}}] = \frac{\sigma^2_{\hat{x}}}{n} M[\hat{\chi}^2_n] = \frac{\sigma^2_{\hat{x}}}{n} n = \sigma^2_{\hat{x}} = D_{\hat{x}}. \qquad (5.2.2)$$

Таким образом, $D^*_{\hat{x}}$ – несмещённая оценка $D_{\hat{x}}$.

2. Поскольку

$$D[D^*_{\hat{x}}] = \frac{\sigma^4_{\hat{x}}}{n^2} D[\hat{\chi}^2_n] = \frac{\sigma^4_{\hat{x}}}{n^2} 2n = \frac{2}{n} D^2_{\hat{x}}; \quad \sigma[D^*_{\hat{x}}] = \sqrt{\frac{2}{n}} \sigma^2_{\hat{x}}, \qquad (5.2.3)$$

то при $n \to \infty$ имеет место $D[D^*_{\hat{x}}] \to 0$. Иначе, дисперсия $D^*_{\hat{x}}$ случайной выборки асимптотически эффективная оценка $D_{\hat{x}}$.

3. Как следует из (5.2.2) и (5.2.3), случайная величина $D^*_{\hat{x}}$ имеет числовые характеристики

$$M[D^*_{\hat{x}}] = D_{\hat{x}}; \quad D[D^*_{\hat{x}}] = \frac{2}{n} D^2_{\hat{x}}; \quad \sigma[D^*_{\hat{x}}] = \sqrt{\frac{2}{n}} \sigma^2_{\hat{x}}.$$

Поскольку

$$\lim_{n \to \infty} D[D^*_{\hat{x}}] = 0,$$

то оценка $D^*_{\hat{x}}$ является состоятельной.

Итак, при $n \to \infty$ дисперсия случайной выборки (5.2.1) представляет собой подходящее значение дисперсии $D_{\hat{x}}$ случайной величины $\hat{x}$. При малых n она в общем случае не вполне эффективна.

<u>П р и м е р 5.4</u>. Полагая $M_{\hat{\tau}}$ = 100 ч, в условиях примера 4.1 найти оценку $\tilde{D}_{\hat{\tau}}$ дисперсии $D_{\hat{\tau}}$ случайной величины $\hat{\tau}$.

▼ Используем данные табл.4.3 и по формуле (5.2.1) получаем

$$\tilde{D}_{\hat{\tau}} = D^*_{\hat{\tau}} = \frac{1}{12} \sum_{i=1}^{12} (\tau_i - 100)^2 = 15115 \text{ ч}^2.$$

▲

Если объём n выборки достаточно велик, то для вычисления оценки $\tilde{D}_{\hat{x}}$ можно пользоваться приближённой формулой

$$\tilde{D}_{\hat{x}} \approx \tilde{\tilde{D}}_{\hat{x}} = \sum_{l=1}^{r} (\bar{x}_l - M_{\hat{x}})^2 P^*_l, \qquad (5.2.4)$$

где $\bar{x}_l$ и P^*_l имеют тот же смысл, что и в формуле (5.1.4).

<u>П р и м е р 5.5</u>. Полагая $M_{\hat{\tau}}$ = 100 ч, в условиях примера 4.2 найти приближённое значение оценки $\tilde{D}_{\hat{\tau}}$ дисперсии случайной величины $\hat{\tau}$.

▼ Используя табл.4.7, по формуле (5.2.4) получаем

$$\tilde{D}_{\hat{\tau}} \approx \tilde{\tilde{D}}_{\hat{\tau}} = \sum_{l=1}^{10} (\bar{\tau}_i - 100)^2 P^*_l = 7975 \text{ ч}^2.$$

▲

Теперь найдем оценку среднего квадратического отклонения $\sigma_{\hat{x}}$. Формула для определения статистического среднего квадратического отклонения имеет вид

$$\sigma^*_{\hat{x}} = \sqrt{D^*_{\hat{x}}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{x}_i - M_{\hat{x}})^2}\,.$$

Требуется выявить основные свойства $\sigma^*_{\hat{x}}$.

1. Из вышеизложенного следует, что $\sigma^*_{\hat{x}}$ является состоятельной и асимптотически эффективной оценкой $\sigma_{\hat{x}}$.

2. Поскольку

$$\sigma^*_{\hat{x}} = \sqrt{\frac{1}{n}\sigma^2_{\hat{x}}\hat{\chi}^2_n} = \frac{\sigma_{\hat{x}}}{\sqrt{n}}\hat{\chi}_n,$$

то

$$M[\sigma^*_{\hat{x}}] = \frac{\sigma_{\hat{x}}}{\sqrt{n}}M[\hat{\chi}_n] = \sigma_{\hat{x}}\sqrt{\frac{2}{n}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \neq \sigma_{\hat{x}},$$

где $\Gamma(n) = \int\limits_0^\infty t^{n-1}e^{-t}dt$ - гамма-функция (интеграл Эйлера 2 рода).

Полученное соотношение указывает на смещённость оценки среднего квадратического отклонения.

Если величину $\sigma^*_{\hat{x}}$ исправить, умножив её на коэффициент

$$k_n = \sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)},$$

то полученная в результате функция случайной выборки

$$\tilde{\sigma}_{\hat{x}} = k_n\sigma^*_{\hat{x}} = k_n\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - M_{\hat{x}})^2} \qquad (5.2.5)$$

будет состоятельной, несмещённой и асимптотически эффективной оценкой среднего квадратического отклонения $\sigma_{\hat{x}}$ случайной величины $\hat{x}$. В табл.5.1 приведены значения коэффициента k_n для некоторых n. Эти значения используются при вычислении оценки (5.2.5).

Таблица 5.1

Значения коэффициента k_n

n	2	3	4	5	6	9	12	18	24
k_n	1,128	1,085	1,064	1,051	1,042	1,028	1,021	1,014	1,010

Из таблицы видно, что необходимость в исправлении оценки возникает лишь при малых объёмах выборки, так как с их увеличением коэффициент k_n достаточно быстро приближается к единице.

5.2.2. Оценивание дисперсии и среднего квадратического отклонения при неизвестном математическом ожидании

Для отыскания оценки дисперсии случайной величины необходимо знать её математическое ожидание $M_{\hat{x}}$. В случае, если данный параметр неизвестен, используют его оценку $\tilde{M}_{\hat{x}} = M^*_{\hat{x}}$.

Рассмотрим функцию случайной выборки в виде статистической дисперсии

$$D^*_{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n}(\hat{x}_i - M^*_{\hat{x}})^2, \qquad (5.2.6)$$

и исследуем её свойства.

1. По аналогии со случаем, когда математическое ожидание известно, можно показать состоятельность оценки (5.2.6).

2. Преобразуем выражение (5.2.6)

$$D^*_{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n}\Big((\hat{x}_i - M_{\hat{x}}) - (M^*_{\hat{x}} - M_{\hat{x}})\Big)^2 =$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n}(\hat{x}_i - M_{\hat{x}})^2 - 2(M^*_{\hat{x}} - M_{\hat{x}})\sum_{i=1}^{n}(\hat{x}_i - M_{\hat{x}}) + \sum_{i=1}^{n}(M^*_{\hat{x}} - M_{\hat{x}})^2\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\hat{x}_i - M_{\hat{x}})^2 - (M^*_{\hat{x}} - M_{\hat{x}})^2 = \frac{\sigma^2_{\hat{x}}}{n}\left(\sum_{i=1}^{n}\left(\frac{\hat{x}_i - M_{\hat{x}}}{\sigma_{\hat{x}}}\right)^2 - \left(\frac{M^*_{\hat{x}} - M_{\hat{x}}}{\sigma_{\hat{x}}/\sqrt{n}}\right)^2\right) =$$

$$= \frac{\sigma^2_{\hat{x}}}{n}(\hat{\chi}^2_n - \hat{\chi}^2_1).$$

Поскольку случайные величины $\hat{\chi}^2_n$ и $\hat{\chi}^2_1$ являются функциями одной и той же выборки, то они зависимы. Причём их зависимость такова, что разность этих случайных величин оказывается подчинённой закону распределения хи-квадрат с $n-1$ степенями свободы. Таким образом

$$D^*_{\hat{x}} = \frac{\sigma^2_{\hat{x}}}{n}\hat{\chi}^2_{n-1}, \qquad (5.2.7)$$

откуда

$$M[D^*_{\hat{x}}] = \frac{\sigma^2_{\hat{x}}}{n}M[\hat{\chi}^2_{n-1}] = \frac{n-1}{n}\sigma^2_{\hat{x}} \neq D_{\hat{x}}.$$

Следовательно, статистическая дисперсия $D^*_{\hat{x}}$ оказывается смещённой оценкой параметра $D_{\hat{x}}$. Для исправления оценки $D^*_{\hat{x}}$ её достаточно умножить на коэффициент $n(n–1)$. С ростом объёма n выборки указанный коэффициент стремится к единице, поэтому при достаточно больших n смещённостью оценки $D^*_{\hat{x}}$ можно пренебречь.

3. Поскольку

$$D[D^*_{\hat{x}}] = \frac{\sigma^4_{\hat{x}}}{n^2} D[\hat{\chi}^2_{n-1}] = \frac{2(n-1)}{n^2} D^2_{\hat{x}}, \quad \sigma[D^*_{\hat{x}}] = \frac{\sqrt{2(n-1)}}{n} \sigma^2_{\hat{x}}, \quad (5.2.8)$$

то при $n \to \infty$ имеет место $D[D^*_{\hat{x}}] \to 0$. Результат, полученный на основе анализа выражения (5.2.8), свидетельствует об асимптотической эффективности оценки $D^*_{\hat{x}}$.

Итак, при $n \to \infty$ исправленная статистическая дисперсия

$$\tilde{D}_{\hat{x}} = \frac{n}{n-1} D^*_{\hat{x}} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \tilde{M}_{\hat{x}})^2 \quad (5.2.9)$$

является подходящим значением дисперсии $D_{\hat{x}}$ случайной величины $\hat{x}$. С уменьшением объёма n выборки эффективность этой оценки несколько падает.

Оценка (5.2.9) имеет следующие числовые характеристики:

$$M[\tilde{D}_{\hat{x}}] = D_{\hat{x}}; \quad D[\tilde{D}_{\hat{x}}] = \frac{2}{n-1} D^2_{\hat{x}}; \quad \sigma[\tilde{D}_{\hat{x}}] = \sqrt{\frac{2}{n-1}} \sigma^2_{\hat{x}}. \quad (5.2.10)$$

Вычисление дисперсии $D[\tilde{D}_{\hat{x}}]$ связано со сложными выкладками, поэтому её выражение приведено без вывода.

При большом объёме n выборки приближённое значение оценки $\tilde{D}_{\hat{x}}$ можно вычислять по формуле

$$\tilde{D}_{\hat{x}} = \tilde{\tilde{D}}_{\hat{x}} = \sum_{l=1}^{r} (\bar{x}_l - \tilde{M}_{\hat{x}})^2 P^*_l. \quad (5.2.11)$$

Перейдём к отысканию оценки для среднего квадратического отклонения $\sigma_{\hat{x}}$ случайной величины $\hat{x}$ в случае неизвестного математического ожидания. Указанная оценка определяется по формуле

$$\tilde{\sigma}'_{\hat{x}} = \sqrt{\tilde{D}_{\hat{x}}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \tilde{M}_{\hat{x}})^2}. \quad (5.2.12)$$

Проанализируем свойства оценки (5.2.12).

1. Из вышеизложенного следует, что данная оценка состоятельна и асимптотически эффективна.

2. Согласно выражениям (5.2.7) и (5.2.9) имеем

$$\tilde{\sigma}'_{\hat{x}} = \frac{\sigma_{\hat{x}}}{\sqrt{n-1}} \hat{\chi}_{n-1}.$$

Поэтому

$$M[\tilde{\sigma}'_{\hat{x}}] = \frac{\sigma_{\hat{x}}}{\sqrt{n-1}} M[\hat{\chi}_{n-1}] = \sqrt{\frac{2}{n-1}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sigma_{\hat{x}} \neq \sigma_{\hat{x}}$$

и, следовательно, (5.2.12) является смещённой оценкой $\sigma_{\hat{x}}$. Для исправления данной оценки её достаточно умножить на коэффициент

$$k_{n-1} = \sqrt{\frac{n-1}{2}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}.$$

Полученная при этом функция случайной выборки

$$\tilde{\sigma}_{\hat{x}} = k_{n-1}\tilde{\sigma}'_{\hat{x}} = k_{n-1}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \tilde{M}_{\hat{x}})^2} \qquad (5.2.13)$$

будет состоятельной, несмещённой и асимптотически эффективной оценкой среднего квадратического отклонения $\sigma_{\hat{x}}$ случайной величины $\hat{x}$.

Подходящее значение $\tilde{\sigma}_{\hat{x}}$ можно получить и непосредственно, используя статистическую дисперсию (5.2.6). Для исправления получаемой при этом оценки $\sigma^*_{\hat{x}} = \sqrt{D^*_{\hat{x}}}$ её необходимо умножить на коэффициент

$$k'_{n-1} = \sqrt{\frac{n}{n-1}} k_{n-1},$$

т.е. величина

$$\tilde{\sigma}_{\hat{x}} = k'_{n-1}\sigma^*_{\hat{x}} = \sqrt{\frac{n}{n-1}} k_{n-1}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \tilde{M}_{\hat{x}})^2}$$

является состоятельной, несмещённой и асимптотически эффективной оценкой.

В заключение отметим, что поскольку известное соотношение для дисперсии

$$D_{\hat{x}} = \nu_2[\hat{x}] - \nu_1^2[\hat{x}]$$

справедливо и для её оценки, т.е.

$$\tilde{D}_{\hat{x}} = \tilde{\nu}_2[\hat{x}] - \tilde{\nu}_1^2[\hat{x}],$$

то формулам (5.2.1), (5.2.4) и (5.2.9), (5.2.11) соответственно можно придать более удобный для практического использования вид:

$$\left.\begin{aligned} \tilde{D}_{\hat{x}} &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - M_{\hat{x}}^2 \\ \tilde{D}_{\hat{x}} &= \tilde{\tilde{D}}_{\hat{x}} = \sum_{l=1}^{r} \bar{x}_l^2 P_l^* - M_{\hat{x}}^2 \end{aligned}\right\}$$

$$\left.\begin{aligned}\tilde{D}_{\hat{x}} &= \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \tilde{M}_{\hat{x}}^2\right)\frac{n}{n-1}\\ \tilde{D}_{\hat{x}} &\approx \tilde{\tilde{D}}_{\hat{x}} = \sum_{l=1}^{r} \bar{x}_l^2 P_l^* - \tilde{M}_{\hat{x}}^2\end{aligned}\right\} \qquad (5.2.14)$$

<u>П р и м е р 5.6</u>. Производятся измерения одного из габаритных размеров $\hat{x}$ однотипных деталей. Данные измерений сведены в табл.5.2.

Т а б л и ц а 5.2

Результаты измерений размера деталей

i	x_i	i	x_i	i	x_i	i	x_i	i	x_i
1	10,5	5	10,4	9	10,3	13	10,5	17	10,8
2	10,8	6	10,6	10	10,8	14	10,7	18	10,7
3	11,2	7	10,9	11	10,6	15	10,8	19	10,9
4	10,9	8	11,0	12	11,3	16	10,9	20	11,0

1. Найти оценку $\tilde{M}_{\hat{x}}$ математического ожидания величины $\hat{x}$ и построить доверительный интервал, соответствующий доверительной вероятности $\beta = 0{,}8$.

2. Определить доверительную вероятность β для математического ожидания случайной величины $\hat{x}$, если максимальная вероятная погрешность $\varepsilon_\beta = 0{,}07$.

▼ 1. В соответствии с выражением (5.1.1) находим оценку математического ожидания

$$\tilde{M}_{\hat{x}} = \frac{1}{20}\sum_{i=1}^{20} x_i = 10{,}78.$$

По первой формуле (5.2.14) находим оценку дисперсии

$$\tilde{D}_{\hat{x}} = \left(\frac{1}{20}\sum_{i=1}^{20} x_i^2 - (10{,}78)^2\right)\frac{20}{19} = 0{,}064.$$

Оценка среднего квадратического отклонения

$$\tilde{\sigma}_{\hat{x}} = \sqrt{\tilde{D}_{\hat{x}}} = \sqrt{0{,}064} = 0{,}253.$$

При заданном $\beta = 0{,}8$ величина $t_\beta = 1{,}282$ (см. приложение 4), тогда максимальное вероятное отклонение математического ожидания найдём по формуле (5.1.16), что составит

$$\varepsilon = \varepsilon_{0,8;\,20} = \frac{0{,}253}{\sqrt{20}} \cdot 1{,}282 = 0{,}072.$$

Выражение (5.1.17) даёт следующий результат:

$I_{0,8;\,20} = [10{,}78–0{,}072;\ 10{,}78+0{,}072] = [10{,}71;\ 10{,}85]$.

2. В соответствии с соотношением (5.1.15) находим

$$\beta = \beta_{0,07;\,20} = 2\Phi_0\left(\frac{0{,}07\sqrt{20}}{0{,}253}\right) = 2\Phi_0(1{,}237) = 2 \cdot 0{,}3912 = 0{,}7824.$$

Значение функции $\Phi_0(x)$ найдены в приложении 2.

▲

<u>П р и м е р 5.7</u>. Габаритный размер $\hat{x}$ деталей измеряется методом, который характеризуется дисперсий $D_{\hat{x}}$ = 0,064. Определить потребный объём выборки, чтобы максимальная вероятная погрешность ε оценки среднего размера деталей не превосходила 0,06 при доверительной вероятности β = 0,93.

▼ Для β = 0,93 в приложении 4 находим t_β = 1,810. Тогда из выражения (5.1.18) получаем

$$n \ge \left(\frac{\sqrt{0{,}064}}{0{,}06}1{,}810\right)^2 = 58.$$

▲

5.2.3. Качество оценивания дисперсии

Качество оценивания дисперсии $D_{\hat{x}}$ характеризуется доверительным интервалом

$$I_{\beta,n} = [D'_{\hat{x}};\, D''_{\hat{x}}]$$

и доверительной вероятностью

$$\beta_{I,n} = P(D'_{\hat{x}} \le D_{\hat{x}} \le D''_{\hat{x}}].$$

При оценивании дисперсии различают случаи известного и неизвестного математического ожидания (см. пп. 5.2.1, 5.2.2). Рассмотрим второй, наиболее распространённый случай. При неизвестном $M_{\hat{x}}$ состоятельная, несмещённая и асимптотически эффективная оценка дисперсии определяется равенством (5.2.9).

При достаточно большом объёме выборки распределение оценки (5.2.9) будет близким к нормальному с параметрами (5.2.10). Следовательно, доверительная вероятность для дисперсии будет приближённо определяться соотношением

$$\beta = \beta_{I,n} = \beta_{\varepsilon,n} = P(|\tilde{D}_{\hat{x}} - D_{\hat{x}}| \le \varepsilon) \approx 2\Phi_0\left(\frac{\varepsilon}{\sigma[D_{\hat{x}}]}\right) = 2\Phi_0\left(\frac{\varepsilon\sqrt{n-1}}{D_{\hat{x}}\sqrt{2}}\right), \quad (5.2.15)$$

где $\varepsilon = \varepsilon_{\beta,n}$ – максимальная с вероятностью β абсолютная погрешность оценки $\tilde{D}_{\hat{x}}$ дисперсии $D_{\hat{x}}$.

Выражаем ε из уравнения (5.2.15) и в результате получим

$$\varepsilon = \varepsilon_{\beta,n} \approx D_{\hat{x}}\sqrt{\frac{2}{n-1}}t_\beta, \quad (5.2.16)$$

откуда доверительный интервал

$$I_{\beta,n} = [D'_{\hat{x}};\, D''_{\hat{x}}] \approx \left[\tilde{D}_{\hat{x}} - D_{\hat{x}}\sqrt{\frac{2}{n-1}}t_\beta;\;\; \tilde{D}_{\hat{x}} + D_{\hat{x}}\sqrt{\frac{2}{n-1}}t_\beta \right]. \qquad (5.2.17)$$

Выражения (5.2.15) – (5.2.17) были получены в предположении, что дисперсия $D_{\hat{x}}$ известна. В действительности известна лишь её оценка $\tilde{D}_{\hat{x}}$, которая только и может фигурировать в этих формулах. С учётом сказанного будем иметь

$$\beta_{\varepsilon,n} \approx 2\Phi_0\left(\frac{\varepsilon\sqrt{n-1}}{\tilde{D}_{\hat{x}}\sqrt{2}} \right); \qquad (5.2.18)$$

$$\varepsilon_{\beta,n} \approx \tilde{D}_{\hat{x}}\sqrt{\frac{2}{n-1}}t_\beta; \qquad (5.2.19)$$

$$I_{\beta,n} \approx \left[\tilde{D}_{\hat{x}} - \tilde{D}_{\hat{x}}\sqrt{\frac{2}{n-1}}t_\beta;\;\; \tilde{D}_{\hat{x}} + \tilde{D}_{\hat{x}}\sqrt{\frac{2}{n-1}}t_\beta \right]. \qquad (5.2.20)$$

Как видно из выражения (5.2.20), доверительный интервал для дисперсии $D_{\hat{x}}$ оказывается симметричным относительно оценки $\tilde{D}_{\hat{x}}$. При этом его ширина зависит от значения $\tilde{D}_{\hat{x}}$ (рис.5.2).

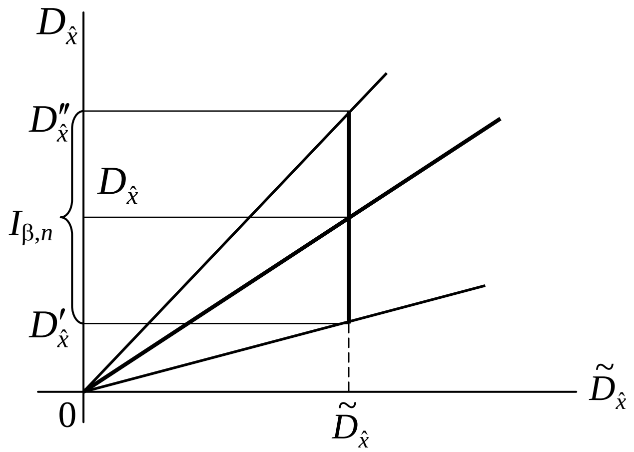

Рис.5.2. Доверительный интервал для дисперсии

Потребный объём экспериментальных данных для оценивания дисперсии $D_{\hat{x}}$ с заданными точностью и надёжностью получим, если выразить n из уравнения (5.2.15):

$$n = n_{\beta,\varepsilon} \geq 2\left(\frac{\tilde{D}_{\hat{x}} t_\beta}{\varepsilon} \right)^2 + 1, \qquad (5.2.21)$$

Следует отметить, что n определяется, когда дисперсия ещё только подлежит оцениванию. Но значения и вероятностные характеристики оценки $\tilde{D}_{\hat{x}}$, входящей в выражение (5.2.21), зависят от объёма выборки. Поэтому объём n определяется методом последовательных приближений.

В первом приближении задаются некоторым ориентировочным значением n_0, при котором вычисляется приближение $\tilde{D}_{\hat{x}}^0$ оценки $\tilde{D}_{\hat{x}}$. Затем оно уточняется в последующих циклах вычислений. Очевидно, что если n_0 превышает найденное по формуле (5.2.21) значение n, то принимается $n = n_0$. В данном случае оценка $\tilde{D}_{\hat{x}}^0 = \tilde{D}_{\hat{x}}$ уже удовлетворяет требованиям по точности и надёжности.

П р и м е р 5.8. В условиях примера 5.6:

1) найти приближённое значение числовых характеристик дисперсии случайной величины $\hat{x}$;

2) построить 80-процентный доверительный интервал для дисперсии;

3) определить доверительную вероятность β для дисперсии, если максимальная с вероятностью β погрешность $\varepsilon_\beta = 0{,}02$.

▼ 1. По формулам (5.2.10) получаем:

$$M[\tilde{D}_{\hat{x}}] \approx \tilde{D}_{\hat{x}} = 0{,}064; \quad D[\tilde{D}_{\hat{x}}] \approx \frac{2}{20-1}(0{,}064)^2 = 0{,}0004;$$

$$\sigma[\tilde{D}_{\hat{x}}] \approx \sqrt{\frac{2}{20-1}}0{,}064 = 0{,}20.$$

2. Используем выражение (5.2.19):

$$\varepsilon = \varepsilon_{0,8;\,20} = 0{,}064\sqrt{\frac{2}{20-1}}1{,}282 = 0{,}027.$$

Доверительный интервал для дисперсии в соответствии с (5.2.20):

$$I = I_{0,8;\,20} \approx [0{,}064 - 0{,}027;\ 0{,}064 + 0{,}027] = [0{,}037;\ 0{,}091].$$

3. По формуле (5.2.18)

$$\beta = \beta_{0,02;\,20} \approx 2\Phi_0\left(\frac{0{,}02\sqrt{20-1}}{0{,}064\sqrt{2}}\right) = 2\Phi_0(0{,}97) = 2\cdot 0{,}3337 = 0{,}6674.$$

Значение $\Phi_0(x)$ найдено в приложении 2.

▲

5.3. Оценивание числовых характеристик и параметров распределения случайных векторов

5.3.1. Двумерный случайный вектор

В разделе 4 были рассмотрены методы оценивания закона распределения случайной величины во всех его возможных формах – ряда, функции и плотности распределения. Аналогичная задача возникает и при обработке экспериментальных данных в виде случайных векторов, т.е. систем стохастически связанных между собой случайных величин. Такие системы характеризуются многомерными законами распределения.

В данном параграфе более подробно остановимся на оценивании параметров распределения случайных векторов. При этом начнём с рассмотрения частного случая – двумерного вектора, т.е. системы двух случайных величин.

Над системой двух случайных величин $(\hat{x};\hat{y})$ произведено n независимых равноточных наблюдений, в результате которых получена последовательность пар чисел $(x_i;y_i)$, $i=\overline{1,n}$ (табл.5.3), которые можно интерпретировать как координаты точек $(x_1;\,y_1)$, $(x_2;\,y_2)$, …, $(x_n;\,y_n)$ плоскости.

Таблица 5.3

Двумерный массив экспериментальных данных

i	1	2	…	i	…	n
x_i	x_1	x_2	…	x_i	…	x_n
y_i	y_1	y_2	…	y_i	…	y_n

Требуется по результатам наблюдений определить состоятельные, несмещённые и эффективные (асимптотически эффективные) оценки числовых характеристик $M_{\hat{x}}$, $M_{\hat{y}}$, $D_{\hat{x}}$, $D_{\hat{y}}$, $K_{\hat{x}\hat{y}}$ системы случайных величин $(\hat{x};\hat{y})$. В данном случае $K_{\hat{x}\hat{y}}$ – корреляционный момент $\hat{x}$ и $\hat{y}$.

Задача определения точечных оценок параметров двумерного распределения решается так же, как и для одной случайной величины. При этом оценки координат $M_{\hat{x}}$, $M_{\hat{y}}$ центра рассеяния системы $(\hat{x};\hat{y})$ находятся по формулам

$$\tilde{M}_{\hat{x}} = M^*_{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n} x_i\,, \quad \tilde{M}_{\hat{y}} = M^*_{\hat{y}} = \frac{1}{n}\sum_{i=1}^{n} y_i\,, \qquad (5.3.1)$$

а оценки элементов её корреляционной матрицы $K_{\hat{x}\hat{y}}$ определяются выражениями

$$\left.\begin{aligned} \tilde{D}_{\hat{x}} &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\tilde{M}_{\hat{x}})^2;\\ \tilde{D}_{\hat{y}} &= \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\tilde{M}_{\hat{y}})^2;\\ \tilde{K}_{\hat{x}\hat{y}} &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\tilde{M}_{\hat{x}})(y_i-\tilde{M}_{\hat{y}}). \end{aligned}\right\} \qquad (5.3.2)$$

Если математические ожидания $M_{\hat{x}}$, $M_{\hat{y}}$ известны, то элементы корреляционной матрицы определяются выборочными дисперсиями и корреляционным моментом (см. п.п. 5.2.1):

$$\left.\begin{aligned}
\tilde{D}_{\hat{x}} &= D^*_{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \tilde{M}_{\hat{x}})^2;\\
\tilde{D}_{\hat{y}} &= D^*_{\hat{y}} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{M}_{\hat{y}})^2;\\
\tilde{K}_{\hat{x}\hat{y}} &= K^*_{\hat{x}\hat{y}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \tilde{M}_{\hat{x}})(y_i - \tilde{M}_{\hat{y}}).
\end{aligned}\right\}$$

При вычислении оценок $\tilde{D}_{\hat{x}}$, $\tilde{D}_{\hat{y}}$, $\tilde{K}_{\hat{x}\hat{y}}$ целесообразно воспользоваться известной связью между центральными и начальными моментами, которая имеет место и для их статистических аналогов

$$\left.\begin{aligned}
D^*_{\hat{x}} &= \mu^*_2[\hat{x}] = \nu^*_2[\hat{x}] - (\nu^*_1[\hat{x}])^2;\\
\tilde{D}_{\hat{y}} &= \mu^*_2[\hat{y}] = \nu^*_2[\hat{y}] - (\nu^*_1[\hat{y}])^2;\\
\tilde{K}_{\hat{x}\hat{y}} &= \mu^*_{11}[\hat{x};\, \hat{y}] = \nu^*_{11}[\hat{x};\, \hat{y}] - \nu^*_1[\hat{x}]\nu^*_1[\hat{y}].
\end{aligned}\right\} \qquad (5.3.3)$$

С учётом (5.3.3) выражения (5.3.2) принимают вид, который обычно используется на практике:

$$\left.\begin{aligned}
\tilde{D}_{\hat{x}} &= \left(\frac{1}{n}\sum_{i=1}^{n}x_i^2 - \tilde{M}^2_{\hat{x}}\right)\frac{n}{n-1};\\
\tilde{D}_{\hat{y}} &= \left(\frac{1}{n}\sum_{i=1}^{n}y_i^2 - \tilde{M}^2_{\hat{y}}\right)\frac{n}{n-1};\\
\tilde{K}_{\hat{x}\hat{y}} &= \left(\frac{1}{n}\sum_{i=1}^{n}x_i y_i - \left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right)\right)\frac{n}{n-1}.
\end{aligned}\right\} \qquad (5.3.4)$$

Очевидно, что оценка $\tilde{r}_{\hat{x}\hat{y}}$ коэффициента корреляции $r_{\hat{x}\hat{y}}$ найдётся по формуле

$$\tilde{r}_{\hat{x}\hat{y}} = \frac{\tilde{K}_{\hat{x}\hat{y}}}{\tilde{\sigma}_{\hat{x}}\tilde{\sigma}_{\hat{y}}}. \qquad (5.3.5)$$

П р и м е р 5.9. Пусть $\hat{x}$ и $\hat{y}$ – координаты пробоины в мишени после выстрела (в сантиметрах). По мишени произведено 10 независимых выстрелов, результаты которых сведены в табл.5.4, где i – номер выстрела. Найти оценки числовых характеристик $M_{\hat{x}}$, $M_{\hat{y}}$, $\sigma_{\hat{x}}$, $\sigma_{\hat{y}}$, $r_{\hat{x}\hat{y}}$, системы случайных величин $(\hat{x};\, \hat{y})$.

Т а б л и ц а 5.4

Координаты пробоин в мишени

i	1	2	3	4	5	6	7	8	9	10
x_i	3	1	2,5	1,5	4	3,5	2	2,5	1,5	1
y_i	5	2	4	3,5	1,5	5,5	2,5	4,5	2	0

▼ По формулам (5.3.1) получим

$$\tilde{M}_{\hat{x}} = \frac{1}{10}\sum_{i=1}^{10} x_i = 2{,}55 \text{ см}, \quad \tilde{M}_{\hat{y}} = \frac{1}{10}\sum_{i=1}^{10} y_i = 2{,}95 \text{ см}.$$

Используя соотношения (5.3.4), имеем:

$$\tilde{D}_{\hat{x}} = \left(\frac{1}{10}\sum_{i=1}^{10} x_i^2 - (2{,}55)^2\right)\frac{10}{10-1} = 1{,}068 \text{ см}^2;$$

$$\tilde{D}_{\hat{y}} = \left(\frac{1}{10}\sum_{i=1}^{10} y_i^2 - (2{,}95)^2\right)\frac{10}{10-1} = 3{,}359 \text{ см}^2;$$

$$\tilde{K}_{\hat{x}\hat{y}} = \left(\frac{1}{10}\sum_{i=1}^{10} x_i y_i - 2{,}55\cdot 2{,}95\right)\frac{10}{10-1} = 0{,}986 \text{ см}^2.$$

Наконец, используя табл.5.1, по формулам (5.2.13) и (5.3.5), получим:

$$\tilde{\sigma}_{\hat{x}} = k_9\sqrt{\tilde{D}_{\hat{x}}} = 1{,}028\sqrt{1{,}068} = 1{,}064 \text{см};$$

$$\tilde{\sigma}_{\hat{y}} = 1{,}028\sqrt{3{,}359} = 1{,}885 \text{ см};$$

$$\tilde{r}_{\hat{x}\hat{y}} = \frac{0{,}986}{1{,}064\cdot 1{,}885} = 0{,}491.$$

▲

5.3.2. Многомерный случайный вектор

Аналогично решается задача оценивания числовых характеристик системы произвольного числа случайных величин.

Пусть имеется m случайных величин $(\hat{x}_1,\ \hat{x}_2, ..., \hat{x}_m)^{\text{т}} = \hat{X}_{<m>}$. Над системой произведено n независимых равноточных наблюдений, результаты которых оформлены в виде табл.5.5. Указанная таблица носит название ***простой статистической матрицы***.

Таблица 5.5

Простая статистическая матрица

i	x_{ij}					
	x_{i1}	x_{i2}	…	x_{ij}	…	x_{im}
1	x_{11}	x_{12}	…	x_{1j}	…	x_{1m}
2	x_{21}	x_{22}	…	x_{2j}	…	x_{2m}
…	…	…	…	…	…	…
i	x_{i1}	x_{i2}	…	x_{ij}	…	x_{im}
…	…	…	…	…	…	…
n	x_{n1}	x_{n2}	…	x_{nj}	…	x_{nm}

В табл.5.5 величины x_{ij}, $i = \overline{1, n}$, $j = \overline{1, m}$ – это значения, принятые случайной величиной $\hat{x}_j$ в i-м опыте.

Требуется найти оценки числовых характеристик системы случайных величин, т.е. оценки математических ожиданий $(M_{\hat{x}_1}, M_{\hat{x}_2},..., M_{\hat{x}_m})^{\mathrm{T}} = M_{\hat{X}_{<m>}}$ и элементов корреляционной матрицы

$$K_{\hat{X}_{<m>}} = \| K_{il} \|_m^m = \begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1m} \\ & K_{22} & \cdots & K_{2m} \\ & & \cdots & \cdots \\ & & & K_{mm} \end{pmatrix},$$

где $K_{jl} = K_{\hat{x}_j \hat{x}_l}$, $K_{jj} = K_{\hat{x}_j \hat{x}_j} = D_{\hat{x}_j}$, $j, l = \overline{1, m}$.

Выведенные ранее формулы для вычисления состоятельных, несмещённых и эффективных (асимптотически эффективных) оценок числовых характеристик в общем случае системы m случайных величин приобретают следующий вид:

$$\left.\begin{aligned} &\tilde{M}_j = \tilde{M}_{\hat{x}_j} = M^*_{\hat{x}_j} = \frac{1}{n}\sum_{i=1}^{n} x_{ij}; \\ &\tilde{D}_j = \tilde{D}_{\hat{x}_j} = \frac{n}{n-1} D^*_{\hat{x}_j} = \frac{1}{n-1}\sum_{i=1}^{n} (x_{ij} - \tilde{M}_{\hat{x}_j})^2; \\ &\tilde{K}_{jl} = \tilde{K}_{\hat{x}_j \hat{x}_l} = \frac{n}{n-1} K^*_{\hat{x}_j \hat{x}_l} = \frac{1}{n-1}\sum_{i=1}^{n} (x_{ij} - \tilde{M}_{\hat{x}_j})(x_{il} - \tilde{M}_{\hat{x}_l}). \end{aligned}\right\} \qquad (5.3.6)$$

Для вычисления оценок (5.3.6) могут быть использованы формулы типа (5.3.4). Оценки средних квадратических отклонений:

$$\tilde{\sigma}_j = \tilde{\sigma}_{\hat{x}_j} = \sqrt{\tilde{D}_{\hat{x}_j}}, \quad j = \overline{1, m}$$

Зная $\tilde{K}_{jl}$ и $\tilde{\sigma}_j$, нетрудно найти оценки элементов нормированной корреляционной матрицы по формулам

$$\tilde{r}_{jl} = \tilde{r}_{\hat{x}_j \hat{x}_l} = \frac{\tilde{K}_{jl}}{\tilde{\sigma}_j \tilde{\sigma}_l}, \quad j, l = \overline{1, m}. \qquad (5.3.7)$$

5.3.3. Качество оценивания числовых характеристик случайных векторов

Основные вероятностные свойства m-мерного случайного вектора

$$\hat{X}_{<m>} = (\hat{x}_1, \hat{x}_2, ..., \hat{x}_m)^{\mathrm{T}}$$

описываются m-мерным вектором математических ожиданий его компонент

$$M_{\hat{X}_{<m>}} = (M_{\hat{x}_1}, M_{\hat{x}_2}, ..., M_{\hat{x}_m})^{\mathrm{T}} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_m)^{\mathrm{T}} = \bar{X}_{<m>}.$$

и корреляционной матрицей m-го порядка

$$K_{\hat{X}_{<m>}} = \| K_{\hat{x}_j \hat{x}_l} \|_m^m = \begin{pmatrix} D_{\hat{x}_1} & K_{\hat{x}_1\hat{x}_2} & \cdots & K_{\hat{x}_1\hat{x}_m} \\ & D_{\hat{x}_2} & \cdots & K_{\hat{x}_2\hat{x}_m} \\ & & \cdots & \cdots \\ & & & D_{\hat{x}_m} \end{pmatrix} = \overline{\dot{X}_{<m>} \dot{X}^{\mathsf{T}}_{<m>}},$$

т.е. всего $m + m^2$ числовыми характеристиками. Поскольку корреляционная матрица симметрична, то число различных её элементов равно $0{,}5(m + m^2)$. Таким образом, при оценивании числовых характеристик m-мерного случайного вектора $\hat{X}_{<m>}$ необходимо вычислить $0{,}5(m^2 + 3m)$ оценок, из которых m оценок $\tilde{M}_{\hat{x}_j}$ математических ожиданий $M_{\hat{x}_j}$, $j = \overline{1,m}$, такое же количество оценок $\tilde{D}_{\hat{x}_j}$ дисперсий $D_{\hat{x}_j}$, $j = \overline{1,m}$ и $\tilde{K}_{\hat{x}_j \hat{x}_l}$ оценок корреляционных моментов $K_{\hat{x}_j \hat{x}_l}$, $i, j = \overline{1,m}$.

Задача анализа качества оценивания числовых характеристик случайного вектора $\hat{X}_{<m>}$ заключается в построении для них $0{,}5(m^2+3m)$ доверительных интервалов

$$I_{\beta_1,n} = [M'_{\hat{x}_j}; M''_{\hat{x}_j}], \quad j = \overline{1,m},$$

$$I_{\beta_2,n} = [D'_{\hat{x}_j}; D''_{\hat{x}_j}], \quad j = \overline{1,m},$$

$$I_{\beta_3,n} = [K'_{\hat{x}_j \hat{x}_l}; K''_{\hat{x}_j \hat{x}_l}], \quad j,l = \overline{1,m},\ j < l$$

и вычислении такого же числа доверительных вероятностей:

$$\beta_{I_1,n} = P(M'_{\hat{x}_j} \le M_{\hat{x}_j} \le M''_{\hat{x}_j}), \quad j = \overline{1,m},$$

$$\beta_{I_2,n} = P(D'_{\hat{x}_j} \le D_{\hat{x}_j} \le D''_{\hat{x}_j}), \quad j = \overline{1,m},$$

$$\beta_{I_3,n} = P(K'_{\hat{x}_j \hat{x}_l} \le K_{\hat{x}_j \hat{x}_l} \le K''_{\hat{x}_j \hat{x}_l}), \quad j,l = \overline{1,m},\ j < l.$$

Методики анализа точности и надёжности оценивания числовых характеристик положения (математических ожиданий и рассеяния (дисперсий и средних квадратических отклонений) были подробно рассмотрены в § 5.2. Поэтому здесь основное внимание будет уделено анализу качества оценивания числовых характеристик связи (корреляционных моментов и коэффициентов корреляции) компонент случайного вектора $\hat{X}_{<m>}$.

При $k_{n-1} \approx 1$ точечные оценки $\tilde{K}_{\hat{x}_j \hat{x}_l}$ и $\tilde{r}_{\hat{x}_j \hat{x}_l}$ вычисляются соответственно по третьей формуле (5.3.6) и (5.3.7). Известно, что корреляционный момент кроме связи компонент случайного вектора $\hat{X}_{<m>}$ характеризует и их рассеяние. Поэтому в качестве основной характеристики связи чаще всего используется коэффициент корреляции $r_{\hat{x}_j \hat{x}_l} = r_{jl} = r$.

При вычислении доверительной вероятности $\beta = \beta_{I,n}$ и построении доверительного интервала $I = I_{\beta,n}$ для коэффициента корреляции необходимо знать закон распределения его оценки $\tilde{r}$. Оказывается, что независимо от распределения случайного вектора $\hat{X}^{il}_{<2>} = (\hat{x}_j; \hat{x}_l)^{\mathrm{T}}$ при достаточно большом объёме n выборки (практически при $n > 30$) закон распределения оценки $\tilde{r}$ близок к нормальному [1] с параметрами

$$M_{\tilde{r}} \approx r, \quad \sigma_{\tilde{r}} \approx \frac{\sqrt{1-r^2}}{\sqrt{n}},$$

иначе

$$\varphi_{\tilde{r}} = \varphi^{\text{н}}_{\tilde{r}}(x; M_{\tilde{r}}; \sigma_{\tilde{r}}) = \varphi^{\text{н}}_{\tilde{r}}\left(x; r; \frac{\sqrt{1-r^2}}{\sqrt{n}}\right).$$

Поэтому доверительная вероятность

$$\beta = \beta_{I,n} = \beta_{\varepsilon,n} = P(|\tilde{r} - r| \le \varepsilon) = 2\Phi_0\left(\frac{\varepsilon}{\sigma_{\tilde{r}}}\right) = \\ = 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sqrt{1-r^2}}\right) \approx 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sqrt{1-\tilde{r}^2}}\right). \tag{5.3.8}$$

Выражая из уравнения (5.3.8) величину ε, будем иметь

$$\varepsilon = \varepsilon_{\beta,n} = \sigma_{\tilde{r}} t_\beta = \frac{\sqrt{1-r^2}}{\sqrt{n}} t_\beta \approx \frac{\sqrt{1-\tilde{r}^2}}{\sqrt{n}} t_\beta, \tag{5.3.9}$$

откуда

$$I = I_{\beta,n} = [r'; r''] = \left[\tilde{r} - \frac{\sqrt{1-r^2}}{\sqrt{n}} t_\beta;\ \tilde{r} + \frac{\sqrt{1-r^2}}{\sqrt{n}} t_\beta\right] \approx \\ \approx \left[\tilde{r} - \frac{\sqrt{1-\tilde{r}^2}}{\sqrt{n}} t_\beta;\ \tilde{r} + \frac{\sqrt{1-\tilde{r}^2}}{\sqrt{n}} t_\beta\right]. \tag{5.3.10}$$

Соотношением (5.3.10) описывается 100β-процентной доверительный интервал для коэффициента корреляции r.

Разрешив уравнение (5.3.9) относительно n, получим выражение для объёма выборки, потребного при оценивании r с точностью ε (с абсолютной ошибкой, не превосходящей ε) и надёжностью β (доверительной вероятностью β):

$$n = n_{\beta,\varepsilon} \ge \frac{1-r^2}{\varepsilon^2} t^2_\beta \approx \frac{1-\tilde{r}^2}{\varepsilon^2} t_\beta. \tag{5.3.11}$$

Следует обратить внимание на то, что во всех выражениях (5.3.8)–(5.3.11) реальными являются лишь приближённые равенства, так как на этапе оценивания (как точечного, так и интервального) истинное значение коэффициента корреляции неизвестно. В связи с этим потребный объём выборки может быть определён лишь методом последовательных приближений. Сущность этого метода была раскрыта в п.п.5.2.3.

П р и м е р 5.10. Пусть $\hat{x}$, $\hat{y}$ – координаты пробоины в мишени. Произведено 40 выстрелов. Коэффициент корреляции случайных величин $\hat{x}$ и $\hat{y}$ составил $r = 0{,}605$. Найти:

1) доверительную вероятность β для r, если максимальная с вероятностью β абсолютная погрешность оценки $\tilde{r}$ должна быть не более 0,1;

2) доверительный интервал для r при доверительной вероятности β = 0,85.

▼ 1. По формуле (5.3.8) находим

$$\beta = \beta_{0,1;\,40} = 2\Phi_0\left(\frac{0{,}1\sqrt{40}}{\sqrt{1-(0{,}605)^2}}\right) = 2\Phi_0(0{,}79) = 0{,}5704.$$

Значение функции $\Phi_0(x)$ – в приложении 2.

2. Для β = 0,85 в приложении 4 находим $t_\beta = t_{0,85} = 1{,}439$.

Для вычисления погрешности ε используем выражение (5.3.9):

$$\varepsilon = \varepsilon_{0,85;\,40} = \frac{\sqrt{1-(0{,}65)^2}}{\sqrt{40}}1{,}439 = 0{,}181.$$

В соответствии с выражением (5.3.10) доверительный интервал

$I = I_{0,85;\,40} = [0{,}605–0{,}181;\ \ 0{,}605+0{,}181] = [0{,}424;\ \ 0{,}786]$.

▲

5.4. Оценивание числовых характеристик случайных функций

Известно, что случайная функция может рассматриваться как обобщение понятия случайного вектора (системы случайных величин) на бесконечное множество составляющих его компонентов (сечений случайной функции). Исчерпывающего вероятностного описания такого случайного объекта не существует, поэтому на практике используются лишь законы распределения и числовые характеристики систем конечного числа сечений случайных функций. При этом из-за сложности построения статистических законов распределения многомерных случайных векторов наиболее широкое применение получила корреляционная теория, в рамках которой изучаются лишь первые и вторые моменты распределений случайных функций, т.е. их математические ожидания, дисперсии и корреляционные функции.

5.4.1. Нестационарные случайные функции

Реализации $x_i(t)$, $i=\overline{1,n}$ случайной функции $\hat{x}(t)$ представляют собой неслучайные функции, значения которых $x_i(t_j)$ в фиксированных точках t_j, j = 1,2,… являются реализациями x_{ij} случайных величин $\hat{x}_j=\hat{x}(t_j)$.

Пусть над случайной функцией $\hat{x}(t)$ произведено n независимых равноточных наблюдений (опытов), в результате которых получено n её реализаций $x_i(t)$, $i=\overline{1,n}$ (рис.5.3).

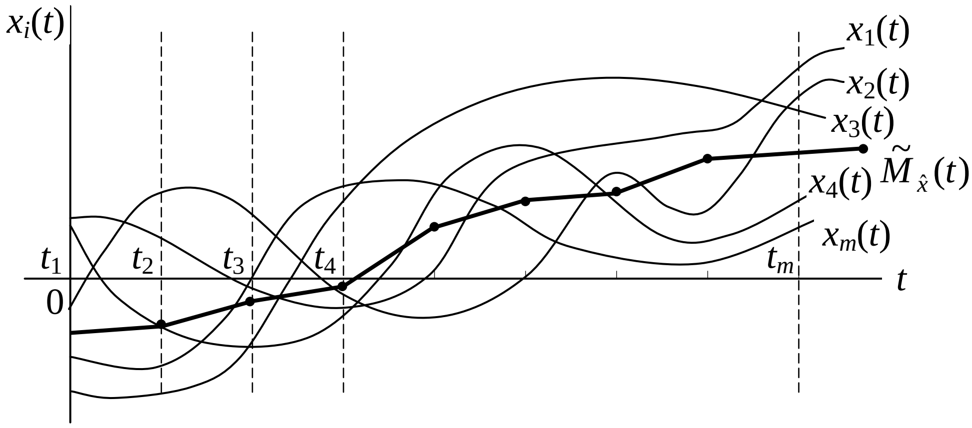

Рис.5.3. Реализации случайной функции

Требуется найти оценки числовых характеристик случайной функции: математического ожидания $M_{\hat{x}}(t)=\overline{x}(t)$, дисперсии $D_{\hat{x}}(t)=\overline{\dot{x}^2}(t)$ и корреляционной функции $K_{\hat{x}}(t';t'')=\overline{\dot{x}(t')\dot{x}(t'')}$, удовлетворяющие требованиям состоятельности, несмещённости и эффективности (асимптотической эффективности).

В ряде сечений случайной функции, соответствующих моментам времени t_1, t_2,…, t_j,…,t_m, фиксируются значения, принятые реализациями $x_i(t)$ функции $\hat{x}(t)$ в эти моменты. Поскольку наблюдалось n реализаций, то каждому из моментов t_j, $j=\overline{1,m}$ будут соответствовать n значений, принятых случайной величиной $\hat{x}_j=\hat{x}(t_j)$. Указанная случайная величина является j-м сечением случайной функции $\hat{x}(t)$. Расстояния

$$h = h_j = t_{j+1} - t_j$$

между фиксируемыми сечениями $\hat{x}(t_j)$ случайной функции $\hat{x}(t)$ обычно берутся одинаковыми и назначаются так, чтобы последовательность $x_i(t_j)$, $j=\overline{1,m}$ позволяла восстановить основной характер зависимости $x_i(t)$. Нередко в основу выбора кладётся теорема В.А. Котельникова, согласно которой для точного восстановления непрерывной функции достаточно её наблюдать в равноотстоящих дискретных точках с частотой, в два раза превышающей максимум её частотного спектра [2,13]. Бывает, что приведённые соображения являются излишними и расстояние задаётся темпом работы регистрирующей аппаратуры.

Для удобства последующей статистической обработки зарегистрированные данные сводятся в таблицу, строки которой соответствуют реализациям, а столбцы – сечениям случайной функции (табл.5.6).

Т а б л и ц а 5.6

Зарегистрированные значения случайной функции

$x_i(t)$	t							
	t_1	t_2	…	t_j	…	t_l	…	t_m
$x_1(t)$	$x_1(t_1)$	$x_1(t_2)$	…	$x_1(t_j)$	…	$x_1(t_l)$	…	$x_1(t_m)$
$x_2(t)$	$x_2(t_1)$	$x_2(t_2)$	…	$x_2(t_j)$	…	$x_2(l_j)$	…	$x_2(t_m)$
…	…	…	…	…	…	…	…	…
$x_i(t)$	$x_i(t_1)$	$x_i(t_2)$	…	$x_i(t_j)$	…	$x_i(t_l)$	…	$x_i(t_m)$
…	…	…	…	…	…	…	…	…
$x_n(t)$	$x_n(t_1)$	$x_n(t_2)$	…	$x_n(t_j)$	…	$x_n(t_l)$	…	$x_n(t_m)$

Приведённая в табл.5.6 совокупность значений случайной функции $\hat{x}(t)$ представляет собой результаты n наблюдений m-мерного случайного вектора

$$\hat{X}_{<m>} = \left(\hat{x}(t_1),\ \hat{x}(t_2),\dots,\hat{x}(t_m)\right)^{\text{T}}$$

и обрабатывается по методике § 5.3.

Так, оценки математических ожиданий сечений $\hat{x}(t_j)$ случайной функции $\hat{x}(t)$ находятся по формулам

$$\widetilde{M}_{\hat{x}}(t_j) = M^*_{\hat{x}}(t_j) = \frac{1}{n}\sum_{i=1}^{n} x_i(t_j)\,, \quad j = \overline{1,m}\,. \tag{5.4.1}$$

Соединяя точки $\left(t_j;\,\widetilde{M}_{\hat{x}}(t_j)\right)$ отрезками прямых, можно построить приближённый график (рис.5.3) оценки $\widetilde{M}_{\hat{x}}(t)$ математического ожидания $M_{\hat{x}}(t)$ случайной функции $\hat{x}(t)$. Очевидно, что возможны и другие виды интерполяции, например, квадратичная.

Несмещенные оценки дисперсий и корреляционных моментов сечений определяются соответственно следующими соотношениями:

$$\widetilde{D}_{\hat{x}}(t_j) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i(t_j) - \widetilde{M}_{\hat{x}}(t_j))^2\,, \quad j = \overline{1,m}\,. \tag{5.4.2}$$

$$\widetilde{K}_{\hat{x}}(t_j;t_l) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i(t_j) - \widetilde{M}_{\hat{x}}(t_l))(x_i(t_l) - \widetilde{M}_{\hat{x}}(t_l)), \quad j,l = \overline{1,m}\,. \tag{5.4.3}$$

Легко заметить, что формула (5.4.2) может быть получена и из выражения (5.4.3) при $l = j$, поскольку $D_{\hat{x}}(t_j) = K_{\hat{x}}(t_j;t_j)$.

В вычислительном отношении более удобны формулы, основанные на связи начальных и центральных моментов, т.е.

$$\widetilde{D}_{\hat{x}}(t_j)=\left(\frac{1}{n}\sum_{i=1}^{n}(x_i^2(t_j)-\widetilde{M}_{\hat{x}}^2(t_j)\right)\frac{n}{n-1},\quad j=\overline{1,m}. \qquad (5.4.4)$$

$$\widetilde{K}_{\hat{x}}(t_j;t_l)=\left(\frac{1}{n}\sum_{i=1}^{n}x_i(t_j)x_i(t_l)-\widetilde{M}_{\hat{x}}(t_l)\widetilde{M}_{\hat{x}}(t_l)\right)\frac{n}{n-1},\quad j,l=\overline{1,m}. \qquad (5.4.5)$$

При практическом использовании формул (5.4.4) и (5.4.5) рекомендуется начало отсчёта значений случайной функции перенести ближе к её математическому ожиданию. Это позволит избежать вычислений разности близких чисел.

П р и м е р 5.11. Результаты наблюдения 11 реализаций случайной функции $\hat{x}(t)$ в момент времени $t_j=\overline{0,10}$ с приведены в табл.5.7.

Т а б л и ц а 5.7

Реализации случайной функции

$x_i(t)$	t										
	0	1	2	3	4	5	6	7	8	9	10
$x_1(t)$	0,7	1,2	2,0	3,2	4,7	6,0	6,4	6,6	6,3	5,6	5,0
$x_2(t)$	1,2	2,0	3,6	4,6	5,1	5,5	6,0	6,2	6,2	6,0	6,0
$x_3(t)$	2,0	3,3	4,1	4,4	4,5	4,5	4,8	5,5	6,0	6,3	6,2
$x_4(t)$	2,5	2,9	3,0	3,2	3,8	4,7	5,4	5,5	5,4	5,7	6,2
$x_5(t)$	2,7	3,8	4,7	5,1	5,3	5,2	5,0	4,9	5,1	5,7	6,6
$x_6(t)$	3,2	3,9	4,1	4,1	4,0	4,2	5,0	6,0	6,3	6,1	5,7
$x_7(t)$	3,8	4,5	5,0	5,4	5,5	5,6	5,5	5,5	5,2	5,0	4,9
$x_8(t)$	4,1	3,8	3,6	3,9	4,8	5,8	6,2	6,1	5,7	5,4	5,4
$x_9(t)$	4,2	4,9	5,0	4,6	4,3	4,0	4,2	4,7	5,7	6,6	6,8
$x_{10}(t)$	5,4	4,3	3,2	2,9	3,1	3,8	4,5	5,5	6,4	7,0	7,2
$x_{11}(t)$	5,8	5,6	5,4	5,2	4,8	4,5	4,3	4,3	4,4	4,5	4,8

Требуется определить оценки числовых характеристик случайной функции $\hat{x}(t)$.

▼ По формуле (5.4.1) вычисляются оценки $\widetilde{M}_{\hat{x}}(t_j)$ и результаты сводятся в табл.5.8.

Т а б л и ц а 5.8

Оценки математических ожиданий сечений случайной функции

t_j	0	1	2	3	4	5	6	7	8	9	10
$\widetilde{M}_{\hat{x}}(t_j)$	3,2	3,7	4,0	4,2	4,5	4,9	5,2	5,5	5,7	5,8	5,9

По формуле (5.4.3) или (5.4.5) вычисляются оценки $\widetilde{K}_{\hat{x}}(t_j;t_l)$ и результаты сводятся в табл.5.9, диагональные элементы $\widetilde{K}_{\hat{x}}(t_j;t_j)$ которой представляют собой оценки $\widetilde{D}_{\hat{x}}(t_j)$ дисперсий $D_{\hat{x}}(t_j)$ в сечениях $\hat{x}(t_j)$ случайной функции $\hat{x}(t)$.

В рассмотренном примере пришлось 11 производить вычисления по формуле (5.4.1) и 66 раз – по формуле (5.4.3) или (5.4.5). Это свидетельствует о большой трудоёмкости задачи оценивания вероятностных характеристик нестационарных случайных функций.

Таблица 5.9

Корреляционные моменты случайной функции

t_l	t_j										
	0	1	2	3	4	5	6	7	8	9	10
0	2,6	1,9	0,9	0,2	–0,3	–0,7	–0,8	–0,7	–0,4	–0,1	–0,1
1		1,6	1,1	0,5	–0,07	–0,5	–0,7	–0,7	–0,5	–0,1	0,07
2			1,0	0,7	0,3	–0,06	–0,5	–0,6	–0,4	–0,2	–0,05
3				0,7	0,7	0,09	–0,2	–0,3	–0,3	–0,3	–0,2
4					0,5	0,4	0,2	–0,03	–0,2	–0,3	–0,3
5						0,6	0,5	0,3	0	–0,3	–0,4
6							0,6	0,4	0,1	–0,1	–0,3
7								0,5	0,3	–0,2	–0,4
8									0,4	0,3	0,2
9										0,5	0,5
10											0,7

▲

5.4.2. Стационарные случайные функции

По определению, случайная функция $\hat{u}(t)$ является стационарной (в широком смысле), если её математическое ожидание и дисперсия постоянны, а корреляционная функция зависит лишь от расстояния между сечениями случайной функции:

$$M_{\hat{u}(t)} = \overline{u(t)} = M_{\hat{u}} = \text{const};$$

$$D_{\hat{u}}(t) = \overline{\dot{u}^2(t)} = D_{\hat{u}} = \text{const};$$

$$K_{\hat{u}}(t';t'') = \overline{\dot{u}(t')\dot{u}(t'')} = K_{\hat{u}}(t';t'+\tau) = K_{\hat{u}}(\tau).$$

Класс стационарных случайных функций достаточно многообразен. Однако в практическом отношении наибольший интерес представляют стационарные случайные функции, обладающие эргодическим свойством, для которых одна реализация достаточно большой продолжительности содержит о случайной функции столько же информации, сколько её содержит и множество реализаций той же суммарной продолжительности. Другими словами, каждая из реализаций эргодической стационарной случайной функции является представителем всего их ансамбля. Как отмечено в работе [12], следует различать эргодические свойства случайных функций по отношению к моментам их распределения различных порядков. При этом под эргодичными обычно понимаются случайные функции, обладающие такими свойствами по отношению к моментам первого и второго порядков, т.е. к математическому ожиданию и корреляционной функции (следовательно, и к дисперсии). Далее рассматриваются только такие случайные функции.

Оценки числовых характеристик эргодичных случайных функций могут быть приближённо определены не как средние по множеству реализаций, а как средние по времени T наблюдения одной реализации по следующим формулам:

$$\tilde{M}_{\hat{u}} = \tilde{M}[\hat{u}(t)] \approx \frac{1}{T}\int_0^T u(z)dz, \quad t \in [0; T]; \qquad (5.4.6)$$

$$\tilde{K}_{\hat{u}}(\tau) = \tilde{M}[\dot{u}(t)\dot{u}(t+\tau)] \approx \frac{1}{T-\tau}\int_0^{T-\tau} \dot{u}(z)\dot{u}(z+\tau)dz, \quad t \in [0; T-\tau]; \qquad (5.4.7)$$

$$\tilde{D}_{\hat{u}} = \tilde{D}[\hat{u}(t)] = \tilde{K}_{\hat{u}}(0) \approx \frac{1}{T}\int_0^T \dot{u}^2(z)dz, \quad t \in [0; T], \qquad (5.4.8)$$

где $\dot{u}(t) = u(t) - \tilde{M}_{\hat{u}}$.

Обоснованием применимости формул (5.4.6) – (5.4.8) служит тот факт, что для эргодичных стационарных случайных функций средние во времени оценки сходятся по вероятности к оцениваемым ими характеристикам $M_{\hat{u}}$, $K_{\hat{u}}(\tau)$, $D_{\hat{u}}$.

Из выражений (5.4.6) – (5.4.8) видно, что для их практического применения требуется интегрировать ряд функций от реализации $u(t)$ случайной функции $\hat{u}(t)$. Чаще всего на практике для нахождения оценок (5.4.6) – (5.4.8) используется следующая методика.

Пусть на интервале времени $[0; T]$ наблюдалась реализация эргодичной стационарной случайной функции $\hat{u}(t)$, значения которой $u(t_j)$ в ряде равноотстоящих опорных моментов времени $t_j = 0{,}5h + kh$, $j = \overline{1, m}$, $k = \overline{0, m-1}$ зарегистрированы и сведены в табл.5.10.

Таблица 5.10

Значения реализации эргодичной стационарной функции в выбранные моменты времени

t	t_1	t_2	…	t_j	…	t_m
$u(t)$	$u(t_1)$	$u(t_2)$	…	$u(t_j)$	…	$u(t_m)$

Требуется по данным этой таблицы определить оценки $\tilde{M}_{\hat{u}}$, $\tilde{K}_{\hat{u}}(\tau)$, $\tilde{D}_{\hat{u}}$ числовых характеристик $M_{\hat{u}}$, $K_{\hat{u}}(\tau)$, $D_{\hat{u}}$ случайной функции $\hat{u}(t)$

Интервал $[0; T]$ наблюдения случайной функции $\hat{u}(t)$ разбивается на m равных подынтервалов длиной $h = T/m$, расположенных симметрично относительно опорных моментов времени $t_1, t_2, \ldots, t_m$ (рис.5.4).

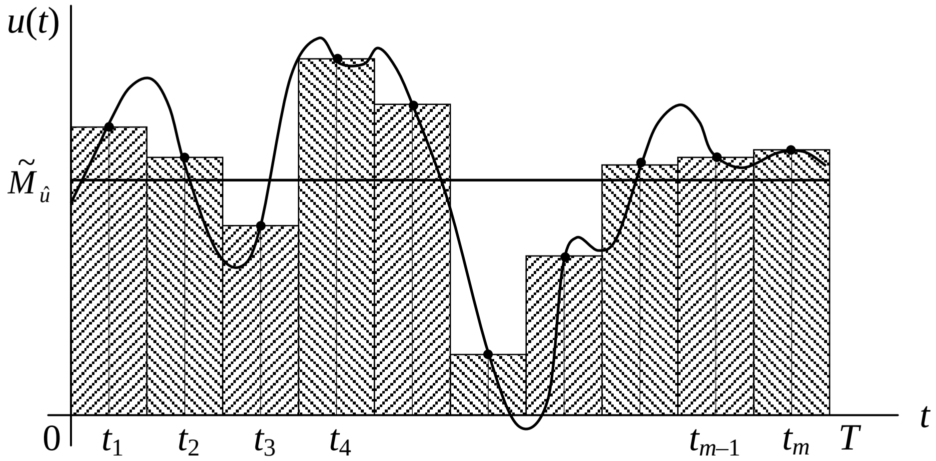

Рис.5.4. Реализация эргодичной стационарной случайной функции

Далее предполагается, что в пределах подынтервала $[t_j–0,5h;\ t_j+0,5h]$ функция, описывающая реализацию случайной функции, постоянна:

$$u(t) = u(t_j) = \text{const},\quad t \in [t_j - 0,5h;\ t_j + 0,5h],\quad j = \overline{1,m},$$

Если h достаточно мало, то можно приближённо полагать, что

$$\int\limits_{t_j-0,5h}^{t_j+0,5h} u(t)dt = hu(t_j),\quad j = \overline{1,m}. \tag{5.4.9}$$

Суммирование результатов (5.4.9) по j даёт

$$M_{\hat{u}} \approx \frac{1}{T}\int\limits_0^T u(t)dt = \frac{h}{T}\sum_{j=1}^{m} u(t_j) = \frac{1}{m}\sum_{j=1}^{m} u(t_j). \tag{5.4.10}$$

Несложно заметить, что выражение (5.4.10) реализует процедуру численного интегрирования по формуле прямоугольников.

Аналогично вычисляется и оценка корреляционной функции для значений аргумента

$$\tau_l = lh = \frac{lT}{m}.$$

Поскольку в выражении (5.4.7) длина интервала интегрирования

$$T - \tau_l = T - \frac{lT}{m} = \frac{m-l}{m}T,$$

то, поделив его на $m - l$ равных участков и вынося на каждом из них за знак интеграла среднее значение функции $\dot{u}(t)\dot{u}(t+\tau)$, получим

$$\tilde{K}_{\hat{u}}(\tau_l) = \tilde{K}_{\hat{u}}\left(\frac{lT}{m}\right) \approx \frac{1}{m-l}\sum_{j=1}^{m-l}\dot{u}(t_j)\dot{u}(t_{j+l}),\quad l = \overline{0,L}, \tag{5.4.11}$$

где $L = m/4$ [2].

При $\tau = 0$ формула (5.4.11) даёт

$$\widetilde{D}_{\hat{u}} = \widetilde{K}_{\hat{u}}(0) \approx \frac{1}{m}\sum_{j=1}^{m} \dot{u}^2(t_j). \qquad (5.4.12)$$

Вычисления по формуле (5.4.11) ведутся последовательно для $l = 0, 1, 2,\ldots$ вплоть до таких значений l_k, при которых функция $\widetilde{K}_{\hat{u}}(lT/m)$ становится практически равной нулю или начинает совершать незначительные колебания около нуля. По полученным точкам может быть построен приближённый график корреляционной функции $K_{\hat{u}}(\tau)$ (рис.5.5).

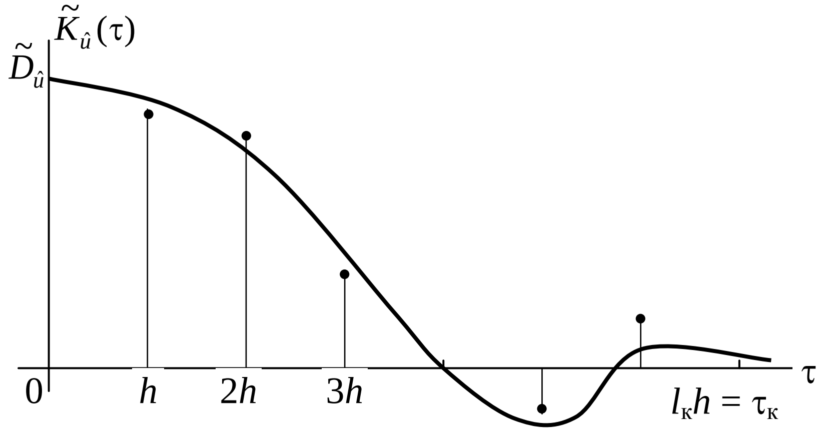

Рис.5.5. Приближённый график корреляционной функции

На представленном рисунке через $\tau_к$ обозначена длина интервала корреляции случайной функции $\hat{u}(t)$, т.е. наименьшее расстояние между сечениями случайной функции, на котором корреляция между ними практически отсутствует ($\tau_к = l_к h$).

П р и м е р 5.12. В табл.5.11 приведены результаты наблюдения эргодичной стационарной случайной функции $\hat{u}(t)$ на интервале времени продолжительностью $T = 28$ с с периодичностью $h = 1$ с в моменты времени t_j, $j = \overline{1, 28}$.

Т а б л и ц а 5.11

Реализация эргодичной стационарной случайной функции

t_j, с	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}
$u(t_j)$	4,0	4.1	3,8	2,0	1,0	–0,3	–0,2	1,0	3,5	2,5	2,6	2,0	0,6	2,5
t_j, с	t_{15}	t_{16}	t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	t_{24}	t_{25}	t_{26}	t_{27}	t_{28}
$u(t_j)$	3,3	3,8	1,2	0,5	–0,5	0,5	3,2	4,0	2,3	0,5	2,4	3,2	2,0	0,5

Требуется определить оценки числовых характеристик функции $\hat{u}(t)$.

▼ Необходимые вычисления производятся по формулам (5.4.10–(5.4.12). Результаты расчётов оформлены в виде табл.5.12, наглядно иллюстрирующей все этапы решения задачи.

▲

Решение рассмотренной задачи потребовало однократного использования формулы (5.4.10) и одиннадцатикратного – формулы (5.4.11). Если интервал корреляции $\tau_к$ случайной функции соизмерим с интервалом T её наблюдения, то аргумент τ корреляционной функции должен варьироваться от 0 до T. В этом случае параметр l в формуле (5.4.11) пробегает значения от 0 до $m-1$ и, следовательно, формула (5.4.11) реализуется m раз.

Расчёты показывают, что по сравнению с нестационарными случайными функциями трудоёмкость оценивания числовых характеристик стационарных эргодичных случайных функций существенно снижается.

5.4.3. Качество оценивания числовых характеристик случайных функций

Известно, что сечения случайной функции $\hat{x}(t)$ представляют собой обычные случайные величины, а совокупности сечений – случайные векторы, вероятностные характеристики которых (законы и параметры распределения) зависят от значений t_j, $j=\overline{1,m}$ параметра t.

Выше отмечалось, что формально случайная функция может интерпретироваться как обобщение понятия случайного вектора на случай бесконечного множества его компонентов. Однако на практике используются законы распределения случайных величин ограниченного и, как правило, невысокого порядка. Такая постановка эквивалентна моделированию случайной функции $\hat{x}(t)$ случайным вектором

$$\hat{X}_{<m>}(T_{<m>})=\left(\hat{x}(t_1),\ \hat{x}(t_2),\dots,\hat{x}(t_m)\right)^{\mathrm{т}}=(\hat{x}_1,\ \hat{x}_2,\dots,\hat{x}_m)^{\mathrm{т}}=\hat{X}_{<m>}. \quad (5.4.13)$$

Таким образом, в основе подходов к оцениванию числовых характеристик случайных функций лежат методы оценивания числовых характеристик случайных величин и векторов, рассмотренные ранее. Поэтому ниже рассматриваются лишь некоторые особенности, присущие только случайным функциям.

Таблица 5.12

Расчётная таблица оценок числовых характеристик стационарной эргодичной случайной функции

j	$u(t_j)$	$u(t_j)-\tilde{M}_{\hat{u}}$	$\dot{u}(t_j)\dot{u}(t_j+\tau_l)=(u(t_j)-\tilde{M}_{\hat{u}})(u(t_j+\tau_l)-\tilde{M}_{\hat{u}})$										
			l										
			0	1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	4,0	2,0	4,00	4,20	3,60	0,00	–2,00	–4,60	–4,40	–2,00	3,00	1,00	1,20
2	4,1	2,1	4,41	3,78	0,00	–2,10	–4,83	–4,62	–2,10	3,15	1,05	1,26	0,00
3	3,8	1,8	3,24	0,00	–1,80	–4,14	–3,96	–1,80	2,70	0,90	1,08	0,00	–2,52
4	2,0	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5	1,0	–1,0	1,00	2,30	2,20	1,00	–1,50	–0,50	–0,60	0,00	1,40	–0,50	–1,30
6	–0,3	–2,3	5,29	5,06	2,30	–3,45	–1,15	–1,38	0,00	3,22	–1,15	–2,99	–4,14
7	–0,2	–2,2	4,84	2,20	–3,30	–1,10	–1,32	0,00	3,08	–1,10	–2,86	–3,96	1,76
8	1,0	–1,0	1,00	–1,50	–0,50	–0,60	0,00	1,40	–0,50	–1,30	–1,80	0,80	1,50
9	3,5	1,5	2,25	0,75	0,90	0,00	–2,10	0,75	1,95	2,70	–1,20	–2,25	–3,75
10	2,5	0,5	0,25	0,30	0,00	–0,70	0,25	0,65	0,90	–0,40	–0,75	–1,25	–0,75
11	2,6	0,6	0,36	0,00	–0,84	0,30	0,78	1,08	–0,48	–0,90	–1,50	–0,90	0,72
12	2,0	0,0	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00
13	0,6	–1.4	1,96	–0,70	–1,82	–2,52	1,12	2,10	3,50	2,10	–1,68	–2,80	–0,42
14	2,5	0,5	0,25	0,65	0,90	–0,40	-0,75	–1,25	–0,75	0,60	1,00	0,15	–0,75
15	3,3	1,3	1,69	2,34	–1,04	–1,95	–3,25	–1,95	1,56	2,60	0,39	–1,95	0,52
16	3,8	1,8	3,24	–1,44	–2,70	–4,50	–2,70	2,16	3,60	0,54	–2,70	0,72	2,16
17	1,2	–0,8	0,64	1,20	2,00	1,20	–0,96	–1,60	–0,24	1,20	–0,32	–0,96	0,00
18	0,5	–1,5	2,25	3,75	2,25	–1,80	–3,00	–0,45	2,25	–0,60	–1,80	0,00	2,25
19	–0,5	–2,5	6,25	3,75	–3,00	–5,00	–0,75	3,75	–1,00	–3,00	0,00	3,75	–
20	0,5	–1,5	2,25	–1,80	–3,00	–0,45	2,25	–0,60	–1,80	0,00	2,25	–	–
21	3,2	1,2	1,44	2,40	0,36	–1,80	0,48	1,44	0,00	–1,80	–	–	–
22	4,0	2,0	4,00	0,60	–3,00	0,80	2,40	0,00	–3,00	–	–	–	–
23	2,3	0,3	0,09	–0,45	0,12	0,36	0,00	–0,045	–	–	–	–	–
24	0,5	–1,5	2,25	–0,60	–1,80	0,00	2,25	–	–	–	–	–	–

j	$u(t_j)$	$u(t_j)-\tilde{M}_{\hat{u}}$	$\dot{u}(t_j)\dot{u}(t_j+\tau_l)=(u(t_j)-\tilde{M}_{\hat{u}})(u(t_j+\tau_l)-\tilde{M}_{\hat{u}})$										
			l										
			0	1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10	11	12	13	14
25	2,4	0,4	0,16	0,48	0,00	–0,60	–	–	–	–	–	–	–
26	3,2	1,2	1,44	0,00	–1,80	–	–	–	–	–	–	–	–
27	2,0	0,0	0,00	0,00	–	–	–	–	–	–	–	–	–
28	0,5	–1,5	2,25	–	–	–	–	–	–	–	–	–	–
Σ	56	0,0	56,80	27,27	–9,97	–27,45	–18,74	–4,87	4,67	5,91	–5,59	–9,88	–3,52
$\tilde{M}_{\hat{u}}=\frac{1}{28}\cdot 56=2$		$m–l$	28	27	26	25	24	23	22	21	20	19	18
		$\tilde{K}_{\hat{u}}(\tau_l)$	2,03	1,01	–0,38	–1,10	–0,78	–0,21	0,21	0,28	–0,28	–0,52	–0,19

Числовые характеристики $M_{\hat{x}}(t)$, $D_{\hat{x}}(t)$, $K_{\hat{x}}(t_1;t_2)$ случайной функции $\hat{x}(t)$ зависят от параметра *t*, следовательно, от него зависят как оценки $\tilde{M}_{\hat{x}}(t)$, $\tilde{D}_{\hat{x}}(t)$, $\tilde{K}_{\hat{x}}(t_1;t_2)$, так и характеристики их качества – доверительные интервалы:

$$I_{\beta_1,n} = [M'_{\hat{x}}(t); M''_{\hat{x}}(t)]; \tag{5.4.14}$$

$$I_{\beta_2,n} = [D'_{\hat{x}}(t); D''_{\hat{x}}(t)]; \tag{5.4.15}$$

$$I_{\beta_3,n} = [K'_{\hat{x}}(t_1;t_2); K''_{\hat{x}}(t_1;t_2)], \tag{5.4.16}$$

а также доверительные вероятности:

$$\beta_{I_1,n} = P(M'_{\hat{x}}(t) \le M_{\hat{x}}(t) \le M''_{\hat{x}}(t)); \tag{5.4.17}$$

$$\beta_{I_2,n} = P(D'_{\hat{x}}(t) \le D_{\hat{x}}(t) \le D''_{\hat{x}}(t)); \tag{5.4.18}$$

$$\beta_{I_3,n} = P(K'_{\hat{x}}(t_1;t_2) \le K_{\hat{x}}(t_1;t_2) \le K''_{\hat{x}}(t_1;t_2)), \tag{5.4.19}$$

Во всех приведённых выражениях *n* – это число наблюдаемых реализаций случайной функции $\hat{x}(t)$. Если случайная функция $\hat{u}(t)$ стационарна, то выражения (5.4.14)–(5.4.19) упрощаются и принимают следующий вид:

$$I_{\beta_1,T} = [M'_{\hat{u}}; M''_{\hat{u}}];$$

$$I_{\beta_2,T} = [D'_{\hat{u}}; D''_{\hat{u}}];$$

$$I_{\beta_3,T} = [K'_{\hat{u}}(\tau); K''_{\hat{u}}(\tau)],$$

$$\beta_{I_1,T} = P(M'_{\hat{u}} \le M_{\hat{u}} \le M''_{\hat{u}});$$

$$\beta_{I_2,T} = P(D'_{\hat{u}} \le D_{\hat{u}} \le D''_{\hat{u}});$$

$$\beta_{I_3,T} = P(K'_{\hat{u}}(\tau) \le K_{\hat{u}}(\tau) \le K''_{\hat{u}}(\tau)),$$

где *T* – продолжительность наблюдения случайной функции $\hat{u}(t)$.

5.4.4. Потребный объём наблюдений

При оценивании числовых характеристик нестационарной случайной функции задача состоит в определении потребного числа *n* её реализаций $x_i(t)$, $i = \overline{1, n}$. Поскольку все числовые характеристики случайной функции, а также характеристики (5.4.14) – (5.4.19) качества оценивания зависят от параметра *t*, то от него будет зависеть и потребный объём *n* выборки реализаций:

$$n_1 = n_{\beta_1,I_1} = n_1(t); \tag{5.4.20}$$

$$n_2 = n_{\beta_2,I_2} = n_2(t); \tag{5.4.21}$$

$$n_3 = n_{\beta_3,I_3} = n_3(t_1; t_2). \tag{5.4.22}$$

Выражениями (5.4.20)–(5.4.22) определяется потребный объём реализаций случайной функции для оценки соответственно математического ожидания, дисперсии и корреляционной функции. Из данных выражений следует, что оценивание числовых характеристик случайной функции для различных её сечений может потребовать различного числа реализаций. Такое требование практически неосуществимо. Поэтому требуемый объём n_k, $k = 1, 2, 3$ определяется как $\max\limits_{t}\{n_k(t)\}$, соответствующий минимуму дисперсии оценки k-й числовой характеристики случайной функции:

$$n_1 = \max_{t\in[0;T]}\{n_1(t)\};$$

$$n_2 = \max_{t\in[0;T]}\{n_2(t)\};$$

$$n_3 = \max_{t_1,t_2\in[0;T]}\{n_3(t_1;t_2)\}.$$

При оценивании числовых характеристик стационарной случайной функции $\hat{u}(t)$ наблюдается всего одна её реализация $u(t)$. Поэтому здесь возникает вопрос о длительности наблюдения T, потребной для оценивания характеристик случайной функции с заданной точностью I или ε и надёжностью β.

Поскольку случайная функция $\hat{u}(t)$ стационарна, то все её сечения распределены одинаково. По этой причине последовательность значений $u(t_j)$, $j = \overline{1, m}$ наблюдаемой реализации $u(t)$ случайной функции $\hat{u}(t)$ может интерпретироваться как однородная выборка $u(t_1)$, $u(t_2)$,…, $u(t_m)$, элементы которой принадлежат одной и той же генеральной совокупности. Теперь задача состоит в определении потребного объёма n этой выборки.

По методикам, изложенным в §§ 5.1 – 5.3, получим соотношения:

$$n_1 = n_{\beta_1, I_1};$$

$$n_2 = n_{\beta_2, I_2};$$

$$n_3 = n_{\beta_3, I_3} = \max_{\tau\in[0;T]}\{n_{\beta_3, I_3}(\tau)\}.$$

Так как наблюдается всего одна реализация случайной функции, то потребное число n её измерений определяется следующим соотношением:

$$n = \max\{n_1; n_2; n_3\}. \tag{5.4.23}$$

Поскольку регистрация значений реализации $u(t)$ стационарной случайной функции $\hat{u}(t)$ обычно производится через равные промежутки времени длительностью h, то потребная длительность наблюдения реализации определяется равенством

$$T = nh,$$

где n находится из соотношения (5.4.23).