

2. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПАРАМЕТРИЧЕСКИХ МЕТОДОВ ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

2.1. Основные понятия

Параметрические методы обработки экспериментальных данных опираются на основополагающий факт, в соответствии с которым свойства результатов экспериментальных исследований, рассматриваемых как случайные объекты, описываются некоторым законом распределения. При этом предполагается, что анализ экспериментальных данных позволяет с достаточной степенью точности определить вид и конкретную форму закона распределения или значения его параметров, если нет необходимости в использовании самого закона. Такая информация даёт возможность в полном объёме использовать методы теории вероятностей для решения задач обработки.

Так как действительный закон распределения и значения его параметров неизвестны, то параметрические методы оперируют с их приближениями – статистическими законами распределения и оценками параметров распределения.

Статистическим законом распределения случайной величины \hat{x} называется закон распределения данной величины, установленный с помощью статистических методов обработки данных.

Статистический закон распределения может быть определён в виде статистической функции распределения $F_{\hat{x}}^*(x)$, статистической плотности распределения $\phi_{\hat{x}}^*(x)$ или статистического ряда распределения $P^*(x_i)$, $i = \overline{1, n}$.

Статистическими оценками параметров закона распределения случайной величины называются приближённые значения данных параметров (статистики), полученные с помощью статистических методов обработки данных.

В дальнейшем статистические оценки для краткости называются просто оценками.

Если некоторый закон распределения характеризуется параметрами a_1, a_2, \dots, a_m , то их оценки будем обозначать в виде $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_m$. Наиболее распространёнными видами параметров законов распределения при обработке экспериментальных данных являются математическое ожидание $M_{\hat{x}}$, дисперсия $D_{\hat{x}}$ или среднее квадратическое отклонение

$\sigma_{\hat{x}}$, а для системы случайных величин – корреляционный момент $K_{\hat{x}_i \hat{x}_j}$ или коэффициент корреляции $r_{\hat{x}_i \hat{x}_j}$. Иногда используются центральные моменты третьего и четвертого порядков. Соответственно при обработке данных используются их статистические аналоги – оценки математического ожидания, корреляционного момента и т.д.

Таким образом, если имеется совокупность экспериментальных данных x_1, x_2, \dots, x_n , то и статистический закон распределения, например функция $F_{\hat{x}}^*(x)$, и оценки его параметров представляют собой некоторые функции этих данных:

$$F_{\hat{x}}^*(x) = \psi(x_1, x_2, \dots, x_n); \quad (2.1.1)$$

$$\tilde{a}_j = f_j(x_1, x_2, \dots, x_n), \quad j = \overline{1, m}. \quad (2.1.2)$$

Вид статистик ψ и f_j определяет качество оценок $F_{\hat{x}}^*(x)$ и \tilde{a}_j . В связи с этим возникает ряд проблем, основной из которых является проблема определения условий, при которых оценки (2.1.1) и (2.1.2) могут с требуемой достоверностью представлять теоретические законы распределения и их параметры. Эти условия формируются **предельными теоремами** теории вероятностей. Именно они служат тем фундаментом параметрических методов обработки экспериментальных данных, на основе которого могут быть получены подходящие оценки законов и параметров распределения наблюдаемых характеристик.

Вторая проблема состоит в выборе **достаточной статистики**, т.е. такой статистики, которая позволяет в конкретных условиях получать оценки заданного качества. Так как на основе результатов наблюдений x_1, x_2, \dots, x_n может быть образован большой спектр статистик (2.1.1) и (2.1.2), данная проблема сводится к выбору из них оптимальной в определённом смысле статистики. Решение проблемы осуществляется методами теории статистических решений.

Как видно из рис.1.1, к проблеме принятия решений при обработке экспериментальных данных сводится не только задача выбора достаточной статистики. Большинство задач обработки данных в разной степени может быть отнесено к задачам принятия решений. В связи с этим фундаментом параметрических методов обработки служат также принципы принятия статистических решений, на основе которых сформированы критерии принятия оптимальных в определённом смысле решений. Особую роль среди данных принципов играет принцип максимального правдоподобия и вытекающий из него для случая нормального закона распределения метод наименьших квадратов.

В настоящей брошюре рассматриваются вопросы параметрической обработки экспериментальных данных.

2.2. Предельные теоремы теории вероятностей

Использование параметрических методов обработки данных предполагает выявление условий, определяющих справедливость априорных предположений о виде закона распределения исследуемой случайной величины и свойствах его параметров. Эти условия формулируются в виде предельных теорем теории вероятностей. Ниже излагаются содержание и сущность теорем без доказательства, а также некоторые рекомендации по их практическому применению.

2.2.1. Теорема Ляпунова

В природе, как известно, широко распространён нормальный закон распределения. Практикой установлено, что этому закону подчиняются ошибки стрельбы и бомбометания, погрешности измерений, погрешности размеров деталей, изготавливаемых промышленными предприятиями, время безотказной работы многих устройств и т.д. Поэтому в процессе обработки экспериментальной информации часто выдвигается предположение о нормальном распределении исследуемой случайной величины. Однако иногда нормальный закон распределения применить нельзя. Ввиду этого необходимо точно знать, когда можно выдвинуть такое предположение и в каких случаях от него следует отказаться. Этому вопросу посвящена центральная предельная теорема и её разновидности (теоремы Ляпунова, Муавра-Лапласа).

Т Е О Р Е М А . Если последовательность независимых случайных величин $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n$ удовлетворяет условию Ляпунова

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n M[|\hat{x}_i - M_{\hat{x}_i}|^3]}{\sqrt{\left(\sum_{i=1}^n D_{\hat{x}_i}\right)^3}} = 0,$$

где $M[|\hat{x}_i - M_{\hat{x}_i}|^3]$ – третий абсолютный центральный момент, то последовательность случайных величин

$$\mathfrak{Y}_n = \frac{\sum_{i=1}^n (\hat{x}_i - M_{\hat{x}_i})}{\sqrt{\sum_{i=1}^n D_{\hat{x}_i}}}$$

сходится по распределению к случайной величине, имеющей нормальное распределение, т.е. существует предел

$$\lim_{n \rightarrow \infty} P \left(\frac{\sum_{i=1}^n (\hat{x}_i - M_{\hat{x}_i})}{\sqrt{\sum_{i=1}^n D_{\hat{x}_i}}} < \vartheta \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\vartheta} e^{-\frac{t^2}{2}} dt.$$

На практике часто пользуются случайными величинами, представляющими собой сумму независимых случайных величин:

$$\hat{z}_n = \sum_{i=1}^n \hat{x}_i = \hat{\vartheta}_n \sqrt{\sum_{i=1}^n D_{\hat{x}_i}} + \sum_{i=1}^n M_{\hat{x}_i}.$$

Поскольку случайная величина \hat{z}_n связана со случайной величиной $\hat{\vartheta}_n$ линейной зависимостью, то в пределе она также будет иметь нормальное распределение. Параметры данного распределения можно выразить с помощью теорем о числовых характеристиках:

$$\left. \begin{aligned} M_{\hat{z}_n} &= M \left[\sum_{i=1}^n \hat{x}_i \right] = \sum_{i=1}^n M_{\hat{x}_i}; \\ D_{\hat{z}_n} &= D \left[\sum_{i=1}^n \hat{x}_i \right] = \sum_{i=1}^n D_{\hat{x}_i}. \end{aligned} \right\}$$

Условие Ляпунова представляет собой требование малости слагаемых

$$\frac{\hat{x}_i - M_{\hat{x}_i}}{\sqrt{\sum_{i=1}^n D_{\hat{x}_i}}}$$

в сумме

$$\frac{\sum_{i=1}^n (\hat{x}_i - M_{\hat{x}_i})}{\sqrt{\sum_{i=1}^n D_{\hat{x}_i}}}.$$

Таким образом, сущность центральной предельной теоремы состоит в следующем: закон распределения суммы независимых случайных величин при неограниченном увеличении числа слагаемых приближается к нормальному, если случайные величины, входящие в сумму, имеют дисперсию одного и того же порядка и конечные математические ожидания. Это означает, что удельный вес каждого слагаемого стремится к нулю при увеличении числа слагаемых.

В реальных условиях любое случайное отклонение от закономерного протекания основного явления вызывается бесчисленным множеством случайных факторов, каждый из которых обычно оказывает малое влияние на суммарное воздействие, и часто эти факторы независимы или

слабо зависимы. Этим и объясняется широкое распространение нормального закона.

На практике теоремой Ляпунова пользуются и тогда, когда n сравнительно невелико. При суммировании непрерывных случайных величин, имеющих одинаковые симметричные законы распределения с одинаковыми числовыми характеристиками, эту теорему можно применять при $n \geq 8$. Если же суммируются случайные величины с различными несимметричными законами и различными числовыми характеристиками, то теоремой Ляпунова можно пользоваться только при числе слагаемых порядка сотни.

Практическое применение теоремы Ляпунова предполагает использование формул для определения вероятности попадания нормально распределённой случайной величины в интервал $[a; b)$. В данном случае можно воспользоваться следующими формулами:

$$P(a \leq \hat{z}_n < b) \approx \Phi_0\left(\frac{b - M_{\hat{z}_n}}{\sigma_{\hat{z}_n}}\right) - \Phi_0\left(\frac{a - M_{\hat{z}_n}}{\sigma_{\hat{z}_n}}\right); \quad (2.2.1)$$

$$P(a \leq \hat{z}_n < b) \approx \Phi_1\left(\frac{b - M_{\hat{z}_n}}{\sigma_{\hat{z}_n}}\right) - \Phi_1\left(\frac{a - M_{\hat{z}_n}}{\sigma_{\hat{z}_n}}\right), \quad (2.2.2)$$

где

$$\Phi_0 = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, \quad (2.2.3)$$

$$\Phi_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \quad (2.2.4)$$

называются функциями нормированного нормального распределения (т. е. распределения с параметрами $M_{\hat{z}} = 0$, $D_{\hat{z}} = 1$) или функциями Лапласа, они являются табличными (см. приложения 2 и 3).

Следует отметить, что формулами (2.2.1) и (2.2.2) можно пользоваться при выполнении условия

$$|z - M_{\hat{z}_n}| < 3\sigma_{\hat{z}_n}, \quad (2.2.5)$$

где z принимает значение a или b . Это требование вызвано тем, что за пределами интервала

$$[M_{\hat{z}_n} - 3\sigma_{\hat{z}_n}; M_{\hat{z}_n} + 3\sigma_{\hat{z}_n}]$$

могут быть существенные ошибки.

Пример 2.1. При обработке информации требуется сложить 1000 чисел, каждое из которых округлено с точностью до 0,01. Полагая, что ошибки округления подчинены равномерному закону распределения,

найти вероятность того, что суммарная ошибка округления не превысит 0,2.

▼ Обозначим через \hat{x}_i , $i = \overline{1, 1000}$ ошибку округления i -го числа, а через \hat{z}_n – суммарную ошибку округления ($n = 1000$).

Далее учитываем, что случайная величина, равномерно распределённая на интервале $[a; b]$ имеет математическое ожидание и дисперсию, которые определяются по формулам

$$M_{\hat{x}} = \frac{a+b}{2}, \quad D_{\hat{x}} = \frac{(b-a)^2}{12}.$$

Ошибки округления в данном случае распределены на интервале $[a; b] = [-0,005; 0,005]$,

следовательно,

$$M_{\hat{x}} = \frac{-0,005 + 0,005}{2} = 0, \quad D_{\hat{x}} = \frac{(0,005 - (-0,005))^2}{12} = \frac{(0,01)^2}{12}.$$

$$M_{\hat{z}_n} = \sum_{i=1}^{1000} M_{\hat{x}_i} = 0; \quad D_{\hat{z}_n} = \sum_{i=1}^{1000} \frac{(0,01)^2}{12} = 0,00833;$$

$$\sigma_{\hat{z}_n} = \sqrt{D_{\hat{z}_n}} = \sqrt{0,00833} = 0,0913.$$

Условие (2.2.5) соблюдается, поэтому

$$\begin{aligned} P(|\hat{z}_n| < 0,2) &= \Phi_0\left(\frac{0,2}{0,0913}\right) - \Phi_0\left(\frac{-0,2}{0,0913}\right) = \\ &= 2\Phi_0\left(\frac{0,2}{0,0913}\right) = 2\Phi_0(2,19) = 2 \cdot 0,4857 = 0,971. \end{aligned}$$

Последнее равенство вытекает непосредственно из формулы (2.2.1), в нём учтено, что функция $\Phi_0(x)$ является нечётной.



2.2.2. Теорема Муавра-Лапласа

Пусть производится n независимых испытаний, в каждом из которых событие A может появляться с одной и той же вероятностью. Тогда случайная величина \hat{z}_n , представляющая собой число появлений события A в n испытаниях, будет иметь биномиальное распределение [4]. Если число испытаний велико, то и случайная величина принимает большое число возможных значений. Пользоваться такой случайной величиной затруднительно из-за сложности вычислений. Поэтому целесообразно применять теорему Муавра-Лапласа, которая доказывает сходимость последовательности случайных величин, имеющих биномиальное распределение, к нормально распределённой случайной величине.

ТЕОРЕМА. Пусть \hat{z}_n число появлений события A в n независимых испытаниях, в каждом из которых вероятность этого события равна p . Тогда при $n \rightarrow \infty$ имеет место соотношение

$$\lim_{n \rightarrow \infty} P\left(\frac{\hat{z}_n - np}{\sqrt{npq}} < \vartheta\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\vartheta} e^{-\frac{t^2}{2}} dt$$

где $q = 1 - p$; $M_{\hat{z}_n} = np$, $D_{\hat{z}_n} = npq$ – соответственно математическое ожидание и дисперсия биномиально распределённой случайной величины.

Таким образом, нормированная случайная величина

$$\hat{\vartheta}_n = \frac{\hat{z}_n - np}{\sqrt{npq}} \quad (2.2.6)$$

согласно теореме Муавра-Лапласа в пределе будет подчиняться нормированному нормальному закону распределения. Отсюда вытекает приближённое равенство, справедливое при больших значениях n :

$$P(a \leq \hat{\vartheta}_n < b) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{\vartheta^2}{2}} d\vartheta. \quad (2.2.7)$$

Найдём вероятность попадания случайной величины \hat{z}_n в интервал $[m_1; m_2)$. Для этого подставим граничные точки m_1 и m_2 в формулу (2.2.6):

$$a = \frac{m_1 - np}{\sqrt{npq}}; \quad b = \frac{m_2 - np}{\sqrt{npq}}.$$

Выражение (2.2.7) принимает вид

$$P(m_1 \leq \hat{z}_n < m_2) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{\vartheta^2}{2}} d\vartheta.$$

Используя табличные функции (2.2.3) и (2.2.4), получаем следующие рабочие формулы:

$$P(m_1 \leq \hat{z}_n < m_2) \approx \Phi_0\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi_0\left(\frac{m_1 - np}{\sqrt{npq}}\right); \quad (2.2.8)$$

$$P(m_1 \leq \hat{z}_n < m_2) \approx \Phi_1\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi_1\left(\frac{m_1 - np}{\sqrt{npq}}\right); \quad (2.2.9)$$

Если n сравнительно мало и разность $|m - np|$ соизмерима с 0,5, то не безразлично, относятся ли граничные точки интервала $[m_1; m_2)$ к числу возможных значений \hat{z}_n или нет. В этом случае вместо z_1 и z_2 следует брать $m_1 - 0,5$, $m_2 - 0,5$. Тогда соотношения (2.2.8) и (2.2.9) примут вид

$$P(m_1 \leq \hat{z}_n < m_2) \approx \Phi_0\left(\frac{m_2 - 0,5 - np}{\sqrt{npq}}\right) - \Phi_0\left(\frac{m_1 - 0,5 - np}{\sqrt{npq}}\right); \quad (2.2.10)$$

$$P(m_1 \leq \hat{z}_n < m_2) \approx \Phi_1\left(\frac{m_2 - 0,5 - np}{\sqrt{npq}}\right) - \Phi_1\left(\frac{m_1 - 0,5 - np}{\sqrt{npq}}\right). \quad (2.2.11)$$

Следует отметить, что формулы (2.2.10) и (2.2.11) дают более точное приближение, чем (2.2.8) и (2.2.9).

Расчёты по приближённым формулам (2.2.8) – (2.2.11) могут производиться при соблюдении условия

$$|m - np| < 3\sqrt{npq},$$

которое непосредственно вытекает из (2.2.5).

Теорема Муавра-Лапласа описывает поведение биномиального распределения при больших значениях n , что позволяет значительно упростить вычисления. Расчёты по точной формуле

$$P(m_1 \leq \hat{z}_n < m_2) = \sum_{m=m_1}^{m_2} C_n^m p^m q^{n-m} \quad (2.2.12)$$

при больших значениях n очень громоздки.

В выражение (2.2.12) входит число сочетаний из n элементов по m :

$$C_n^m = \frac{n!}{m!(n-m)!},$$

где $n! = 1 \cdot 2 \cdot \dots \cdot n$, $m! = 1 \cdot 2 \cdot \dots \cdot m$.

Пример 2.2. По линии связи независимо друг от друга передаётся 90 сообщений, каждое из которых состоит из пяти двоичных чисел. Вероятность искажения хотя бы одного числа в сообщении равна 0,06. Определить вероятность того, что число принятых без искажения сообщений будет находиться в интервале [82; 87].

▼ Обозначим через \hat{x}_i число неискажённых сообщений при i -й передаче. Эта случайная величина принимает значение ноль с вероятностью 0,06 и единица – с вероятностью 0,94. Следовательно

$$p_i = p = 0,94; \quad q_i = q = 0,06.$$

Обозначим $\hat{z}_n = \sum_{i=1}^{90} \hat{x}_i$ – общее число неискажённых сообщений. Тогда

где

$$M_{\hat{z}_n} = np = 90 \cdot 0,94 = 84,6,$$

$$\sigma_{\hat{z}_n} = \sqrt{npq} = \sqrt{90 \cdot 0,94 \cdot 0,06} = 2,25.$$

По формуле (2.2.9) получаем

$$P(82 \leq \hat{z}_n \leq 87) \approx \Phi_1\left(\frac{87 - 84,6}{2,25}\right) - \Phi_1\left(\frac{82 - 84,6}{2,25}\right) = \\ = \Phi_1(1,067) - \Phi_1(-1,156) = 0,857 - 0,124 = 0,733.$$



2.2.3. Неравенство Чебышева

Для любой случайной величины, имеющей конечное математическое ожидание и дисперсию, при каждом $\varepsilon > 0$ имеет место неравенство

$$P(|\hat{x} - M_{\hat{x}}| \geq \varepsilon) \leq \frac{D_{\hat{x}}}{\varepsilon^2}. \quad (2.2.13)$$

Для противоположного события неравенство Чебышева принимает вид

$$P(|\hat{x} - M_{\hat{x}}| < \varepsilon) \geq 1 - \frac{D_{\hat{x}}}{\varepsilon^2}. \quad (2.2.14)$$

Неравенства (2.2.13) и (2.2.14) можно использовать для получения оценок вероятностей отклонения случайной величины от своего математического ожидания, если закон распределения случайной величины неизвестен.

Пример 2.3. Найти нижнюю границу вероятности того, что случайная величина \hat{x} , имеющая произвольный закон распределения, отклоняется от своего математического ожидания меньше чем на $\pm 3\sigma_{\hat{x}}$.

▼ По формуле (2.2.14) получим

$$P(|\hat{x} - M_{\hat{x}}| < 3\sigma_{\hat{x}}) \geq 1 - \frac{D_{\hat{x}}}{(3\sigma_{\hat{x}})^2} = 1 - \frac{D_{\hat{x}}}{9D_{\hat{x}}} = 1 - \frac{1}{9} = \frac{8}{9}.$$

Известно, что для нормального закона распределения существует так называемое «правило трёх сигм», согласно которому вероятность попадания случайной величины в интервал

$$[M_{\hat{x}} - 3\sigma_{\hat{x}}; M_{\hat{x}} + 3\sigma_{\hat{x}}]$$

близка к единице ($\approx 0,997$). Подобное правило существует и для случайных величин, имеющих распределение, отличное от нормального, но при этом вероятность указанного события будет не меньше $8/9$.



2.2.4. Теоремы Чебышева и Маркова

Частная теорема Чебышева. При неограниченном увеличении независимых испытаний среднее арифметическое полученных при испытаниях значений случайной величины, имеющей конечную дисперсию, сходится по вероятности к её математическому ожиданию:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n \hat{x}_i - M_{\hat{x}}\right| < \varepsilon\right) = 1. \quad (2.2.15)$$

Из (2.2.15) и (2.2.1) следует, что при ограниченном n справедливо приближённое равенство

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n \hat{x}_i - M_{\hat{x}}\right| < \varepsilon\right) \approx 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sigma_{\hat{x}}}\right), \quad (2.2.16)$$

где $\frac{\sigma_{\hat{x}}}{\sqrt{n}} = \sigma[M_{\hat{x}}]$ – среднее квадратическое отклонение математического ожидания случайной величины \hat{x} .

В выражении (2.2.16) учтено, что математическое ожидание случайной величины

$$\left|\frac{1}{n} \sum \hat{x} - M_{\hat{x}}\right|$$

равно нулю.

Применяя неравенство Чебышева (2.2.14) для случайной величины $1/n \sum_{i=1}^n \hat{x}_i$, получим

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n \hat{x}_i - M_{\hat{x}}\right| < \varepsilon\right) \geq 1 - \frac{D_{\hat{x}}}{n\varepsilon^2}. \quad (2.2.17)$$

Формулой (2.2.16) можно пользоваться, когда применима теорема Ляпунова или когда закон распределения каждой случайной величины \hat{x}_i , $i = \overline{1, n}$ нормальный. Если же теорема Ляпунова не применима или законы распределения \hat{x}_i , $i = \overline{1, n}$ неизвестны, то приходится определять нижние границы соответствующих вероятностей из соотношения (2.2.17).

При решении практических задач с применением теоремы Чебышева часто возникают трудности, связанные с невозможностью обеспечить независимость испытаний. Теорема Маркова определяет условия, при которых закон больших чисел справедлив и для зависимых испытаний.

ТЕОРЕМА МАРКОВА. Если случайная величина \hat{x} такова, что

$$\lim_{n \rightarrow \infty} \frac{D\left[\sum_{i=1}^n \hat{x}_i\right]}{n^2} = 0,$$

то для любого числа $\varepsilon > 0$ существует предельное соотношение

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

2.2.5. Теорема Бернулли

Если производится n независимых испытаний и вероятность появления события A в каждом из них равна p , то при любом $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{\hat{z}_n}{n} - p\right| < \varepsilon\right) = 1,$$

где \hat{z}_n – число появления события A в n испытаниях.

При $n \rightarrow \infty$ согласно теореме Ляпунова можно считать, что случайная величина

$$P^* = \frac{\hat{z}_n}{n}$$

будет иметь нормальное распределение, поэтому справедливо приближённое равенство

$$P(|P^* - p| < \varepsilon) \approx 2\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right). \quad (2.2.17)$$

В соотношении (2.2.17) учитывается, что

$$D[|P^* - p|] = D[P^*] = D\left[\frac{\hat{z}_n}{n}\right] = \frac{1}{n^2} D_{\hat{z}_n} = \frac{npq}{n^2} = \frac{pq}{n} \quad \text{или} \quad \sigma[|P^* - p|] = \sqrt{\frac{pq}{n}}.$$

К случайной величине P^* применимо неравенство Чебышева

$$P(|P^* - p| < \varepsilon) \geq 1 - \frac{pq}{n\varepsilon^2}.$$

Теорема Бернулли устанавливает факт устойчивости частоты, когда вероятность появления события от испытания к испытанию не меняется.

2.3. Элементы теории статистических решений

2.3.1. Задачи принятия статистических решений при обработке экспериментальных данных

Конечной целью обработки экспериментальных данных являются некоторые выводы о состоянии или свойствах исследуемого процесса или объекта. Например, это могут быть выводы о законах распределения случайных величин или параметрах законов распределения, о справедливости каких-либо гипотез и, наконец, выводы о наличии и особенностях взаимосвязей различных показателей, характеризующих свойства объекта. В любом случае из множества возможных выводов необходимо выбрать один, оптимальный в каком-либо смысле. Иначе говоря, необходимо принять решение.

Решением называется некоторое заключение, вывод об исследуемом объекте или его свойствах.

Обработка данных всегда осуществляется в условиях неопределённости, обусловленной неполнотой информации об исследуемом объекте, помехами как естественного, так и искусственного характера. В связи с этим принимаемые решения являются статистическими.

Статистическим решением называется некоторое заключение (вывод) об исследуемом объекте или его свойствах, полученное в результате обработки экспериментальных данных.

Основой для принятия решения является **решающее правило** (решающая функция), которое служит для выбора из множества возможных решений одного предпочтительного.

Пусть $\{E\}$ - множество возможных состояний исследуемого объекта; $\{X_{<n>}\}$ – множество возможных результатов наблюдений, а $\{\tilde{E}\}$ – множество возможных решений.

Функция

$$R: \{X_{<n>}\} \rightarrow \{\tilde{E}\},$$

отображающая множество $\{X_{<n>}\}$ результатов наблюдений в множество решений $\{\tilde{E}\}$, называется решающим правилом (решающей функцией).

Существует много решающих правил, но из этих правил выбирается такое, которое обеспечивает принятие решения требуемого качества, т.е. решения в определённом смысле оптимального.

В связи с этим возникает задача определения оптимального решающего правила. Выбор такого правила определяется рядом факторов:

- а) требованиями, которые предъявляются к качеству решения;
- б) свойствами экспериментальных данных;
- в) условиями, в которых получены данные;
- г) дополнительной априорной информацией, которая может быть использована при принятии решения.

Требования к качеству решения определяются потребителем решения (см. рис.1.1) и могут быть сформированы в виде требований минимизации:

- потерь, которые может понести потребитель при неправильном решении;
- риска, связанного с принятием неправильного решения;
- вероятности принятия неправильного решения.

Потерями называются отрицательные последствия, сопровождающие реализацию принятого решения.

Риск – это возможность некоторых потерь со стороны потребителя.

Условия, в которых получены данные, можно разделить на две группы: условия пассивного эксперимента и условия активного эксперимента. В первом случае планирование экспериментальных работ с целью получения данных с необходимыми свойствами отсутствует. Во втором случае эксперимент организуется так, чтобы полученные результаты обладали требуемыми свойствами.

Наиболее важными свойствами совокупности экспериментальных данных, существенно влияющими на качество решающей функции являются объём выборки и её представительность, статистическая устойчивость, однородность, отсутствие аномальных результатов.

Влияние данных свойств на качество решения и вид решающей функции рассмотрено в последующих разделах.

Наиболее важными свойствами априорной информации, оказывающими влияние на формирование решающей функции, являются объём данной информации и её достоверность. Так, в идеальном случае априорная информация I_1 (см. рис.1.1) позволяет получить ряд распределения вероятностей состояний исследуемого объекта, наиболее полной формой информации I_4 является закон распределения вектора результатов наблюдений.

Описание свойств решений может осуществляться с различных позиций. Чаще всего для этого применяется функция потерь.

Пусть множество решений $\{\tilde{E}\}$ является дискретным и состоит из l альтернативных решений \tilde{E}_j , $j = \overline{1, l}$. Такими же свойствами пусть обладает и множество $\{E\}$ состояний объекта, мощность этого множества обозначим через m .

Функцией потерь называется функция $\pi(E_i; \tilde{E}_j)$, $i = \overline{1, m}$, $j = \overline{1, l}$, характеризующая последствия принятия решения \tilde{E}_j при условии, что объект находится в состоянии E_i :

$$\pi(E_i; \tilde{E}_j) = \pi_{ij}. \quad (2.3.1)$$

Функция (2.3.1) записывается в виде матрицы потерь

$$\Pi_{[m;l]} = \begin{pmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1l} \\ \cdots & \cdots & \cdots & \cdots \\ \pi_{i1} & \pi_{i2} & \cdots & \pi_{il} \\ \cdots & \cdots & \cdots & \cdots \\ \pi_{m1} & \pi_{m2} & \cdots & \pi_{ml} \end{pmatrix}.$$

Следует заметить, что для определённого состояния E_i исследуемого объекта величина потерь в общем случае будет величиной случайной.

В дальнейшем будем полагать, что информация I_6 (рис.1.1) представляет собой сведения о функции потерь (2.3.1).

Учитывая изложенное выше, под объёмом априорной информации будем понимать число видов информации, используемой при формировании решающего правила, а под качеством априорной информации – степень её соответствия объективным свойствам исследуемого объекта.

Оптимальным решающим правилом называется правило, обеспечивающее выполнение требований, предъявляемых к качеству решения в конкретных условиях применения данного правила.

Под условиями принятия решения понимается совокупность перечисленных выше факторов, определяющих выбор правила решения. Обозначим вектор условий, в которых применяется решающее правило $R_I(X)$, символом $K_{<4>}$. Компонентами данного вектора являются объём K_{X1} и качество K_{X2} результатов наблюдений, объём K_{I1} и качество K_{I2} априорной информации:

$$K_{<4>} = (K_{X1}, K_{X2}, K_{I1}, K_{I2})^T.$$

С учётом этого решающее правило можно представить в виде функции двух переменных – результатов наблюдений и априорной информации, т.е.

$$R_I(X) = R(X_{<n>}; I),$$

а качество (оптимальность) данного правила применительно к условиям $K_{<4>}$ охарактеризовать показателем

$$L = f(R; K_{<4>}),$$

представляющим собой эффект, который достигается в результате использования решающего правила R в условиях $K_{<4>}$. Конкретное выражение показателя L при решении различных задач обработки может быть различным. Поэтому оптимальным решающим правилом R_0 для некоторой совокупности условий обработки $K_{<4>j}$, $j = \overline{1, I}$, целесообразно считать правило, обеспечивающее экстремум показателя L в условиях $K_{<4>j}$:

$$R_0(X_{<n>}; I) = \arg \operatorname{extr}_{R_i \in \{R\}} \{L(R_i; K_{<4>j})\},$$

где $\{R\}$ – множество решающих правил.

На практике подбор оптимального решающего правила для некоторых условий обработки выполняется на основе принципов принятия решений. В настоящее время сформулирован ряд таких принципов, основными из которых являются принцип максимального правдоподобия и принцип минимальной вероятности ошибки.

Рассмотрим их более подробно.

2.3.2. Принцип максимального правдоподобия

Данный принцип используется в тех случаях, когда известен только условный закон распределения результатов наблюдений относительно состояния E исследуемого объекта.

Пусть случайная величина \hat{x} имеет плотность распределения $\varphi_{\hat{x}/E}(x; E)$, а результаты наблюдения над величиной \hat{x} представляют собой простую (повторную) случайную выборку

$$(\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n)^T = \hat{X}_{<n>}.$$

Вероятность появления конкретной реализации $X_{<n>}$ пропорциональна элементу вероятности распределения случайного вектора $\hat{X}_{<n>}$:

$$P(\hat{X}_{<n>} = X_{<n>}) = P\left(\bigcap_{i=1}^n (x_i \leq \hat{x}_i < x_i + dx_i)\right) = \varphi_{\hat{X}_{<n>}/E}(X_{<n>}; E) dX_{<n>}. \quad (2.3.2)$$

Плотность распределения для простой выборки представляется выражением

$$\varphi_{\hat{X}_{<n>}/E}(X_{<n>}; E) = \prod_{i=1}^n \varphi_{\hat{x}_i/E}(x_i; E) = \prod_{i=1}^n \varphi_{\hat{x}/E}(x_i; E).$$

Функция оцениваемого состояния E вида

$$L(E; X_{<n>}) = \varphi_{\hat{X}_{<n>}/E}(X_{<n>}; E) \quad (2.3.3)$$

называется **функцией правдоподобия**.

Принцип максимального правдоподобия состоит в утверждении, что при данной реализации $X_{<n>}$ вектора $\hat{X}_{<n>}$ наиболее правдоподобным, т.е. наиболее близким к действительному состоянию, является то значение \tilde{E} оценки состояния исследуемого объекта, при котором вероятность (2.3.2), а следовательно, и функция правдоподобия (2.3.3) максимальны.

Исходя из данного утверждения, для определения наиболее правдоподобного значения состояния E необходимо найти такое значение \tilde{E} , которое обеспечивает экстремум функции правдоподобия. Если оценивается несколько параметров E_1, E_2, \dots, E_k какого-либо состояния, то решение данной задачи сводится к нахождению корней системы уравнений

$$\frac{\partial L(E_1; E_2; \dots; E_k; X_{<n>})}{\partial E_j} = 0, \quad j = \overline{1, k}. \quad (2.3.4)$$

Выражения (2.3.4) представляют собой необходимое условие экстремума функции многих переменных, которое заключается в равенстве нулю всех частных производных данной функции.

Решение уравнений (2.3.4) даёт оценки исследуемых параметров, которые называются **оценками максимального правдоподобия**.

Таким образом, принятие решения в соответствии с принципом максимального правдоподобия заключается в следующем. На основе выборки $X_{<n>}$ определяются значения функции правдоподобия для всех возможных состояний E_i , $i = \overline{1, n}$. Среди данных состояний выбирается то, которое обеспечивает максимум функции правдоподобия.

2.3.3. Принцип минимальной вероятности ошибки

Рассматриваемый принцип, так же как и предыдущий, применим в тех случаях, когда известен только условный закон распределения результатов наблюдений. Сущность принципа состоит в том, что минимизируется вероятность принятия неправильного решения. Введём следующие определения.

Ошибкой первого рода называется ошибка, представляющая собой принятие решения о том, что исследуемый объект не находится в предполагаемом состоянии, в то время как в действительности он пребывает именно в этом состоянии.

Ошибкой второго рода называется ошибка, представляющая собой принятие решения о том, что исследуемый объект находится в предполагаемом состоянии, в то время как в действительности он пребывает в другом состоянии.

В общем случае правило решения должно быть таким, чтобы обеспечивалась минимально возможная вероятность принятия ошибочных решений. Если бы была известна функция потерь или функция риска, соответствующая каждому из исходов, то задачу поиска решающего правила, минимизирующего вероятность принятия ошибочного решения, можно было бы переформулировать как задачу нахождения решающего правила, минимизирующего вероятность ошибки либо первого, либо второго рода, в зависимости от того, какая из них связана с большими потерями или большим риском.

Так как функция потерь (риска) неизвестна, то задача поиска решающего правила формулируется как задача минимизации суммы вероятностей ошибок первого и второго рода. Метод решения данной задачи рассмотрим для случая, когда исследуемый объект может находиться в одном из двух состояний E_1 или E_2 . Пусть наблюдаемая переменная \hat{x} является скалярной, а кривые условных законов распределения $\varphi_{\hat{x}/E_i}(x; E_i)$ при условии, что объект может находиться в состоянии E_i , $i = \overline{1, 2}$, имеют вид, изображённый на рис.2.1.

Как видно из рисунка, состояниям E_1 и E_2 соответствуют некоторые подмножества значений переменной \hat{x} , попадание в которые результата наблюдения с наибольшей вероятностью соответствует тому или

иному состоянию объекта. Поэтому, фиксируя попадание наблюдаемого результата в одно или другое подмножество, можно судить с некоторой вероятностью о состоянии, которое принял объект. Пусть такими подмножествами являются $(-\infty; x_\alpha)$ и $[x_\alpha; +\infty)$. Величина x_α является границей данных подмножеств. Тогда при $x \in [x_\alpha; +\infty)$ принимается решение о нахождении объекта в состоянии E_1 , а если $x \in (-\infty; x_\alpha)$ – о нахождении объекта в состоянии E_2 .

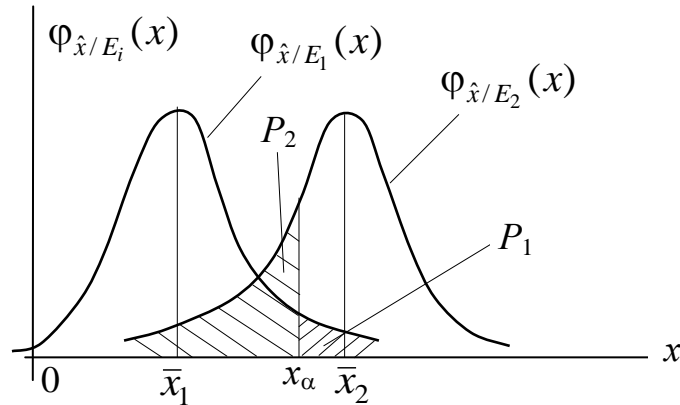


Рис.2.1. Условные законы распределения случайной величины

Поскольку кривые частично перекрываются, существует вероятность ошибки. Так, результат наблюдения с вероятностью

$$P_1 = \int_{x_\alpha}^{\infty} \varphi_{\hat{x}/E_1}(x; E_1) dx \quad (2.3.5)$$

может попасть в область $[x_\alpha; +\infty)$, если объект находится в состоянии E_1 , но при этом будет принято решение, что он пребывает в состоянии E_2 . Это означает ошибку первого рода. Наоборот, с вероятностью

$$P_2 = \int_{-\infty}^{x_\alpha} \varphi_{\hat{x}/E_2}(x; E_2) dx \quad (2.3.6)$$

результат наблюдения может попасть в область $(-\infty; x_\alpha)$, если объект находится в состоянии E_2 , но принимается решение о пребывании объекта в состоянии E_1 . Будет допущена ошибка второго рода.

Если последствия ошибочных решений оценить невозможно, то очевидно, что решающее правило должно обеспечивать минимум суммы вероятностей (2.3.5) и (2.3.6). Так как данное правило состоит в определении границы x_α , она должна выбираться таким образом, чтобы минимизировать величину

$$u = \int_{-\infty}^{x_\alpha} \varphi_{\hat{x}/E_2}(x; E_2) dx + \int_{x_\alpha}^{\infty} \varphi_{\hat{x}/E_1}(x; E_1) dx.$$

Общих правил выбора оптимального значения x_α не существует. На практике чаще всего минимизируется вероятность ошибки первого рода до определённой заранее назначенной величины. На основе этого и выбирается значение критической границы.

В том случае, когда объект может находиться более чем в двух состояниях, применение рассмотренного принципа существенно усложняется. Ввиду этого используется не сам наблюдаемый параметр, а построенная на его основе специальная функция, так называемый показатель согласованности (см. раздел 7).

Следует отметить, что на основе рассмотренных выше принципов может быть сформировано большое число показателей и критериев принятия решений, специфика построения которых определяется особенностями конкретной задачи.

2.4. Элементы теории оценивания

Первичной задачей обработки экспериментальных данных является задача оценивания. При её решении наибольшее распространение получил принцип максимального правдоподобия и вытекающие из него критерии и алгоритмы оценивания.

Пусть схема наблюдения имеет вид

$$\hat{Y} = F(A) + \hat{E}, \quad (2.4.1)$$

а вектор ошибок измерений \hat{E} имеет нормальное распределение

$$\varphi_{\hat{E}}(E) = C \exp\left(-\frac{1}{2} E^T K_{\hat{E}}^{-1} E\right),$$

где $C = (2\pi)^{n/2} |K_{\hat{E}}|^{-1}$ – нормирующий множитель; $K_{\hat{E}}$ – корреляционная матрица вектора ошибок измерений.

Учитывая, что

$$\hat{E} = \hat{Y} - F(A),$$

плотность распределения $\varphi_{\hat{E}}(E)$ можно выразить через \hat{Y} и $F(A)$:

$$\varphi_{\hat{Y}}(Y / A) = C \exp\left(-\frac{1}{2} (Y - F(A))^T K_{\hat{E}}^{-1} (Y - F(A))\right).$$

Тогда принцип максимального правдоподобия приводит к следующей функции потерь:

$$V(\hat{Y}; A) = (\hat{Y} - F(A))^T K_{\hat{E}}^{-1} (\hat{Y} - F(A)). \quad (2.4.2)$$

Таким образом, при нормальном законе распределения выборки функция потерь является квадратичной. В частном случае, когда все эле-

менты выборки имеют одинаковое распределение с дисперсией σ^2 и независимы, функция потерь (2.4.2) принимает вид

$$V(\hat{Y}; A) = (\hat{Y} - F(A))^T (\hat{Y} - F(A)). \quad (2.4.3)$$

Метод оценивания, основанный на минимизации квадратичной функции потерь вида (2.4.2) или (2.4.3), называется **методом наименьших квадратов** (см. раздел 8). Этот метод является оптимальным и для ряда других распределений ошибок наблюдений.

Если рассматривать схему наблюдения (2.4.1) в предположении, что вектор ошибок измерений имеет распределение Лапласа, то получим функцию потерь в виде суммы модулей ошибок. Метод оценивания вектора параметров, основанный на минимизации функции потерь как суммы модулей ошибок измерений, называется **методом наименьших модулей** [8].

В настоящей брошюре он не рассматривается. Следует только отметить, что данный метод является оптимальным и в ряде других задач оценивания.