3. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ И ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

3.1. Понятия интервальной оценки и доверительного интервала

При оценивании неизвестных параметров наряду с рассмотренными выше *точечными оценками* используются также *интервальные оценки*. В отличие от точечной оценки интервальная оценка позволяет получить вероятностную характеристику точности оценивания неизвестного параметра.

Пусть $\vec{X}_n$ — *случайная выборка объема* n из *генеральной совокупности* X с функцией распределения $F(x;\theta)$, зависящей от параметра θ, значение которого неизвестно. Предположим, что для параметра θ построен интервал $\big(\underline{\theta}(\vec{X}_n),\ \overline{\theta}(\vec{X}_n)\big)$, где $\underline{\theta}(\vec{X}_n)$ и $\overline{\theta}(\vec{X}_n)$ являются функциями случайной выборки $\vec{X}_n$, такими, что выполняется равенство

$$\mathbf{P}\left\{\underline{\theta}(\vec{X}_n) < \theta < \overline{\theta}(\vec{X}_n)\right\} = \gamma. \tag{3.1}$$

В этом случае интервал $(\underline{\theta}(\vec{X}_n), \overline{\theta}(\vec{X}_n))$ называют ***интервальной оценкой*** для параметра θ с ***коэффициентом доверия*** γ (или, сокращенно, ***γ-доверительной интервальной оценкой***), а $\underline{\theta}(\vec{X}_n)$ и $\overline{\theta}(\vec{X}_n)$ соответственно ***нижней*** и ***верхней границами*** интервальной оценки.

Интервальная оценка $(\underline{\theta}(\vec{X}_n), \overline{\theta}(\vec{X}_n))$ представляет собой интервал со случайными границами, который с заданной вероятностью γ накрывает неизвестное истинное значение параметра θ. Таким образом, для различных *реализаций случайной выборки* $\vec{X}_n$, т.е. для различных элементов *выборочного пространства* $\mathcal{X}_n$, *статистики* $\underline{\theta}(\vec{X}_n)$ и $\overline{\theta}(\vec{X}_n)$ могут принимать различные значения. Более того, согласно (3.1), существует подмножество $\mathcal{K} \subset \mathcal{X}_n$, такое, что если $\vec{x}_n \in \mathcal{K}$, то $\theta \notin \left(\underline{\theta}(\vec{x}_n),\, \overline{\theta}(\vec{x}_n)\right)$.

При этом вероятностной характеристикой точности оценивания параметра θ является случайная величина

$$l(\vec{X}_n) = \overline{\theta}(\vec{X}_n) - \underline{\theta}(\vec{X}_n),$$

которая для любой реализации $\vec{x}_n$ случайной выборки $\vec{X}_n$ есть длина интервала $\left(\underline{\theta}(\vec{x}_n),\, \overline{\theta}(\vec{x}_n)\right)$.

Интервал $\left(\underline{\theta}(\vec{x}_n),\, \overline{\theta}(\vec{x}_n)\right)$ называют ***доверительным интервалом*** для параметра θ с коэффициентом доверия γ или ***γ-доверительным интервалом.***

Заметим, что наряду с термином „коэффициент доверия" широко используют также термины ***доверительная вероятность*** и ***уровень доверия.*** При этом коэффициент доверия γ чаще всего выбирают равным 0,9, 0,95 или 0,99, т.е. близким к 1.

В некоторых ситуациях (например, при рассмотрении дискретных случайных величин) вместо равенства (3.1) удается обеспечить лишь неравенство

$$\mathbf{P}\left\{\underline{\theta}(\vec{X}_n) < \theta < \overline{\theta}(\vec{X}_n)\right\} \geqslant \gamma,$$

т.е. построить интервальную оценку для параметра θ с коэффициентом доверия, не меньшим γ. Иногда требуется оценить параметр θ только снизу или только сверху. При этом, если

$$\mathbf{P}\left\{\underline{\theta}(\vec{X}_n) < \theta\right\} = \gamma,$$

то статистику $\underline{\theta}(\vec{X}_n)$ называют ***односторонней нижней γ-доверительной границей*** для параметра θ. Аналогично, если

$$\mathbf{P}\left\{\theta < \overline{\theta}(\vec{X}_n)\right\} = \gamma,$$

то статистику $\overline{\theta}(\vec{X}_n)$ называют ***односторонней верхней γ-доверительной границей*** для параметра θ.

Пример 3.1. Пусть θ — среднее значение предела прочности X некоторого материала, которое оценивают независимо друг от друга в каждой из N различных лабораторий по результатам n независимых натурных испытаний. Иначе говоря, среднее значение предела прочности в каждой лаборатории оценивают по „своим“ *экспериментальным данным*, представленным *выборкой объема* n, и в каждой лаборатории получают „свои“ значения верхней и нижней границ γ-доверительного интервала (рис. 3.1).

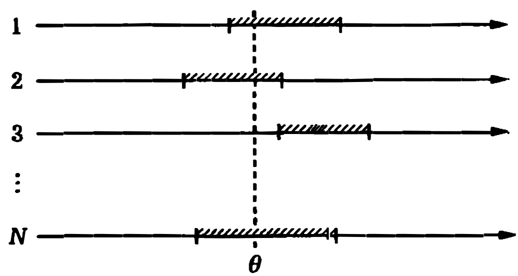

Рис. 3.1

Возможны случаи, когда γ-доверительный интервал для параметра θ не накрывает его истинного значения. Если M — число таких случаев, то при больших значениях N должно выполняться приближенное равенство $\gamma \approx (N - M)/N$. Таким образом, если опыт — получение выборки объема n в лаборатории, то уровень доверия γ — доля тех опытов (при их многократном независимом повторении), в каждом из которых γ-доверительный интервал накрывает истинное значение оцениваемого параметра.

3.2. Построение интервальных оценок

Пусть $\vec{X}_n$ — *случайная выборка объема* n из *генеральной совокупности* X с функцией распределения $F(x;\theta)$, зависящей от параметра θ, значение которого неизвестно. Рассмотрим один из наиболее распространенных методов построения *интервальных оценок* для θ, связанный с использованием ***центральной статистики*** — любой *статистики* $T(\vec{X}_n, \theta)$, функция распределения которой

$$F_T(t) = \mathbf{P}\{T(\vec{X}_n, \theta) < t\}$$

не зависит от параметра θ. Примеры центральных статистик приведены в **3.3**.

Для упрощения дальнейших рассуждений будем предполагать следующее:

1) функция распределения $F_T(t)$ является непрерывной и возрастающей;

2) заданы такие положительные числа α и β, что *коэффициент доверия* $\gamma = 1 - \alpha - \beta$;

3) для любой выборки $\vec{x}_n$ из генеральной совокупности X функция $T(\vec{x}_n, \theta)$ является непрерывной и возрастающей (убывающей) функцией параметра $\theta \in \Theta$.

Согласно допущению 1, для любого $q \in (0, 1)$ существует единственный корень h_q уравнения $F_T(t) = q$, который называют квантилью уровня q функции распределения $F_T(t)$ случайной величины $T(\vec{X}_n, \theta)$. Таким образом, согласно допущению 2, имеют место равенства

$$\mathbf{P}\{h_\alpha < T(\vec{X}_n, \theta) < h_{1-\beta}\} = \\ = F_T(h_{1-\beta}) - F_T(h_\alpha) = 1 - \beta - \alpha = \gamma, \quad (3.2)$$

которые справедливы для любых возможных значений параметра θ, так как $T(\vec{X}_n, \theta)$ — центральная статистика, и ее функция распределения $F_T(t)$ не зависит от θ. Для преобразования (3.2) в (3.1), т.е. для построения искомой интервальной оценки, воспользуемся следующими соображениями.

Пусть для определенности функция $T(\vec{x}_n, \theta)$ является возрастающей функцией параметра θ. Тогда, согласно допущению 3, для каждой выборки $\vec{x}_n \in \mathcal{X}_n$ уравнения $T(\vec{x}_n, \theta) = h_\alpha$ и

$T(\vec{x}_n,\theta) = h_{1-\beta}$ имеют единственные решения $\underline{\theta}(\vec{x}_n)$ и $\overline{\theta}(\vec{x}_n)$ соответственно. При этом неравенства

$$h_\alpha < T(\vec{x}_n,\theta) < h_{1-\beta}, \qquad \underline{\theta}(\vec{x}_n) < \theta < \overline{\theta}(\vec{x}_n)$$

являются равносильными, т.е. для любой выборки $\vec{x}_n \in \mathcal{X}_n$ они выполняются или не выполняются одновременно. Таким образом,

$$\gamma = \mathbf{P}\big\{h_\alpha < T(\vec{X}_n,\theta) < h_{1-\beta}\big\} = \mathbf{P}\big\{\underline{\theta}(\vec{X}_n) < \theta < \overline{\theta}(\vec{X}_n)\big\}$$

и $\big(\underline{\theta}(\vec{X}_n),\, \overline{\theta}(\vec{X}_n)\big)$ — искомая интервальная оценка.

Завершая рассуждения, заметим, что фактически построение *доверительного интервала* сводится к выполнению следующих действий:

1) построение центральной статистики $T(\vec{X}_n,\theta)$ с известной функцией распределения $F_T(t)$;

2) представление заданного коэффициента доверия γ в виде $\gamma = 1-\alpha-\beta$;

3) нахождение квантилей h_α и $h_{1-\beta}$ уровня α и $1-\beta$ функции распределения $F_T(t)$;

4) нахождение значений *нижней* $\underline{\theta}(\vec{x}_n)$ и *верхней* $\overline{\theta}(\vec{x}_n)$ *границ* искомой интервальной оценки путем решения уравнений

$$T(\vec{x}_n,\underline{\theta}) = h_\alpha, \qquad T(\vec{x}_n,\overline{\theta}) = h_{1-\beta} \tag{3.3}$$

соответственно в случае, когда $T(\vec{x}_n,\theta)$ — возрастающая функция параметра θ. Если же $T(\vec{x}_n,\theta)$ — убывающая функция параметра θ, то $\underline{\theta}(\vec{X}_n)$ и $\overline{\theta}(\vec{X}_n)$ получают путем решения уравнений

$$T(\vec{x}_n,\underline{\theta}) = h_{1-\beta}, \qquad T(\vec{x}_n,\overline{\theta}) = h_\alpha \tag{3.4}$$

соответственно.

3.3. Примеры построения интервальных оценок

Рассмотрим построение *интервальной оценки* для параметров некоторых часто используемых распределений.

Экспоненциальное распределение. Пусть $\vec{X}_n$ — *случайная выборка объема* n из *генеральной совокупности* X с *экспоненциальным законом* распределения, имеющим плотность распределения

$$p(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0; \\ 0, & x < 0, \end{cases}$$

где $\lambda > 0$ — неизвестный параметр. Требуется построить интервальную оценку для параметра λ по данным случайной выборки $\vec{X}_n$.

В данном случае $\theta = \lambda$. Рассмотрим *статистику*

$$T(\vec{X}_n, \lambda) = 2\lambda n\overline{X},$$

где $\overline{X}$ — *выборочное среднее* для $\vec{X}_n$. Эта статистика имеет χ^2*-распределение с* $2n$ *степенями свободы* (см. **Д.3.1**), т.е. является *центральной статистикой*. Уравнения (3.3) в данном случае принимают вид

$$2\underline{\lambda} n\overline{x} = \chi^2_\alpha(2n), \qquad 2\overline{\lambda} n\overline{x} = \chi^2_{1-\beta}(2n),$$

где $\chi^2_q(2n)$ — квантиль уровня q для χ^2-распределения с $2n$ степенями свободы.

Получаем, что *нижняя* и *верхняя границы* интервальной оценки с *коэффициентом доверия* $\gamma = 1 - \alpha - \beta$ для параметра экспоненциального распределения λ имеют вид

$$\underline{\lambda}(\vec{X}_n) = \frac{\chi^2_\alpha(2n)}{2n\overline{X}}, \qquad \overline{\lambda}(\vec{X}_n) = \frac{\chi^2_{1-\beta}(2n)}{2n\overline{X}}.$$

Нормальное распределение. Пусть $\vec{X}_n$ — случайная выборка объема n из генеральной совокупности X, распределенной по нормальному закону с параметрами μ и σ^2. Рассмотрим некоторые варианты построения интервальных оценок для параметров μ, σ.

Вариант 1 — оценка для математического ожидания при известной дисперсии. В данном случае статистика

$$T(\vec{X}_n,\mu)=\frac{\overline{X}-\mu}{\sigma}\sqrt{n}$$

имеет стандартное нормальное распределение с параметрами $\mu=0$, $\sigma^2=1$, т.е. является центральной статистикой. Функция $T(\vec{X}_n,\mu)$ является убывающей функцией по μ, и система уравнений (3.4) принимает вид

$$\frac{\sqrt{n}\big(\overline{x}-\underline{\mu}(\vec{x}_n)\big)}{\sigma}=u_{1-\beta},\qquad \frac{\sqrt{n}\big(\overline{x}-\overline{\mu}(\vec{x}_n)\big)}{\sigma}=u_{\alpha},$$

где u_q — квантиль уровня q стандартного нормального распределения. Учитывая, что для нормального закона $u_{1-\alpha}=-u_{\alpha}$, получаем следующие нижнюю и верхнюю границы *γ-доверительного интервала* для параметра μ при $\gamma=1-\alpha-\beta$:

$$\underline{\mu}(\vec{x}_n)=\overline{x}-\frac{\sigma}{\sqrt{n}}u_{1-\beta},\qquad \overline{\mu}(\vec{x}_n)=\overline{x}+\frac{\sigma}{\sqrt{n}}u_{1-\alpha}.$$

Вариант 2 — оценка математического ожидания при неизвестной дисперсии. При неизвестной дисперсии статистика

$$T(\vec{X}_n,\mu)=\frac{\overline{X}-\mu}{S(\vec{X}_n)}\sqrt{n}$$

является центральной, так как имеет *распределение Стьюдента* с $n-1$ степенями свободы (см. **Д.3.1**), которое не зависит от μ и σ^2. Система уравнений (3.4) в данном случае принимает вид

$$\frac{\sqrt{n}\big(\overline{x}-\underline{\mu}(\vec{x}_n)\big)}{S(\vec{x}_n)}=t_{1-\beta}(n-1),\qquad \frac{\sqrt{n}\big(\overline{x}-\overline{\mu}(\vec{x}_n)\big)}{S(\vec{x}_n)}=t_{\alpha}(n-1),$$

где $t_q(n-1)$ — квантиль уровня q распределения Стьюдента с $n-1$ степенями свободы. Поскольку плотность распределения Стьюдента — четная функция, то $t_\alpha(n-1) = -t_{1-\alpha}(n-1)$. Отсюда заключаем, что нижняя и верхняя границы интервальной оценки с коэффициентом доверия $\gamma = 1-\alpha-\beta$ для параметра μ в случае с неизвестной дисперсией можно определить по формулам

$$\underline{\mu}(\vec{X}_n) = \overline{X} - \frac{S(\vec{X}_n)}{\sqrt{n}} t_{1-\beta}(n-1), \quad \overline{\mu}(\vec{X}_n) = \overline{X} + \frac{S(\vec{X}_n)}{\sqrt{n}} t_{1-\alpha}(n-1).$$

Вариант 3 — оценка среднего квадратичного отклонения. Рассмотрим статистику

$$T(\vec{X}_n, \sigma) = \frac{(n-1)S^2(\vec{X}_n)}{\sigma^2}.$$

Эта статистика является центральной, так как имеет χ^2-распределение с $n-1$ степенями свободы (см. **Д.3.1**), которое не зависит от μ и σ^2. При этом $T(\vec{x}_n, \sigma)$ — убывающая функция параметра σ. Исходя из этого, согласно (3.4), находим нижнюю и верхнюю границы интервальной оценки для параметра σ с коэффициентом доверия $\gamma = 1-\alpha-\beta$:

$$\underline{\sigma}(\vec{X}_n) = \frac{S(\vec{X}_n)\sqrt{n-1}}{\sqrt{\chi^2_{1-\beta}(n-1)}}, \qquad \overline{\sigma}(\vec{X}_n) = \frac{S(\vec{X}_n)\sqrt{n-1}}{\sqrt{\chi^2_{\alpha}(n-1)}},$$

где $\chi^2_q(n-1)$ — квантиль уровня q для χ^2-распределения с $n-1$ степенями свободы.

Приближенные интервальные оценки. Сначала рассмотрим два частных случая построения таких оценок.

Пусть требуется найти интервальную оценку для математического ожидания в случае, когда закон распределения генеральной совокупности X неизвестен. Предполагаем, что существуют конечные математическое ожидание $\mu = \mathbf{M}X$ и дисперсия $\sigma^2 = \mathbf{D}X$.

Рассмотрим статистику

$$T(\vec{X}_n) = \frac{\overline{X} - \mu}{\sigma}\sqrt{n}.$$

В соответствии с центральной предельной теоремой эта статистика при больших объемах n случайной выборки $\vec{X}_n$ имеет закон распределения, близкий к стандартному нормальному. Поэтому при достаточно больших n неравенства

$$-u_{1-\beta} \leqslant \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \leqslant u_{1-\alpha}$$

выполняются с вероятностью, близкой к величине $\gamma = 1 - \alpha - \beta$, где u_q — квантиль уровня q стандартного нормального распределения. Приведенные неравенства эквивалентны следующим:

$$\overline{X} - \frac{\sigma}{\sqrt{n}}u_{1-\alpha} \leqslant \mu \leqslant \overline{X} + \frac{\sigma}{\sqrt{n}}u_{1-\beta}.$$

Эти неравенства не дают еще интервальной оценки для параметра μ, так как их левая и правая части содержат неизвестный параметр σ. Применяя еще одно приближение, а именно: подставляя в указанные неравенства вместо неизвестного точного значения σ его *оценку* $S(\vec{X}_n)$, получаем нижнюю и верхнюю границы (приближенной) интервальной оценки с коэффициентом доверия $\gamma = 1 - \alpha - \beta$ для математического ожидания μ:

$$\underline{\mu}(\vec{X}_n) = \overline{X} - \frac{S(\vec{X}_n)}{\sqrt{n}}u_{1-\beta}, \qquad \overline{\mu}(\vec{X}_n) = \overline{X} + \frac{S(\vec{X}_n)}{\sqrt{n}}u_{1-\alpha}.$$

Пусть проводится серия из n испытаний по схеме Бернулли и X_i, $i = \overline{1, n}$, — исход i-го испытания („успех“ или „отказ“). По данным случайной выборки $\vec{X}_n = (X_1, \ldots, X_n)$ построим доверительный интервал для вероятности p „успеха“ в каждом отдельном испытании.

Рассмотрим суммарное число „успехов" в серии из n испытаний, т.е. введем случайную величину

$$K(\vec{X}_n) = X_1 + \ldots + X_n,$$

которая имеет биномиальное распределение с параметром p. Для построения доверительного интервала для p воспользуемся статистикой

$$T(\vec{X}_n, p) = \frac{K(\vec{X}_n) - np}{\sqrt{np(1-p)}}.$$

В соответствии с предельной теоремой Муавра — Лапласа статистика $T(\vec{X}_n, p)$ при больших объемах n случайной выборки $\vec{X}_n$ имеет закон распределения, близкий к стандартному нормальному. Тем самым неравенства

$$-u_{1-\beta} \leqslant \frac{K(\vec{X}_n) - np}{\sqrt{np(1-p)}} \leqslant u_{1-\alpha}$$

выполняются с вероятностью, которую при больших n можно считать приближенно равной $\gamma = 1 - \alpha - \beta$. Указанные неравенства могут быть записаны в виде

$$\frac{K(\vec{X}_n)}{n} - \frac{u_{1-\alpha}}{\sqrt{n}} \sqrt{p(1-p)} \leqslant p \leqslant \frac{K(\vec{X}_n)}{n} + \frac{u_{1-\beta}}{\sqrt{n}} \sqrt{p(1-p)}.$$

Эти неравенства еще не дают интервальной оценки параметра p, так как их левая и правая части содержат этот параметр. Поэтому на практике в указанные части неравенств часто подставляют вместо неизвестного точного значения p его оценку $\widehat{p}(\vec{X}_n) = K(\vec{X}_n)/n$. В результате получают следующие верхнюю и нижнюю границы интервальной оценки с коэффициентом доверия $\gamma = 1 - \alpha - \beta$ для параметра p:

$$\underline{p}(\vec{X}_n) = \frac{K(\vec{X}_n)}{n} - \frac{u_{1-\alpha}}{\sqrt{n}} \sqrt{\frac{K(\vec{X}_n)}{n}\left(1 - \frac{K(\vec{X}_n)}{n}\right)},$$

$$\overline{p}(\vec{X}_n) = \frac{K(\vec{X}_n)}{n} + \frac{u_{1-\beta}}{\sqrt{n}} \sqrt{\frac{K(\vec{X}_n)}{n}\left(1 - \frac{K(\vec{X}_n)}{n}\right)}.$$

Подчеркнем, что эти доверительные границы являются приближенными и могут использоваться при достаточно больших объемах наблюдений n.

Приведенный способ построения приближенного доверительного интервала для параметра p биномиального распределения может применяться и в следующей более общей ситуации. Пусть $\widehat{\theta}(\vec{X}_n)$ — *точечная несмещенная оценка* для параметра θ, построенная по данным случайной выборки $\vec{X}_n$. Обозначим через

$$V_n(\theta) = \mathbf{M}\big(\widehat{\theta}(\vec{X}_n) - \theta\big)^2$$

значение дисперсии оценки $\widehat{\theta}(\vec{X}_n)$. Предположим, что оценка $\widehat{\theta}(\vec{X}_n)$ имеет *асимптотически нормальное распределение*. Другими словами, нормированная случайная величина

$$\eta_n = \frac{\widehat{\theta}(\vec{X}_n) - \theta}{\sqrt{V_n(\theta)}}$$

имеет распределение, которое при $n \to \infty$ сходится к стандартному нормальному распределению. В этом случае неравенства

$$-u_{1-\beta} \leqslant \frac{\widehat{\theta}(\vec{X}_n) - \theta}{\sqrt{V_n(\theta)}} \leqslant u_{1-\alpha},$$

где u_q — квантиль уровня q стандартного нормального закона распределения, выполняются с вероятностью, которую при достаточно больших n можно считать приближенно равной

$\gamma = 1 - \alpha - \beta$. Указанные неравенства эквивалентны (см. **3.2**) следующим:

$$\widehat{\theta}(\vec{X}_n) - u_{1-\alpha}\sqrt{V_n(\theta)} \leqslant \theta \leqslant \widehat{\theta}(\vec{X}_n) + u_{1-\beta}\sqrt{V_n(\theta)}.$$

Записанные неравенства еще не дают интервальной оценки для θ, так как их левая и правая части содержат неизвестный параметр θ. Подставляя в левую и правую части указанных неравенств вместо θ оценку $\widehat{\theta}(\vec{X}_n)$, получаем окончательно следующие *нижнюю* и *верхнюю границы* для параметра θ с *коэффициентом доверия* $\gamma = 1 - \alpha - \beta$:

$$\underline{\theta}(\vec{X}_n) = \widehat{\theta}(\vec{X}_n) - u_{1-\alpha}\sqrt{V_n(\widehat{\theta}(\vec{X}_n))},$$

$$\overline{\theta}(\vec{X}_n) = \widehat{\theta}(\vec{X}_n) + u_{1-\beta}\sqrt{V_n(\widehat{\theta}(\vec{X}_n))}.$$

Изложенный метод является приближенным и может применяться при достаточно большом объеме случайной выборки. Заметим, что его использование фактически связано с „двойным приближением", а именно: закон распределения оценки $\widehat{\theta}(\vec{X}_n)$ заменяют нормальным и, кроме того, в приведенных формулах для границ $\underline{\theta}(\vec{X}_n)$, $\overline{\theta}(\vec{X}_n)$ интервальной оценки в дисперсию $V_n(\theta)$ вместо точного значения θ подставляют его оценку $\widehat{\theta}(\vec{X}_n)$. При малых и средних объемах случайной выборки применение указанного метода может приводить к значительным ошибкам. Поэтому использовать его следует с достаточной степенью осторожности и лишь в качестве первого приближения.

Пример 3.2. Рассмотрим построение приближенного доверительного интервала для параметра p биномиального распределения. Пусть проводилось $n = 16$ независимых испытаний с неизвестной вероятностью p „успеха" в каждом испытании, при этом наблюдалось $k = 8$ „успехов". Определим значения границ доверительного интервала для p с коэффициентом доверия $\gamma = 0{,}9$.

Значение точечной оценки параметра p определяется как

$$\widehat{p} = \frac{k}{n},$$

дисперсия этой оценки [XVI]

$$V_n(p) = \frac{p(1-p)}{n}.$$

Применяя приведенные выше формулы, получаем следующие значения для нижней и верхней границ доверительного интервала:

$$\underline{p} = \widehat{p} - u_{0,95}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 0{,}294,$$

$$\overline{p} = \widehat{p} + u_{0,95}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 0{,}706.$$

3.4. Метод доверительных множеств

Пусть $\vec{X}_n$ — *случайная выборка объема* n из *генеральной совокупности* X, закон распределения которой зависит от r-мерного вектора параметров $\vec{\theta}$. Каждому фиксированному значению вектора параметров $\vec{\theta}$ поставим в соответствие такое множество $H_{\vec{\theta}}$ из *выборочного пространства* $\mathcal{X}_n$, что

$$\mathbf{P}\{\vec{X}_n \in H_{\vec{\theta}}\} \geqslant \gamma,$$

где γ — заданный *коэффициент доверия.*

Как известно (см. **3.1**), *нижняя* и *верхняя границы интервальной оценки* (множества в $\mathbb{R}$) являются случайными величинами, поскольку они функции случайной выборки. Теперь в $\mathbb{R}^r$ рассмотрим такое множество $D_{\vec{X}_n}$ со случайной границей, чтобы при каждом фиксированном значении вектора параметров $\vec{\theta}$ случайные события

$$\{\vec{\theta} \in D_{\vec{X}_n}\}, \qquad \{\vec{X}_n \in H_{\vec{\theta}}\}$$

были эквивалентны, т.е.

$$\mathbf{P}\{\vec{\theta} \in D_{\vec{X}_n}\} \geqslant \gamma.$$

Полученную таким образом совокупность множеств $D_{\vec{X}_n}$ называют ***системой γ-доверительных множеств***, а рассмотренную процедуру — ***методом доверительных множеств*** (методом Неймана) для параметра θ.

Если θ — скаляр, то метод доверительных множеств имеет простую и наглядную графическую интерпретацию (рис. 3.2). Поэтому все дальнейшие рассуждения проведем именно для этого случая, т.е. при $r = 1$.

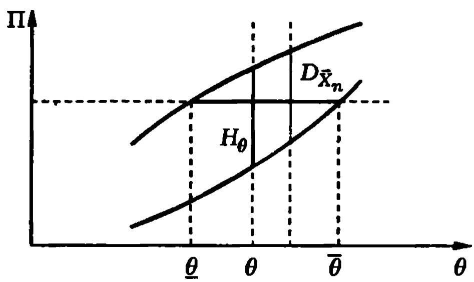

Рис. 3.2

Заметим, что процедура построения доверительных множеств $D_{\vec{X}_n}$ основана на выборе множеств H_θ, а это может быть реализовано различными способами, в том числе и с использованием некоторой *статистики* $\Pi = \Pi(\vec{X})$. Зачастую в качестве статистики $\Pi(\vec{X}_n)$ используют *несмещенную точечную оценку* параметра θ.

Для упрощения дальнейших рассуждений функцию распределения $F_\Pi(t, \theta)$ статистики $\Pi(\vec{X}_n)$ будем предполагать непрерывной, возрастающей по t и убывающей по θ.

Каждому возможному значению параметра θ поставим в соответствие значения $t_1 = t_1(\theta)$, $t_2 = t_2(\theta)$, выбираемые из условий

$$F_\Pi(t_1, \theta) = \alpha, \qquad F_\Pi(t_2, \theta) = 1 - \beta. \tag{3.5}$$

Таким образом, $t_1(\theta)$, $t_2(\theta)$ являются соответственно квантилями уровней α и $1-\beta$ для функции распределения $F_{\Pi}(t,\theta)$ статистики $\Pi(\vec{X}_n)$. При этом выполняется равенство

$$\mathbf{P}\left\{t_1(\theta) \leqslant \Pi(\vec{X}_n) \leqslant t_2(\theta)\right\} = \gamma,$$

где $\gamma = 1-\alpha-\beta$. Множество значений статистики $\Pi(\vec{X}_n)$, принадлежащих отрезку $[t_1(\theta),\, t_2(\theta)]$, обозначим H_θ и назовем ***γ-зоной*** (***гамма-зоной***) для θ (см. рис. 3.2). Как следует из этого определения, для любого возможного значения параметра θ вероятность того, что статистика $\Pi(\vec{X}_n)$ попадет в γ-зону, равна γ.

Далее, каждому значению статистики $\Pi(\vec{X}_n)$ поставим в соответствие *интервал* тех значений θ, для которых данное значение статистики $\Pi(\vec{X}_n)$ попадает в γ-зону (см. рис. 3.2). Значения нижней $\underline{\theta}(\vec{X}_n)$ и верхней $\overline{\theta}(\vec{X}_n)$ границ этого интервала определяются из условий

$$t_2\big(\underline{\theta}(\vec{x}_n)\big) = \Pi(\vec{x}_n), \qquad t_1\big(\overline{\theta}(\vec{x}_n)\big) = \Pi(\vec{x}_n),$$

которые в силу (3.5) эквивалентны следующим:

$$F\big(\Pi(\vec{x}_n), \overline{\theta}\big) = \alpha, \qquad F\big(\Pi(\vec{x}_n), \underline{\theta}\big) = 1-\beta. \tag{3.6}$$

Построенный таким образом интервал является γ-доверительной оценкой для параметра θ. Действительно, при любом возможном, а следовательно, и при неизвестном истинном значении θ интервал $\big(\underline{\theta}(\vec{X}_n),\, \overline{\theta}(\vec{X}_n)\big)$ накрывает значение θ тогда и только тогда, когда наблюдаемое значение статистики $\Pi(\vec{X}_n)$ попадает в γ-зону H_θ для данного значения θ. Тем самым, согласно определению γ-зоны, выполняется равенство

$$\mathbf{P}\left\{\underline{\theta}(\vec{X}_n) \leqslant \theta \leqslant \overline{\theta}(\vec{X}_n)\right\} = \gamma.$$

Если функция распределения $F(t;\theta)$ возрастает по параметру θ, то границы $t_1(\theta)$ и $t_2(\theta)$ γ-зоны убывают по θ. Повторяя предыдущие рассуждения, заключаем, что в этом случае значения нижней и верхней границ формально (см. **3.2**) определяются из условий

$$F_{\Pi}\big(\Pi(\vec{x}_n),\underline{\theta}(\vec{x}_n)\big)=\alpha,\qquad F_{\Pi}\big(\Pi(\vec{x}_n),\overline{\theta}(\vec{x}_n)\big)=1-\beta. \qquad (3.7)$$

Этот метод применяют аналогичным образом и в тех случаях, когда статистика $\Pi(\vec{X}_n)$ является дискретной случайной величиной. Рассмотрим, например, случай, когда статистика $\Pi(\vec{X}_n)$ принимает неотрицательные целые значения 0, 1, 2, ...

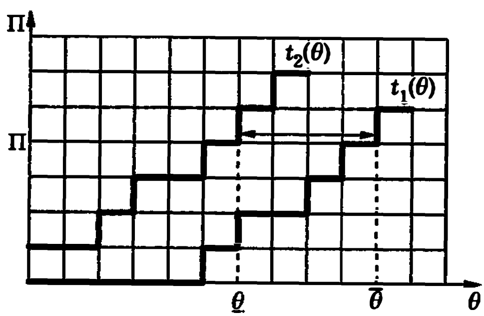

Рис. 3.3

В отличие от непрерывного случая, рассмотренного выше, границы γ-зоны теперь становятся ступенчатыми кривыми (рис. 3.3). При данном фиксированном значении θ границу $t_1(\theta)$ γ-зоны определим как максимальное из чисел k, таких, что выполняется неравенство

$$\mathbf{P}\big\{\Pi(\vec{X}_n)\geqslant k\big\}=1-F(k;\theta)\geqslant 1-\alpha.$$

Границу $t_2(\theta)$ γ-зоны определим как минимальное из чисел k, удовлетворяющих неравенству

$$\mathbf{P}\big\{\Pi(\vec{X}_n)\leqslant k\big\}=F(k+1;\theta)\geqslant 1-\beta.$$

Нижнюю $\underline{\theta}(\vec{X}_n)$ и верхнюю $\overline{\theta}(\vec{X}_n)$ границы интервальной оценки параметра θ с коэффициентом доверия не меньше $\gamma=$

$= 1-\alpha-\beta$ определим как минимальное и максимальное значения среди всех θ, удовлетворяющих неравенствам

$$t_1(\theta) \leqslant \Pi(\vec{X}_n) \leqslant t_2(\theta), \tag{3.8}$$

т.е. среди всех θ, принадлежащих γ-зоне при данном значении статистики $\Pi(\vec{X}_n)$, полученном в результате эксперимента. Неравенства (3.8) эквивалентны неравенствам

$$\begin{cases} F_\Pi\big(\Pi(\vec{x}_n)+1, \theta\big) \geqslant \alpha, \\ F_\Pi\big(\Pi(\vec{x}_n), \theta\big) \leqslant 1-\beta. \end{cases}$$

Отсюда получаем, что если функция распределения статистики $\Pi(\vec{X}_n)$ убывает по θ, то нижнюю и верхнюю границы γ-интервальной оценки для θ формально (см. **3.2**) можно определить из уравнений

$$\begin{cases} F_\Pi\big(\Pi(\vec{X}_n), \underline{\theta}(\vec{X}_n)\big) = 1-\beta, \\ F_\Pi\big(\Pi(\vec{X}_n)+1, \overline{\theta}(\vec{X}_n)\big) = \alpha. \end{cases} \tag{3.9}$$

Аналогично, если функция распределения статистики $\Pi(\vec{X}_n)$ возрастает по θ, то границы γ-интервальной оценки для θ формально (см. **3.2**) можно найти из уравнений

$$\begin{cases} F_\Pi\big(\Pi(\vec{X}_n)+1, \underline{\theta}(\vec{X}_n)\big) = \alpha, \\ F_\Pi\big(\vec{X}_n, \overline{\theta}(\vec{X}_n)\big) = 1-\beta, \end{cases} \tag{3.10}$$

где коэффициент доверия $\gamma = 1-\alpha-\beta$.

Рассмотрим далее в качестве примеров построение интервальных оценок для параметров биномиального распределения и распределения Пуассона.

Интервальная оценка Клоппера — Пирсона для параметра биномиального распределения. Пусть дискретная случайная величина X_i, $i=\overline{1,n}$, характеризует исход i-го испытания в серии из n испытаний, проводимых по схеме Бернулли. Тогда случайная величина $K = X_1 + \ldots + X_n$ — число успехов в n испытаниях. При этом $K = K(\vec{X}_n)$ — функция случайной выборки $\vec{X}_n = (X_1, \ldots, X_n)$. В рассматриваемом случае $\Pi(\vec{X}_n) = K(\vec{X}_n)$.

Функция распределения статистики $K(\vec{X}_n)$ имеет вид

$$F(x;p) = \begin{cases} \sum\limits_{j<x} C_n^j p^j (1-p)^{n-j}, & x > 0; \\ 0, & x \leqslant 0. \end{cases}$$

Эта функция убывающая по p. Применяя общую формулу (3.9), получаем, что нижняя и верхняя границы интервальной оценки с коэффициентом доверия $\gamma = 1 - \alpha - \beta$ для параметра p (см. **3.2**) определяются из следующих уравнений:

$$\sum_{j=0}^{K(\vec{X}_n)-1} C_n^j \underline{p}^j(\vec{X}_n)(1 - \underline{p}(\vec{X}_n))^{n-j} = 1 - \beta \quad \text{при} \quad K(\vec{X}_n) \geqslant 1,$$

$$\sum_{j=0}^{K(\vec{X}_n)} C_n^j \overline{p}^j(\vec{X}_n)(1 - \overline{p}(\vec{X}_n))^{n-j} = \alpha \quad \text{при} \quad K(\vec{X}_n) \leqslant n - 1.$$

Эти уравнения называются ***уравнениями Клоппера — Пирсона***. При $K(\vec{X}_n) = 0$ нижняя граница $\underline{p}(\vec{X}_n) = 0$. При $K(\vec{X}_n) = n$ верхняя граница $\overline{p}(\vec{X}_n) = 1$.

Заметим, что приведенные уравнения Клоппера — Пирсона могут быть также выражены через неполную бета-функцию (см. **Д.3.1**):

$$\mathrm{B}_{\underline{p}(\vec{X}_n)}(K(\vec{X}_n), n - K(\vec{X}_n) + 1) = \beta,$$

$$\mathrm{B}_{\overline{p}(\vec{X}_n)}(K(\vec{X}_n) + 1, n - K(\vec{X}_n)) = 1 - \alpha.$$

Пример 3.3. Пусть число испытаний $n = 16$, а число наблюдаемых „успехов" $K = 8$, коэффициент доверия $\gamma = 0{,}95$.

Полагая $\alpha=\beta=0{,}025$, получаем

$$\underline{p}=0{,}247, \qquad \overline{p}=0{,}753.$$

Доверительный интервал для параметра распределения Пуассона. Пусть X — дискретная случайная величина, имеющая распределение Пуассона с неизвестным параметром λ. Требуется построить доверительный интервал для параметра λ на основе наблюдаемого значения d_* случайной величины X.

Согласно предположению, функция распределения случайной величины X имеет вид

$$F(x;\lambda)=\begin{cases}\sum\limits_{j<x}\dfrac{\lambda^j}{j!}e^{-\lambda}, & x>0;\\ 0, & x\leqslant 0.\end{cases}$$

Это функция, убывающая по λ. Применяя снова формулы (3.9), получаем уравнения

$$e^{-\underline{\lambda}}\sum_{j=0}^{d_*-1}\frac{\underline{\lambda}^j}{j!}=1-\beta, \qquad e^{-\overline{\lambda}}\sum_{j=0}^{d_*}\frac{\overline{\lambda}^j}{j!}=\alpha,$$

решая которые находим значения нижней и верхней границ доверительного интервала для λ с коэффициентом доверия $\gamma=1-\alpha-\beta$. При $d_*=0$ значение нижней границы $\underline{\lambda}=0$.

3.5. Решение типовых примеров

Пример 3.4. При помощи вольтметра, точность которого характеризуется средним квадратичным отклонением 0,2В, проведено 10 измерений напряжения бортовой батареи. Найдем *доверительный интервал* для истинного значения напряжения батареи с *коэффициентом доверия* $\gamma=0{,}95$, если среднее арифметическое результатов наблюдений $\overline{x}=50{,}2$В. Контролируемый признак имеет нормальный закон распределения.