3. МЕТОДЫ СТАТИСТИЧЕСКОГО ОЦЕНИВАНИЯ

Одной из важнейших задач обработки данных является задача оценивания (экспериментального определения) вероятностных характеристик случайных объектов.

3.1. Постановка задачи оценивания законов и параметров распределения случайных величин

Пусть $\hat{x}$ – случайная величина, характеризующая свойство исследуемого объекта. Требуется на основе случайной выборки $\hat{x}_1,\ \hat{x}_2,\ldots,\hat{x}_n$ выработать объективное суждение о вероятностных свойствах случайной величины $\hat{x}$.

Как было указано в § 1.3, любая функция случайной выборки

$$\hat{s} = s(\hat{x}_1,\ \hat{x}_2,...,\hat{x}_n) \qquad (3.1.1)$$

называется статистикой. Если статистика (3.1.1) используется в качестве приближения неизвестной вероятностной характеристики (закона или параметра распределения) случайной величины, то её значение

$$\tilde{s} = s(x_1,\ x_2,...,x_n), \qquad (3.1.2)$$

полученное в результате обработки экспериментальных данных по формуле (3.1.2), называется ***оценкой*** этой характеристики.

Известно, что исчерпывающей характеристикой вероятностного поведения случайной величины $\hat{x}$ является закон её распределения $F_{\hat{x}}(x)$ или $\varphi_{\hat{x}}(x)$. Поэтому основной целью в рассматриваемой здесь задаче является построение закона распределения случайной величины $\hat{x}$ по экспериментальным данным, т.е. его представление как функции выборки

$$\tilde{F}_{\hat{x}}(x) = F_{\hat{x}}(x;\ x_1,\ x_2,...,x_n) = s(x;\ X_{<n>}),$$

которая может служить в качестве оценки функции $F_{\hat{x}}(x)$, обладающей требуемой точностью и надёжностью (достоверностью).

В общем случае функция $F_{\hat{x}}(x)$ зависит как от своего аргумента, так и от параметров распределения:

$$F_{\hat{x}}(x) = F_{\hat{x}}(x;\ a_1,\ a_2,...,a_m) = F_{\hat{x}}(x;\ A_{<m>}).$$

Параметрическая обработка данных опирается на предположение о том, что класс распределений, которому принадлежит функция $F_{\hat{x}}(x)$, априорно известен. Конкретные значения параметров $A_{<m>}$ этого распределения, выделяющие его в рассматриваемом классе, неизвестны. Тогда оценивание функции $F_{\hat{x}}(x)$ сводится к оцениванию её параметров $A_{<m>}$, т.е. к отысканию такой статистики

$$S_{<m>}(X_{<n>}) = \tilde{A}_{<m>} \approx A_{<m>},$$

которая обеспечивала бы приближённое равенство

$$F_{\hat{x}}(x) = F_{\hat{x}}(x;\ A_{<m>}) \approx F_{\hat{x}}(x;\ \tilde{A}_{<m>}).$$

Поскольку вся информация об исследуемом объекте содержится в выборке объёма n, то для однозначного решения задачи статистического оценивания m параметров требуется выполнение условия $n > m$.

В качестве критериев оценивания истинных значений характеристик используются соотношения следующего вида:

$$\begin{cases} F_{\hat{x}}(x) \approx s(x;\ X_{<n>}); \\ A_{<m>} \approx S_{<m>}(X_{<n>}) \end{cases}. \qquad (3.1.3)$$

или

$$\begin{cases} s'(x;\ X_{<n>}) \le F_{\hat{x}}(x) \le s''(x;\ X_{<n>}); \\ S'_{<m>}(X_{<n>}) \le A_{<m>} \le S''_{<m>}(X_{<n>}), \end{cases} \qquad (3.1.4)$$

где $s'(x;\ X_{<n>})$, $s''(x;\ X_{<n>})$ – нижняя и верхняя границы интервалов; $S'_{<m>}(X_{<n>})$, $S''_{<m>}(X_{<n>})$ – границы m-мерных областей.

Оценивание вероятностных характеристик в соответствии с критерием (3.1.3) называется ***точечным***, а в соответствии с критерием (3.1.4) – интервальным. Строго говоря оценки всегда являются точечными. Что же касается интервальных оценок, то их назначение - характеризовать качество точечных оценок.

Предположим, что распределение $F_{\hat{x}}(x)$ однопараметрическое, т.е. $A_{<m>} = A_{<1>} = a$. Принятое допущение позволяет существенно повысить наглядность рассуждений, которые без затруднений распространяются на случай многопараметрического распределения. Кроме того, будем считать, что класс распределений, которому принадлежит функция $F_{\hat{x}}(x)$, известно, но неизвестно значение параметра a. В этом случае задача оценивания функции распределения $F_{\hat{x}}(x;\ a)$ сводится к оцениванию параметра, т.е. к определению соотношения вида

$$F_{\hat{x}}(x) = F_{\hat{x}}(x;\ a) \approx F_{\hat{x}}(x;\ \tilde{a}) = \tilde{F}_{\hat{x}}(x),$$

где $\tilde{a} = S_{<1>}(X_{<n>}) = s(X_{<n>})$ – оценка параметра a.

Поскольку результаты $\hat{X}_{<n>}$ наблюдений над случайной величиной $\hat{x}$ априори являются случайными, то случайной оказывается и оценка $\tilde{a}$:

$$\tilde{a} = s(\hat{X}_{<n>}) = \hat{s}.$$

В общем случае $\tilde{a} \ne a$, следовательно, и после получения оценки $\tilde{a}$ параметра a его неопределённость для исследователя полностью не снимается. В то же время исследователь даёт вероятностное суждение об истинном значении a согласно результату эксперимента так, чтобы соответствовать ему наилучшим (в некотором смысле) образом. Оценка будет объективной характеристикой параметра, если она удовлетворяет требованиям несмещённости, состоятельности и эффективности.

Оценка $\tilde{a}$ параметра a называется ***несмещённой***, если её математическое ожидание равно оцениваемому параметру:

$$M_{\tilde{a}} = a. \qquad (3.1.5)$$

Если $M_{\tilde{a}} \neq a$, то оценка называется ***смещённой***.

Оценка $\tilde{a}$ параметра a называется ***состоятельной***, если она сходится по вероятности к оцениваемому параметру:

$$\lim_{n\to\infty} P(|\tilde{a} - a| \le \varepsilon) = 1, \quad \forall \varepsilon > 0, \qquad (3.1.6)$$

где n – объём выборки.

Очевидно, что состоятельной может быть только несмещённая оценка. Поскольку согласно известному неравенству Чебышева

$$P(|\tilde{a} - a| \le \varepsilon) = 1 - \frac{D_{\tilde{a}}}{\varepsilon^2},$$

то из выражений (3.1.5) и (3.1.6) следует, что

$$\lim_{n\to\infty} D_{\tilde{a}} = 0,$$

т.е. с ростом объёма выборки дисперсия состоятельной оценки стремится к нулю, и наоборот, если с ростом n дисперсия стремится к нулю, то оценка $\tilde{a}$ состоятельная.

Несмещённая оценка[1] параметра a называется ***эффективной***, если её дисперсия минимальна:

$$D_{\tilde{a}} = \min_k \{D_{\tilde{a}_k} \mid k = 1,2,...\}, \qquad (3.1.7)$$

где $\tilde{a}_k = s_k(\hat{X}_{<n>})$ – оценка параметра a с помощью статистики k-го вида.

Если равенство (3.1.7) выполняется только в пределе при $n \to \infty$, то соответствующая оценка называется ***асимптотически эффективной***. Из последнего определения следует, что состоятельная оценка асимптотически эффективна.

Оценки, удовлетворяющие всем трём перечисленным требованиям, называются ***подходящими значениями*** оцениваемых параметров. На практике достичь совместного выполнения всех трёх условий (3.1.5), (3.1.6) и (3.1.7) удаётся не всегда, так как формулы для вычисления эффективной и несмещённой оценки могут оказаться слишком сложными. Поэтому для упрощения расчётов нередко используются незначительно смещённые и не вполне эффективные оценки. Однако, выбор той или иной оценки должен опираться на её критическое рассмотрение со всех указанных выше точек зрения.

[1] Для смещённой оценки понятие эффективности не определено.

Следует отметить, что определение эффективной оценки имеет аналитическое выражение (3.1.7) лишь в случае оценивания единственного параметра распределения $\varphi_{\hat{x}}(x;a)$. Если число m оцениваемых параметров $(a_1, a_2,\ldots, a_m)^{\text{т}} = A_{<m>}$ больше единицы, то в качестве характеристики рассеяния их оценок $(\tilde{a}_1, \tilde{a}_2,\ldots,\tilde{a}_m)^{\text{т}} = \tilde{A}_{<m>}$ должна использоваться обобщённая дисперсия

$$D_{\tilde{A}_{<m>}} \overset{d}{=} |\, K_{\tilde{A}_{<m>}} \,|. \qquad (3.1.8)$$

В правой части соотношения (3.1.8) определитель корреляционной матрицы вектора $\tilde{A}_{<m>}$ оценок параметров $A_{<m>}$. Более полное раскрытие понятия обобщённой дисперсии можно найти, например, в монографии [6].

Ещё раз подчеркнём, что следует различать две фазы оценивания вероятностных характеристик случайных объектов: априорную (доопытную) и апостериорную (послеопытную). На первой фазе оценки рассматриваются как функции случайной выборки и, следовательно, сами случайны. На второй фазе они не случайны, так как представляют собой функции выборки (реализации случайной выборки), элементы которой не случайны. Очевидно, что все требования к оценкам как законов, так и параметров распределения случайной величины предъявляются на первой фазе их оценивания, т.е. априори.

3.2. Качество статистического оценивания

Как отмечалось в § 3.1, при обработке данных принято различать точечное и интервальное оценивание вероятностных характеристик случайных объектов. Однако, строго говоря, собственно оценки $\tilde{F}_{\hat{x}}(x)$, $\tilde{A}_{<m>}$ и т.п., представляющие практический интерес, могут быть только точечными и определяются приближёнными равенствами типа (3.1.3). Что касается оценок (3.1.4), то они оценивают не характеристики $F_{\hat{x}}(x)$, $A_{<m>}$, а интервалы, в которых эти характеристики могут находиться, а могут и не находиться. Последнее утверждение обосновывается тем, что статистики $\hat{s}'$, $\hat{s}''$, $\hat{S}'_{<m>}$, $\hat{S}''_{<m>}$ априори случайны (как функции случайных аргументов – элементов случайной выборки $\hat{X}_{<n>}$). Поэтому ясно, что интервальное оценивание сугубо вероятностное и служит для характеристики качества точечного оценивания. Компонентами, которые характеризуют это качество, являются точность и надёжность (достоверность).

Раскроем существо задачи исследования точности и надёжности статистического оценивания. Пусть по результатам n наблюдений случайной величины $\hat{x}$ получена точечная оценка $\tilde{a}$. Возникает вопрос: насколько эта оценка точна и надёжна.

Точность статистического оценивания характеризуется абсолютной погрешностью (ошибкой)

$$\Delta\tilde{a} = \tilde{a} - a.$$

Истинную погрешность Δa определить невозможно даже при известной оценке $\tilde{a}$. Это объясняется тем, что исследователь не знает истинное значение параметра a. Поэтому вводится понятие вероятной погрешности статистической оценки параметра a.

Максимальной вероятной погрешностью статистической оценки называется её максимально возможное отклонение $\varepsilon_\beta > 0$ от оцениваемой характеристики случайного объекта, гарантируемое с вероятностью не менее β. Данная величина (ε_β) имеет также эквивалентное название - ***максимальная с вероятностью β погрешность оценки*** какой – либо характеристики.

Если оценка несмещённая, то её математическое ожидание равно оцениваемой характеристике, т.е. справедливо равенство (3.1.5). Тогда при симметричном распределении оценки $\tilde{a}$ относительно математического ожидания имеет место следующее соотношение:

$$\beta = \beta(a) = P(a' \le \tilde{a} \le a'') = P(a - \varepsilon_\beta \le \tilde{a} \le a + \varepsilon_\beta) = P(|\tilde{a} - a| \le \varepsilon_\beta), \quad (3.2.1)$$

где $\varepsilon_\beta = \varepsilon_\beta(a)$ – максимальная с вероятностью β погрешность оценки $\tilde{a}$ параметра a.

Соотношение (3.2.1) введено в предположении, что параметр a известен. Но тогда задача его оценивания теряет смысл, её просто не существует. На практике дело обстоит иначе. На основе экспериментальных данных определяется оценка параметра, истинное значение которого остаётся неизвестным. Затем вводится соотношение

$$\beta = P(|\tilde{a} - a| \le \varepsilon_\beta) = P(\tilde{a} - \varepsilon_\beta \le a \le \tilde{a} + \varepsilon_\beta) = P(\tilde{a}' \le a \le \tilde{a}''). \quad (3.2.2)$$

Следует подчеркнуть, что похожие на первый взгляд выражения (3.2.1) и (3.2.2) имеют различный вероятностный смысл. Так, (3.2.1) определяет вероятность того, что случайная величина $\tilde{a}$ попадает в неслучайный интервал $[a - \varepsilon_\beta;\ a + \varepsilon_\beta]$, а (3.2.2) – вероятность того, что неслучайное (хотя и неизвестное) значение a оцениваемого параметра окажется в пределах случайного интервала $[\tilde{a} - \varepsilon_\beta;\ \tilde{a} + \varepsilon_\beta]$. Данный интервал является случайным как по величине, так и по расположению на вещественной оси, т.е. он накрывает точку a.

Интервал $I_\beta = [\tilde{a} - \varepsilon_\beta;\ \tilde{a} + \varepsilon_\beta]$ называется ***доверительным интервалом***, соответствующим ***доверительной вероятности*** β, или 100β-процентным доверительным интервалом. Его границы называются доверительными границами для параметра a.

Очевидно, чем у́же доверительный интервал, тем меньше максимальная с вероятностью β погрешность ε_β оценки параметра, тем она точнее. С другой стороны, чем больше доверительная вероятность, тем более надёжна (достоверна) оценка, тем с бо́льшим доверием можно к ней относиться.

Абсолютная достоверность оценивания характеризуется доверительной вероятностью $\beta = 1$. В условиях воздействия случайных факторов такая достоверность не достижима, поэтому реальная доверительная вероятность определяется на основе принципа практической уверенности. Согласно этому принципу события, имеющие вероятности, близкие к единице, считаются практически достоверными, а имеющие вероятности, близкие к нулю, – практически невозможными. Иначе говоря, если вероятность случайного события близка к единице (к нулю), то практически можно быть уверенным, что при однократном проведении опыта это событие произойдёт (не произойдёт).

Вероятность практически достоверного события определяется сущностью решаемой задачи. При анализе качества статистического оценивания обычно принимают $\beta \in [0{,}8;\ 0{,}99]$.

Как было показано в § 3.1 [см. формулу (3.1.4)], доверительные границы представляют собой статистики, т.е. некоторые функции элементов выборки $(x_1,\ x_2,\dots,x_n)^{\mathrm{т}} = X_{<n>}$. В случае одномерного параметра a соотношение (3.1.4) принимает вид

$$s'(X_{<n>}) \le a \le s''(X_{<n>}),$$

где $s'(X_{<n>}) = a'$; $s''(X_{<n>}) = a''$.

Поскольку выборка $\hat{X}_{<n>}$ априори случайна, то и статистики $\hat{s}', \hat{s}''$, а следовательно, и доверительные границы $\hat{a}'$, $\hat{a}''$ априори случайны:

$$\hat{a}' = s'(\hat{X}_{<n>}); \qquad \hat{a}'' = s''(\hat{X}_{<n>}).$$

При анализе качества статистических оценок вся информация об исследуемой переменной содержится в случайной выборке $\hat{X}_{<n>}$. Поэтому не только оценка $\tilde{a}$ параметра, но и её максимальная вероятная погрешность ε_β определяется через выборку:

$$\begin{cases} \hat{a}' = s'(\hat{X}_{<n>}) = \tilde{a}(\hat{X}_{<n>}) - \varepsilon'_\beta(\hat{X}_{<n>}); \\ \hat{a}'' = s''(\hat{X}_{<n>}) = \tilde{a}(\hat{X}_{<n>}) + \varepsilon''_\beta(\hat{X}_{<n>}). \end{cases} \tag{3.2.4}$$

Соотношения (3.2.4) носят общий характер и в явном виде никогда не формируются. На практике погрешность ε_β оценки $\tilde{a}$ выражается через саму оценку:

$$\varepsilon_\beta = \varepsilon_\beta(\tilde{a}).$$

Таким образом, доверительные границы a', a'' для параметра a определяются его оценкой $\tilde{a}$:

$$\begin{cases} \hat{a}' = \tilde{a} - \varepsilon'_\beta(\tilde{a}); \\ \hat{a}'' = \tilde{a} + \varepsilon''_\beta(\tilde{a}), \end{cases} \tag{3.2.5}$$

где функции $\varepsilon'_\beta(\tilde{a})$ и $\varepsilon''_\beta(\tilde{a})$ в общем случае различные. Соотношения (3.2.5) иллюстрируются рис.3.1.

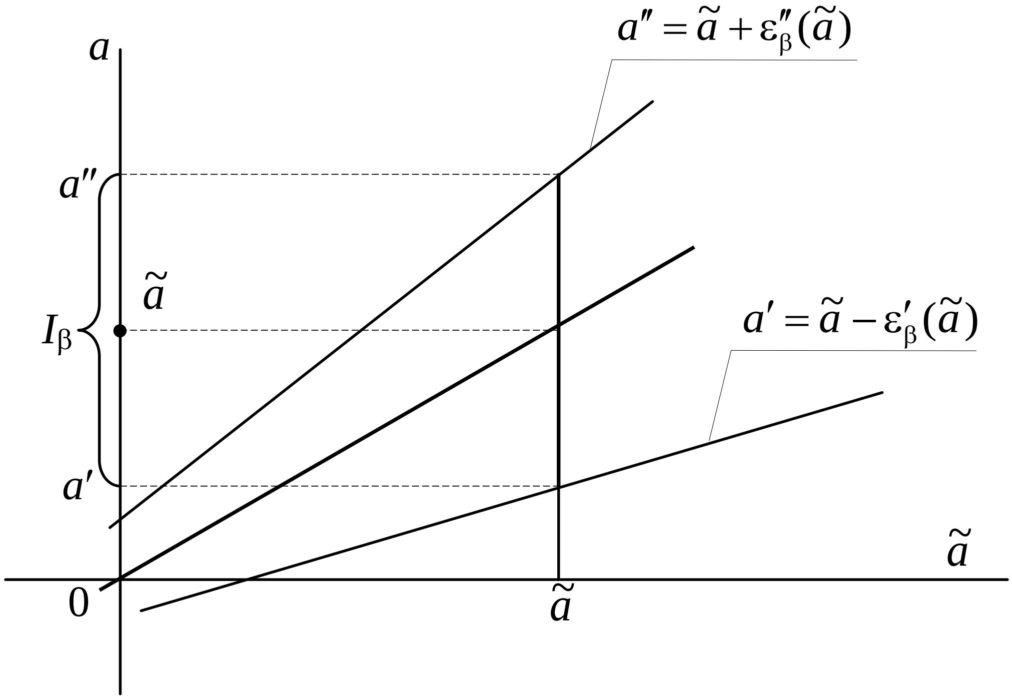

Рис.3.1. Зависимость доверительных границ параметра от его оценки

Выражения (3.2.1) и (3.2.2) были получены в предположении, что оценка $\tilde{a}$ имеет симметричное относительно параметра a распределение. При этом доверительный интервал $I_\beta = [a';\, a'']$ оказывается также симметричным. В этом случае должны выполняться условия $\varepsilon'_\beta = \varepsilon''_\beta = \varepsilon_\beta$, которые означают равенство положительной и отрицательной максимальных вероятных погрешностей. В общем случае это требование не обязательно и при определении ε'_β и ε''_β в основу могут быть положены различные соображения. Чаще всего выдвигается требование

$$P(a < \hat{a}') = P(a > \hat{a}'') = \frac{1-\beta}{2}, \tag{3.2.6}$$

означающее, что истинное значение a оцениваемого параметра может оказаться левее или правее его доверительного интервала с одинаковой вероятностью.

При условии (3.2.6) и симметричном распределении оценки $\tilde{a}$ доверительный интервал для параметра a будет симметричным:

$$I_\beta(a) = [a';\ a''] = [\tilde{a} - \varepsilon_\beta;\ \tilde{a} + \varepsilon_\beta].$$

Следовательно, имеют место равенства

$$\beta = P(a \in I_\beta(a)) = P(\tilde{a} - \varepsilon_\beta \le a \le \tilde{a} + \varepsilon_\beta) = P(|\tilde{a} - a| \le \varepsilon_\beta) = \beta_\varepsilon(a). \qquad (3.2.7)$$

Если распределение оценки $\tilde{a}$ не симметрично относительно оцениваемого параметра a, то при условии (3.2.6) доверительный интервал не симметричен, и соотношение (3.2.7) принимает вид

$$\beta = P(a \in I_\beta(a)) = P(\tilde{a} - \varepsilon'_\beta \le a \le \tilde{a} + \varepsilon''_\beta),$$

где ε'_β и ε''_β – соответственно абсолютные значения максимальных с вероятностью β отрицательной и положительной погрешностей оценки $\tilde{a}$ параметра a.

Итак, доверительный интервал (его составляющие ε'_β и ε''_β) характеризует точность, доверительная вероятность β – надёжность (достоверность) оценки $\tilde{a}$, а вместе они определяют качество оценивания параметра a.

В § 3.1 отмечалось, что с ростом объёма n выборки оценка $\tilde{a}$ сходится по вероятности к оцениваемому параметру a (закон больших чисел) и при этом её дисперсия стремится к нулю. Это значит, что с увеличением n растёт как точность, так и надёжность оценивания. В результате оказываются связанными между собой три характеристики качества статического оценивания:

– доверительный интервал $I_{\beta,n}(a)$;

– доверительная вероятность $\beta_{\varepsilon,n}(a)$;

– объём $n_{\beta,\varepsilon}(a)$ выборки, потребный для оценивания параметра a с заданной точностью и надёжностью.

Указанные характеристики связаны соотношениями, позволяющими управлять качеством статистического оценивания[1]

$$\left.\begin{array}{l}
n\uparrow \Rightarrow \begin{cases} \varepsilon = \text{const} \Rightarrow \beta\uparrow \\ \beta = \text{const} \Rightarrow \varepsilon\downarrow \end{cases} \\
\beta\uparrow \Rightarrow \begin{cases} n = \text{const} \Rightarrow \varepsilon\uparrow \\ \varepsilon = \text{const} \Rightarrow n\uparrow \end{cases} \\
\varepsilon\downarrow \Rightarrow \begin{cases} n = \text{const} \Rightarrow \beta\downarrow \\ \beta = \text{const} \Rightarrow n\uparrow \end{cases}
\end{array}\right\} \qquad (3.2.8)$$

---

[1] Символы ↑, ↓ означают соответственно возрастание и убывание.

Таким образом, при исследовании качества статистического оценивания решается одна из трёх основных задач:

– определение доверительного интервала $I_{\beta,n}$ (или половины его длины $\varepsilon_{\beta,n}$) для параметра a при заданной доверительной вероятности β и фиксированном объёме n выборки;

– определение доверительной вероятности $\beta_{I,n}$ (или $\beta_{\varepsilon,n}$) при заданном доверительном интервале I (или максимальной вероятной погрешности ε) и фиксированном объёме n выборки;

– определение объёма $n_{\beta,I}$ (или $n_{\beta,\varepsilon}$) выборки, потребного для оценивания параметра a с требуемой надёжностью β и точностью I (или ε).

3.3. Оценивание вероятности случайного события

В результате реализации определённого комплекса условий может произойти некоторое случайное событие A, вероятность $P(A) = p$ появления которого неизвестна. Требуется по результатам наблюдений данного события в некотором эксперименте оценить вероятность p.

Для решения поставленной задачи проводится серия n независимых и однородных испытаний – схема Бернулли, т.е. осуществляется n независимых реализаций одного и того же комплекса условий. Подсчитывается число $m(A) = m$ испытаний, в которых событие A появилось. Отношение

$$\frac{m}{n} = P^*(A) = p^*(n) \qquad (3.3.1)$$

называется *частотой* события A в серии n испытаний или его *статистической вероятностью*. Проанализируем свойства частоты p^* как оценки вероятности p.

1. Поскольку число $\hat{m}$ появлений события A в n независимых и однородных испытаниях подчинено биномиальному закону распределения, то

$$M_{p^*} = \frac{1}{n} M_{\hat{m}} = \frac{np}{n} = p\,. \qquad (3.3.2)$$

Из выражения (3.3.2) следует, что частота (3.3.1) является несмещённой оценкой вероятности p.

2. Согласно теореме Бернулли

$$\lim_{n\to\infty} P(|\, p^* - p \,| < \varepsilon) = 1, \quad \forall \varepsilon > 0\,,$$

т.е. частота p^* сходится по вероятности к вероятности p. Следовательно, рассматриваемая частота – это состоятельная оценка вероятности p.

3. Дисперсия частоты

$$D_{p^*} = \frac{1}{n^2} D_{\hat{m}} = \frac{npq}{n^2} = \frac{pq}{n}, \qquad (3.3.3)$$

где $q = 1 - p$. Из соотношения (3.3.3) вытекает, что при $n \to \infty$ дисперсия $D_{p^*} \to 0$. Это означает асимптотическую эффективность указанной оценки. Можно показать, что при любом n дисперсия частоты - минимально возможная величина, следовательно, p^* является эффективной оценкой p.

Таким образом, частота p^* события A в серии n независимых однородных испытаний есть подходящее значение его вероятности, т.е. наилучшая точечная оценка.

Исследуем качество оценивания вероятности p по его частоте p^*. Итак, полагаем, что

$$p \approx \tilde{p} = p^* = \frac{m}{n}.$$

Априори число $\hat{m}$ случайно и подчинено биномиальному закону распределения с параметрами n, p. Согласно теореме Муавра-Лапласа при достаточно больших n (практически при $np(1-p) > 9$) биномиальное распределение может быть с достаточной точностью аппроксимировано нормальным распределением с параметрами $M_{\hat{z}} = np$, $\sigma_{\hat{z}} = \sqrt{np(1-p)}$. В этом случае справедливо соотношение

$$F_{\hat{m}}(z) = F_{\hat{m}}^{\text{б}}(z; n; p) \approx F_{\hat{z}}^{\text{н}}\left(z; np; \sqrt{np(1-p)}\right) = \Phi_1\left(\frac{z - np}{\sqrt{np(1-p)}}\right).$$

Поскольку оценка $\tilde{p} = \hat{m}/n$ связана с $\hat{m}$ линейной зависимостью, она будет распределена приближённо нормально с параметрами

$$M_{\tilde{p}} = p; \quad \sigma_{\tilde{p}} = \sqrt{\frac{p(1-p)}{n}}.$$

Тогда справедливо

$$F_{\tilde{p}}(\vartheta) \approx F_{\tilde{\vartheta}}\left(\vartheta; p; \sqrt{\frac{p(1-p)}{n}}\right) = \Phi_1\left(\frac{(\vartheta - p)\sqrt{n}}{\sqrt{p(1-p)}}\right). \qquad (3.3.4)$$

Так как закон распределения (3.3.4) оценки $\tilde{p}$ симметричен относительно оцениваемой вероятности p, доверительный интервал $I_{\beta,n}(p)$ будет симметричен относительно оценки $\tilde{p}$. Для определения данного интервала достаточно знать половину его длины, которая равна максимальной с доверительной вероятностью $\beta(p)$ абсолютной погрешности $\varepsilon(p)$:

$$\varepsilon'_{\beta,n} = \varepsilon''_{\beta,n} = \varepsilon_{\beta,n} = \varepsilon.$$

В результате доверительная вероятность для p будет определяться следующим равенством:

$$\beta = \beta_{I,n} = \beta_{\varepsilon,n} = P(|\,\tilde{p} - p\,| \le \varepsilon) = 2\Phi_0\left(\frac{\varepsilon}{\sigma_{\tilde{p}}}\right) =$$
$$= 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) \approx 2\Phi_0\left(\frac{\varepsilon\sqrt{n}}{\sqrt{\tilde{p}(1-\tilde{p})}}\right). \qquad (3.3.5)$$

Разрешив уравнение (3.3.5) относительно ε, получим

$$\varepsilon = \varepsilon_{\beta,n} = t_\beta\sqrt{\frac{p(1-p)}{n}} \approx t_\beta\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}} = \tilde{\varepsilon}\,, \qquad (3.3.6)$$

откуда

$$I = I_{\beta,n} = [\,p';\, p''\,] = [\,\tilde{p} - \varepsilon;\; \tilde{p} + \varepsilon\,] \approx [\,\tilde{p} - \tilde{\varepsilon};\; \tilde{p} + \tilde{\varepsilon}\,]. \qquad (3.3.7)$$

В выражении (3.3.6) величина t_β – квантиль нормированного нормального распределения:

$$t_\beta = \Phi_0^{-1}\left(\frac{\beta}{2}\right) \quad \text{или} \quad t_\beta = \Phi_1^{-1}\left(\frac{1+\beta}{2}\right).$$

Значения функции t_β приведены в приложении 4.

Если необходимые точность ε и надёжность β заданы, то потребное для их обеспечения число $n_{\beta,n}$ испытаний находится из уравнения (3.3.6):

$$n = n_{\beta,n} \ge \frac{p(1-p)}{\varepsilon^2}t_\beta^2 \approx \frac{\tilde{p}(1-\tilde{p})}{\varepsilon^2}t_\beta^2\,, \qquad (3.3.8)$$

Формулы (3.3.5) – (3.3.8) определяют решения трёх основных задач исследования качества статистического оценивания (см. § 3.2) применительно к оценке вероятности случайного события по его частоте в серии n независимых однородных испытаний.

Из соотношения (3.3.8) видно, что потребный объём выборки обратно пропорционален квадрату максимальной вероятной погрешности ε оценки $\tilde{p}$ и пропорционален квадрату функции t_β, который растёт быстрее, чем β. Поэтому для оценивания вероятности случайного события по его частоте с достаточной точностью и надёжностью требуется проведение довольно длинной серии испытаний. Сказанное иллюстрируется табл.3.1, в которой приведены потребные числа $n_{0,95;\varepsilon}$ испытаний, обеспечивающие с доверительной вероятностью β = 0,95 необходимую точность ε оценивания различных значений вероятности p.

Таблица 3.1

Зависимость числа испытаний от требуемой доверительной вероятности

ε	*p*				
	0,9	0,8	0,7	0,6	0,5

ε	p				
	0,9	0,8	0,7	0,6	0,5
0,05	139	246	323	369	385
0,01	3458	6147	8068	9220	9604

Из табл.3.1 видно, что потребное число $n_{\beta;\varepsilon}$ испытаний растёт не только с увеличением неоходимой точности оценивания, но и с приближением истинного значения p оцениваемой вероятности к 0,5. Это объяснимо, поскольку при $p = 0{,}5$ дисперсия оценки $\tilde{p} = p^*$ достигает максимального значения, равного $0{,}25/n$ [см. формулу (3.3.3)]. Указанный факт используется для определения верхней границы потребного числа испытаний. Так, полагая $p = 0{,}5$, $\beta = 0{,}95$, имеем значение $t_\beta = t_{0,95} = 1{,}96 \approx 2$ (см. приложение 4). В соответствии с выражением (3.3.8) получаем

$$n = n_{0,95;n} \geq \frac{0{,}5 \cdot 0{,}5}{\varepsilon^2} 1{,}96^2 \approx \frac{1}{\varepsilon^2}. \qquad (3.3.9)$$

П р и м е р 3.1. В процессе эксперимента выполнено 200 опытов, частота события A оказалась $p^* = 0{,}34$.

1. Построить 85%-й доверительный интервал для вероятности события A.

2. Найти доверительную вероятность β для вероятности события A, если максимальная вероятная погрешность $\varepsilon_\beta = 0{,}1$.

▼ 1) Для $\beta = 0{,}85$ в приложении 4 находим $t_\beta = 1{,}439$. Тогда по формуле (3.3.6) оценка максимальной вероятной ошибки составит

$$\tilde{\varepsilon} = 1{,}439 \sqrt{\frac{0{,}34(1 - 0{,}34)}{200}} = 0{,}048.$$

Находим доверительный интервал из соотношения (3.3.7)

$$I_{0,85;\,200} \approx [0{,}34 - 0{,}048;\ 0{,}34 + 0{,}048] = [0{,}292;\ 0{,}388].$$

2) По формуле (3.3.5) находим доверительную вероятность

$$\beta_{0,1;200} \approx 2\Phi_0\left(\frac{0{,}1\sqrt{200}}{\sqrt{0{,}34(1-0{,}34)}}\right) = 2\Phi_0\left(\frac{1{,}414}{0{,}474}\right) = 2\Phi_0(2{,}98) = 2 \cdot 0{,}4986 = 0{,}9972.$$

Значение функции $\Phi_0(x)$ взято из приложения 2.

▲

П р и м е р 3.2. В процессе эксперимента выполняются опыты, частота события составляет $p^* = 0{,}7$.

1. Определить потребный объём выборки, чтобы максимальная вероятная погрешность оценки p^* составляла $\varepsilon \le 0{,}05$ при доверительной вероятности $\beta = 0{,}9$.

2. Найти верхнюю границу потребного числа опытов при любой частоте события .

▼ 1) По заданному β находим $t_\beta = 1{,}643$. Тогда в соответствии с формулой (3.3.8) потребный объём выборки составит

$$n \ge \frac{0{,}7(1-0{,}7)}{(0{,}05)^2} \cdot 1{,}643 = 138 \,.$$

2) Из выражения (3.3.9) имеем

$$n \ge \frac{1}{0{,}0025} = 400 \,.$$

▲