Интервальные методы оценивания для ограниченных по величине погрешностей часто более точны, чем усреднение выборки

реферат статьи Дж.У. Уолстера и В. Крейновича

С.П. Шарый

ФИЦ ИВТ и НГУ, Новосибирск, Россия

26 января 2026 г.

Резюме работ

Дывак Н.П. Интервальные модели ошибок в прикладных задачах // Материалы Всесоюзной конференции «Актуальные проблемы прикладной математики», 20–22 мая 1991 года, Саратов. Т. 1. – Саратовское областное правление Союза НИО СССР, Дом науки и техники Союза НИО СССР, СГУ им. Н.Г. Чернышевского: Саратов, 1991. – С. 70–75.

Walster G.W., Kreinovich V. For unknown-but-bounded errors, interval estimates are often better than averaging // *ACM SIGNUM Newsletter*. – 1996. – Vol. 31, Issue 2. – P. 6–19.

Постановка задачи

Для многих измерительных устройств мы знаем лишь их максимально возможную погрешность $\varepsilon > 0$.

Погрешность $\Delta x = \tilde{x} - x$ случайна, она может становиться равной $-\varepsilon$ или ε, но мы не имеем никакой другой информации о погрешности.

Обычно можно наверняка считать, что погрешности последовательных измерений — это независимые случайные величины. В этом случае методы математической статистики позволяют улучшить оценку для x, выполнив её измерения несколько раз и скомбинировав их результаты.

В этой работе обсуждаются два метода:

- простейший статистический метод усреднения результатов измерений, наиболее широко применяемый на практике, и
- метод, основанный на идеях интервального анализа.

Усреднение или пересечение?

Этот доклад — по материалам статьи Дж.У. Уолстера и В. Крейновича, которая продолжила работу Н.П. Дывака.

Основной математический результат — теорема Дывака–Крейновича, которая впервые сформулирована и обоснована для частного случая в работе Н.П. Дывака, а в статье Дж.У. Уолстера и В. Крейновича дано её полное доказательство.

Рассмотрим ситуацию, когда погрешность $\Delta x = \tilde{x} - x^*$ некоторого измерительного устройства ограничена по величине предельным значением $\varepsilon > 0$, но подчиняется при этом законам теории вероятностей, т. е. является вероятностно-случайной величиной, хотя и с не вполне известной функцией распределения.

Определение 1.

Пусть задано число $\alpha \in]0,1[$, которое назовём *уровнем значимости* (уровнем доверительной вероятности). Предположим также, что неизвестная величина x^* оценивается с помощью какого-либо метода вероятностной статистики по выборке случайных результатов измерений $\tilde{x}_1$, $\tilde{x}_2$, ..., $\tilde{x}_n$, так что получена оценка $\hat{x}$, которая также является случайной величиной.

Назовём *вероятностной точностью* оценки $\hat{x}$ наименьшее число δ, для которого неравенство

$$|\hat{x} - x^*| \leq \delta$$

выполняется с вероятностью $\geq \alpha$.

Иными словами, вероятностная точность — наименьшее из чисел δ, для которых $\mathrm{P}\left(|\hat{x} - x^*| \leq \delta\right) \geq \alpha$.

Усреднение в статистике

Рассмотрим далее ситуацию, когда математическое ожидание погрешности измерений равно нулю, т. е. $\mathsf{E}\left(\Delta x\right) = 0$.

Её обычно интерпретируют как отсутствие в измерениях систематической компоненты погрешности.

Тогда после n измерений можем взять оценкой для x^* величину

$$\overline{x} := \frac{\tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_n}{n}$$

— среднее арифметическое всех результатов выполненных измерений.

Усреднение в статистике

Смысл усреднения в том, что случайные погрешности, которые могут принимать разные знаки, компенсируют в этом выражении друг друга и результат в самом деле получится точнее.

В капитальной книге

> Маликов М.Ф. *Основы метрологии. Часть первая. Учение об измерении. – Москва: Комитет по делам мер и измерительных приборов при Совете Министров СССР, 1949. – 479 с.*

дано метрологическое обоснование того, что усреднение (возможно, с весами, отражающими неравноценность измерений) является наиболее адекватным выражением для обработки многократных наблюдений.

Погрешность этой оценки —

$$\Delta\overline{x} = \overline{x} - x^* = \overline{x} - \frac{nx^*}{n} = \frac{(\tilde{x}_1 - x^*) + (\tilde{x}_2 - x^*) + \ldots + (\tilde{x}_n - x^*)}{n}$$

$$= \frac{\Delta x_1 + \Delta x_2 + \ldots + \Delta x_n}{n}.$$

Здесь все Δx_i — независимые случайные величины с одинаковыми распределениями, имеющие нулевые средние.

В силу Центральной Предельной Теоремы теории вероятностей случайная величина $\Delta\overline{x}$ при достаточно больших n близка к нормальному гауссовскому распределению с плотностью

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right),$$

где a и σ — некоторые постоянные.

Нормальное распределение определяется двумя параметрами — средним (математическим ожиданием), равным a, и стандартным (среднеквадратичным) отклонением σ.

Для больших n достаточно оценить эти два параметра, и затем сможем использовать полученные оценки, работая с предельными нормальными распределениями.

Для получения интервальных оценок погрешности Δx теория вероятностей рекомендует тогда, задавшись каким-то числом k (обычно равным 1, 2, 3, ...), брать оценку в виде

$$\left[\, \mathsf{E}\,(\Delta x) - k\,\sigma(\Delta x), \mathsf{E}\,(\Delta x) + k\,\sigma(\Delta x)\,\right]$$

в зависимости от желаемого уровня доверительной вероятности.

Например, если $k = 2$, то выписанный интервал содержит истинное значение с вероятностью $\approx 95\%$, а если $k = 3$, то с вероятностью $\approx 99.7\%$, и т. д.

Результирующая погрешность $\Delta\overline{x}$ имеет нулевое среднее значение, так как равны нулю средние для всех слагаемых Δx_i, $i = 1, 2, \ldots, n$.

Поскольку погрешности Δx_i являются независимыми случайными величинами с одинаковыми распределениями вероятностей, то в силу известных свойств дисперсии

$$\mathsf{D}(\Delta\overline{x}) = \mathsf{D}\left(\frac{\Delta x_1 + \Delta x_2 + \ldots + \Delta x_n}{n}\right) =$$

$$= \frac{1}{n^2}\left(\mathsf{D}(\Delta x_1) + \mathsf{D}(\Delta x_2) + \ldots + \mathsf{D}(\Delta x_n)\right) = \frac{1}{n}\,\mathsf{D}(\Delta x_i).$$

Поэтому стандартное отклонение $\sigma(\Delta\overline{x})$ для $\Delta\overline{x}$, т. е. $\sqrt{\mathsf{D}(\Delta\overline{x})}$, равно

$$\frac{\sigma(\Delta x_i)}{\sqrt{n}}\,,$$

где $\sigma(\Delta x_i)$ — стандартное отклонение каждого Δx_i.

Стандартное отклонение = квадратный корень из дисперсии случайной величины, её второго центрального момента, равного

$$\mathsf{E}\left(\left(\Delta x_i - \mathsf{E}\left(\Delta x_i\right)\right)^2\right).$$

Поскольку $\mathsf{E}\left(\Delta x_i\right) = 0$, можем заключить, что σ^2 равен среднему значению величины $(\Delta x_i)^2$.

Из-за того, что погрешность $\leq \varepsilon$, получается

$$(\Delta x_i)^2 \leq \varepsilon^2.$$

Следовательно, среднее значение $\mathsf{E}\left((\Delta x_i)^2\right)$ тоже не превосходит ε^2.

Итак, $\sigma^2 \leq \varepsilon^2$, и потому $\sigma \leq \varepsilon$.

Отсюда вытекает

$$\sigma(\Delta\overline{x}) \;\leq\; \frac{\varepsilon}{\sqrt{n}}\,.$$

Усреднение в статистике

В силу $\mathsf{E}\,(\Delta\overline{x}) = 0$ приходим к выводу, что погрешность $\Delta\overline{x}$ лежит в

$$[-k\,\sigma(\Delta\overline{x}), k\,\sigma(\Delta\overline{x})]$$

для заданного k.

Иными словами, $|\Delta\overline{x}| \leq k\,\sigma(\Delta\overline{x})$.

Вместе с полученной выше оценкой это даёт

$$|\Delta\overline{x}| \;\leq\; \frac{k\,\varepsilon}{\sqrt{n}}\,,$$

т. е. после n измерений стандартное отклонение усреднённого результата и получающийся для него доверительный интервал (при фиксированном уровне значимости) уменьшаются в $\sqrt{n}$ раз.

Зададим уровень доверительной вероятности $\alpha \in \,]0, 1[$.

Пусть на интервале $[-\varepsilon, \varepsilon]$ задана функция плотности $\rho(x)$ некоторого вероятностного распределения с нулевым средним значением, так что

$$\int_{-\infty}^{+\infty} x\,\rho(x)\,\mathrm{d}x = 0.$$

Пусть $\tilde{x}_1$, $\tilde{x}_2$, ..., $\tilde{x}_n$ — независимые и одинаково распределённые случайные величины с плотностью вероятности $\rho(x - x^*)$ для некоторой постоянной x^*. Согласно Определению 1 назовём *вероятностной точностью* усредняющей оценки $\overline{x}$,

$$\overline{x} := \frac{\tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_n}{n},$$

наименьшее положительное число $\delta_{\text{уср}}$, для которого неравенство

$$|\,\overline{x} - x^*| \leq \delta_{\text{уср}}$$

выполняется с вероятностью $\geq \alpha$.

Иными словами, $\delta_{\text{уср}}$ — это наименьшее из вещественных чисел δ, для которых

$$\mathsf{P}\left(\,|\,x - x^*| \le \delta\,\right) \;\ge\; \alpha.$$

Чтобы подчеркнуть зависимость вероятностной точности оценки от количества измерений n будем обозначать её также $\delta_{\text{уср}}(n)$ (для заданного фиксированного α).

Далее мы хотим описать убывание $\delta_{\text{уср}}(n)$ с ростом n, и для этого необходимо напомнить некоторые стандартные понятия, касающиеся сходимости последовательностей.

Говорят, что две последовательности $\{a_n\}$ и $\{b_n\}$ *асимптотически эквивалентны* и пишут $\{a_n\} \sim \{b_n\}$, если

$$\lim_{n\to\infty} a_n/b_n = 1.$$

Будут также полезны «символы Ландау» — O-большое и o-малое, которые широко применяются в математическом анализе для описания асимптотического поведения различных величин.

Обозначение $a_n = O(b_n)$ означает, что существует такая константа $C > 0$, что $|a_n| \leq C\,|b_n|$ для всех n, начиная с некоторого номера.

Обозначение $a_n = o(b_n)$ означает, что $a_n/b_n \to 0$ при неограниченном возрастании номера n.

Теорема

Пусть на интервале $[-\varepsilon, \varepsilon]$ задана функция плотности вероятности $\rho(x)$ с нулевым средним значением, так что

$$\int_{-\infty}^{+\infty} x\,\rho(x)\,\mathrm{d}x = 0.$$

Тогда для любого уровня значимости α существует такая константа C, что

$$\delta_{\text{уср}}(n) \sim \frac{C}{\sqrt{n}}.$$

Этот результат хорошо известен в математической статистике.

См., к примеру,

Ибрагимов И.А., Хасьминский Р.З. *Асимптотическая теория оценивания. – М.: Наука, 1979.*

Интервальные оценки

Выше мы применяли традиционные рассуждения вероятностной статистики, но не совсем типичным образом, поскольку никак не использовалась информация о конкретном виде распределения, которому подчиняются рассматриваемые случайные погрешности.

Было вполне естественно опираться лишь на факт ограниченности погрешностей, т. е. то, что они принадлежат интервалу $[-\varepsilon, \varepsilon]$ и подчинены законам теории вероятностей.

Существует раздел статистической теории — «робастная статистика», в котором рассматриваются методы оценивания в ситуациях, где неизвестны типы распределений или их параметры, и мы знаем лишь класс возможных распределений:

Хьюбер П. *Робастность в статистике. – М.: Мир, 1984.*

Но оценивание параметров для случая, когда нам известны лишь интервалы возможных значений измеряемых величин, развивается в интервальном анализе данных с начала 60-х годов прошлого века.

Начало — пионерская работа Л.В. Канторовича

> Канторович Л.В. *О некоторых новых подходах к вычислительным методам и обработке наблюдений // Сибирский математический журнал. – 1962. – Т. 3, №5. – С. 701–709.*

В некотором смысле интервальный анализ данных можно

поэтому рассматривать как продолжение робастной статистики.

Определение 2.

Пусть задано число $\alpha \in]0,1[$, называемое уровнем значимости (уровнем доверительной вероятности) и пусть $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ — случайные результаты измерений постоянной величины x^* с абсолютной погрешностью ε, так что имеем интервальную выборку $[\tilde{x}_i - \varepsilon, \tilde{x}_i + \varepsilon]$, $i = 1, 2, \ldots, n$, при обработке которой получается интервальная оценка $[x^-, x^+]$ искомой величины x^*.

Назовём *точностью* интервальной оценки $[x^-, x^+]$ такое наименьшее положительное число $\delta_{\text{инт}}$, для которого оба неравенства

$$|x^- - x^*| \le \delta_{\text{инт}} \qquad \text{и} \qquad |x^+ - x^*| \le \delta_{\text{инт}}$$

выполнены с вероятностью $\ge \alpha$. Иными словами, $\delta_{\text{инт}}$ — наименьшее число δ, для которого $\mathrm{P}\left(|x^- - x^*| \le \delta \ \&\ |x^+ - x^*| \le \delta\right) \ge \alpha$.

Обработка накрывающей выборки

Для каждого измерения $\tilde{x}_i$ мы знаем, что его погрешность $\tilde{x}_i - x^*$ принадлежит интервалу $[-\varepsilon, \varepsilon]$.

Поэтому истинное значение x^* лежит в каждом из интервалов $\boldsymbol{x}_i := [\tilde{x}_i - \varepsilon, \tilde{x}_i + \varepsilon]$, $i = 1, 2, \ldots, n$, так что все эти интервалы измерений — накрывающие.

Следовательно, более точную оценку можно получить как пересечение

$$\bigcap_{i=1}^{n} \boldsymbol{x}_i = \bigcap_{i=1}^{n} [\tilde{x}_i - \varepsilon, \tilde{x}_i + \varepsilon].$$

Обработка накрывающей выборки

Если обозначить

$$x^- := \max_{1\le i\le n}\left(\tilde{x}_i - \varepsilon\right) = \max_{1\le i\le n}\tilde{x}_i - \varepsilon,$$

$$x^+ := \min_{1\le i\le n}\left(\tilde{x}_i + \varepsilon\right) = \min_{1\le i\le n}\tilde{x}_i + \varepsilon,$$

то интересующее нас значение x^* лежит в интервале $[x^-, x^+]$.

Операция пересечения интервалов накрывающей выборки настолько естественна и интуитивно понятна, что её происхождение нельзя персонально приписать кому-либо из специалистов, как это сделано в работе Дж.У. Уолстера и В. Крейновича. Это фольклор.

Автор, к примеру, узнал об этом способе обработки интервалов измерений из книги П.Е. Эльясберга (§2.10):

> Эльясберг П.Е. *Измерительная информация: сколько её нужно? как её обрабатывать? – М.: Наука, 1983,*

которая ссылается на более раннюю работу А.И. Шеховцова 1975 года.

И это касается только обработки измерений.

В интервальном анализе идея пересечения результатов для уточнения решения восходит к самым первым работам из 60-х годов XX века.

Какова точность оценки пересечения?

Вычислительные эксперименты, проведённые Дж.У. Уолстером, показывают, что эта оценка часто более точна, чем усреднение.

Н.П. Дываком в работе [1] впервые показано, что при некоторых естественных допущениях стандартное отклонение интервальной оценки убывает как $1/n$ с ростом n, что гораздо лучше точности усреднения, погрешность которого убывает как $1/\sqrt{n}$.

Какие распределения возможны?

Точность интервальной оценки зависит
от неизвестной плотности вероятности погрешности $\rho(x)$.

Единственное, что будем предполагать о ней — ограниченность её носителя, т. е., что она отличается от нуля лишь на интервале $[-\varepsilon, \varepsilon]$.

Кроме того, считаем крайние значения $(-\varepsilon)$ и ε возможными
для погрешности и даже реально достигаемыми ею.

Как можно сформулировать это дополнительное знание
в терминах функции плотности вероятности $\rho(x)$?

Типичный способ определения плотности вероятности из экспериментальных данных таков.

Пусть имеется выборка $\alpha_1, \ldots, \alpha_N$.

Тогда делим интервал $[\min_i \alpha_i, \max_i \alpha_i]$ на несколько подинтервалов и для каждого подинтервала I оцениваем частоту f_I, с которой значения α_i попадают в этот подинтервал:

$$f_I := \frac{N(I)}{N},$$

где $N(I)$ — общее количество тех значений α_i, попавших в I.

Можем определить эмпирическую функцию плотности вероятности $\tilde{\rho}(x)$, которая приближает реальную функцию плотности $\rho(x)$, следующим образом: для $x \in I$ полагаем $\tilde{\rho}(x) = f_I$.

Если зафиксируем конечный набор подинтервалов и применим описанную процедуру к нашему случаю, то в силу того факта, что оба значения $(-\varepsilon)$ и ε возможны, вероятность неравенства $N(I) > 0$ для подинтервалов I, которые содержат $(-\varepsilon)$ или ε стремится к 1 при неограниченном увеличении N.

Следовательно, для достаточно больших N

с вероятностью, близкой к 1,

будем иметь $f_I > 0$ для этих интервалов.

В результате

$$\tilde{\rho}(-\varepsilon) > 0 \qquad \text{и} \qquad \tilde{\rho}(\varepsilon) > 0$$

для эмпирического приближения $\tilde{\rho}(x)$

функции плотности вероятности $\rho(x)$.

Кроме того, описанная выше приближённая функция $\tilde{\rho}(x)$, определённая на интервале $[-\varepsilon, \varepsilon]$, непрерывна в некоторых окрестностях точек $-\varepsilon$ и ε, справа и слева соответственно.

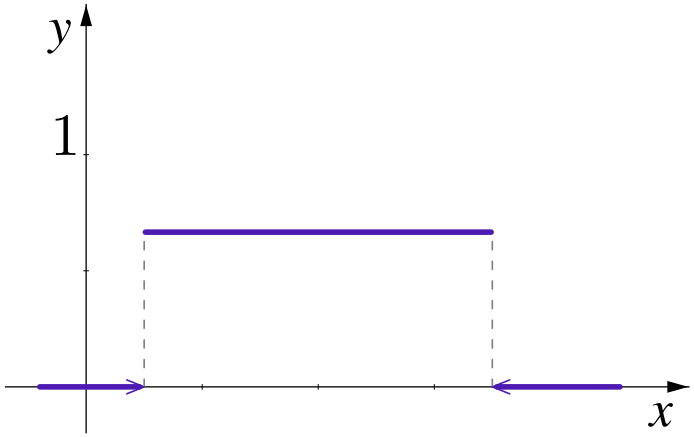
Из сказанного следует, что разумно предполагать те же самые свойства для действительной функции плотности вероятности $\rho(x)$.

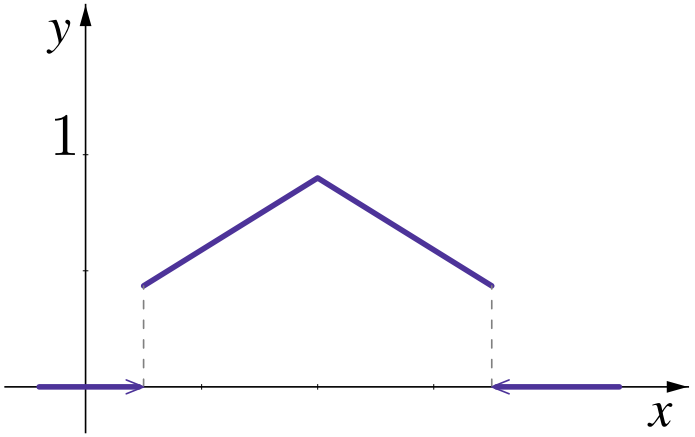
Более точно, будем считать, что

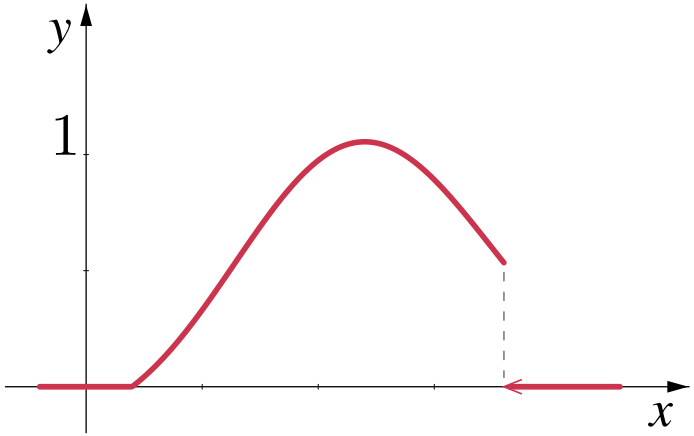
$$\rho(-\varepsilon) > 0 \qquad \text{и} \qquad \rho(\varepsilon) > 0,$$

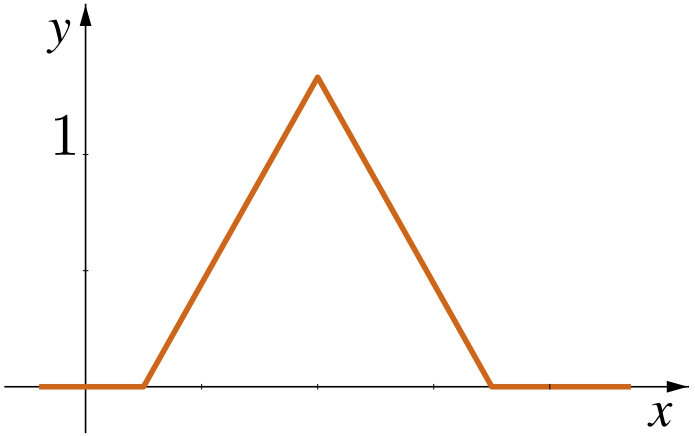
а также что $\rho(x)$ непрерывна в некоторой окрестности точки $(-\varepsilon)$ справа от неё и в некоторой окрестности точки ε слева от неё.

Какие распределения возможны?









Плотности распределений с ограниченным носителем

Какие распределения возможны?

Строго говоря, изложенная выше мотивация допущений на то, что плотность вероятности положительна и непрерывна на концах интервала её носителя, не является совершенно строгой.

Поэтому нельзя заявлять, что все реальные распределения должны удовлетворять этим свойствам.

Какие распределения возможны?

Строго говоря, изложенная выше мотивация допущений на то, что плотность вероятности положительна и непрерывна на концах интервала её носителя, не является совершенно строгой.

Поэтому нельзя заявлять, что все реальные распределения должны удовлетворять этим свойствам.

Но эти условия необходимы

для доказательство теоремы Дывака-Крейновича.

По этой причине мы всё-таки утверждаем, что существует разумный класс распределений, для которых наши допущения выполнены и для которых интервальные оценки обладают лучшим свойствами.

С математической точки зрения сформулированные ограничения означают, что из всех возможных значений $\rho(-\varepsilon)$ и $\rho(\varepsilon)$, т. е. из всех неотрицательных чисел, мы исключаем 0 с близкими значениями.

Основной результат

Теорема Дывака–Крейновича

Пусть $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ — независимые и одинаково распределённые случайные величины с плотностью вероятности $\rho(x - x^*)$ для некоторой постоянной x^*, а функция плотности вероятности $\rho(x)$ удовлетворяет следующим условиям:

(i) $\rho(x) \geq 0$ лишь на интервале $[-\varepsilon, \varepsilon]$, а вне его равна нулю;

(ii) $\rho(\varepsilon) > 0$ и $\rho(-\varepsilon) > 0$,

(iii) $\rho(x)$ является непрерывной в некоторой окрестности точки $(-\varepsilon)$ справа от неё и в некоторой окрестности точки ε слева от неё.

Тогда точность интервальной оценки $\delta_{\text{инт}}(n) = O(1/n)$.

Точность интервальной оценки убывает не медленнее C/n, что с ростом n гораздо лучше точности усредняющей оценки, убывание которого эквивалентно всего лишь $1/\sqrt{n}$.

Кроме того, усреднение применимо лишь к случаю $\mathsf{E}\,(\Delta x) = 0$, тогда как интервальная оценка работает всегда.

Точность интервальной оценки убывает не медленнее C/n, что с ростом n гораздо лучше точности усредняющей оценки, убывание которого эквивалентно всего лишь $1/\sqrt{n}$.

Кроме того, усреднение применимо лишь к случаю $\mathsf{E}\,(\Delta x) = 0$, тогда как интервальная оценка работает всегда.

Доказательство результата несложно, хотя и не совсем тривиально. Он находится в хорошем согласии с известными результатами.

Известно, например, что оценка

$$\frac{1}{2}\left(\min_i \tilde{x}_i + \max_i \tilde{x}_i\right)$$

для равномерных распределений на интервале $[-\varepsilon, \varepsilon]$ является оптимальной в некотором разумном смысле. Более точно, она является *минимаксной оценкой*.

Какие условия теоремы важны?

Если условия $\rho(-\varepsilon) > 0$ или $\rho(\varepsilon) > 0$ не выполнены, то заключение теоремы может оказаться неверным.

В самом деле, если, к примеру, $\rho(x)$ — равномерное распределение на интервале $[-\varepsilon/2, \varepsilon/2]$, то

$$-\tfrac{1}{2}\,\varepsilon \le \tilde{x}_i \le \tfrac{1}{2}\,\varepsilon,$$

и потому

$$x^- = \max_i \tilde{x}_i - \varepsilon \le \tfrac{1}{2}\,\varepsilon - \varepsilon = -\tfrac{1}{2}\,\varepsilon,$$

$$x^+ = \min_i \tilde{x}_i + \varepsilon \ge -\tfrac{1}{2}\,\varepsilon + \varepsilon = \tfrac{1}{2}\,\varepsilon.$$

Тогда для всех n точность интервальной оценки оценивается снизу как $\delta_{\text{инт}}(n) \ge \frac{1}{2}\,\varepsilon$, и она не стремится к нулю.

Какие условия теоремы важны?

При нарушении других условий теоремы, например, когда $\rho(\pm\varepsilon) = 0$, но производные $\rho'(\pm\varepsilon)$ отличны от нуля, сходимость $\delta_{\text{инт}}(n)$ к нулю имеет место, хотя и существенно более медленная, чем утверждается теоремой.

Но имеются также примеры противоположного свойства.

Тот факт, что для точности $\delta_{\text{инт}}(n)$ получена лишь верхняя оценка, не случаен.

В некоторых случаях точность интервального оценивания может быть гораздо более высокой.

Когда оценка точности ещё выше

Рассмотрим, к примеру, ситуацию, когда погрешность принимает с вероятностью 0.5 одно из двух значений $-\varepsilon$ и ε.

Если мы измеряем величину x несколько раз, то можем получить два возможных результата $x - \varepsilon$ и $x + \varepsilon$.

При независимости результатов отдельных измерений вероятность того, что n последовательных результатов равны $x + \varepsilon$, есть 2^{-n}.

Аналогично, вероятность того, что n последовательных результатов равны $x - \varepsilon$, есть 2^{-n}.

Когда оценка точности ещё выше

Во всех остальных случаях, имеющих суммарную вероятность $1 - 2 \cdot 2^{-n}$, имеются как минимум два различных результата измерений.

При этом один из них равен $x + \varepsilon$, а другой — $x - \varepsilon$.

По этой причине $x^- = \max \tilde{x}_i - \varepsilon = x + \varepsilon - \varepsilon = x$,

и аналогично $x^+ = x = x^-$.

Таким образом, в этом случае интервальная оценка

абсолютно точна, т. е. $\delta_{\text{инт}}(n) = 0$.

Когда оценка точности ещё выше

Это происходит, когда

$$1 - 2 \cdot 2^{-n} = 1 - 2^{-(n-1)} \geq \alpha,$$

т. е. когда

$$2^{-(n-1)} \leq 1 - \alpha,$$

так что

$$n \;\geq\; 1 - \log_2(1-\alpha) \;=\; \log_2\left(\frac{2}{1-\alpha}\right).$$

Получающиеся из этой оценки значения n вполне разумны.

Например, для $\alpha = 95\%$ имеем $n \geq 6$;

для $\alpha = 99.9\%$ имеем $n \geq 11$ и т. д.

Доказательство теоремы Дывака-Крейновича

Чтобы оценить $\delta_{\text{инт}}(n)$,

оценим сначала для заданного δ вероятность P того,

что $|x^- - x^*| \le \delta$ и $|x^+ - x^*| \le \delta$.

Эту вероятность более удобно оценивать через вероятность p противоположного события, как

$$P = 1 - p,$$

где p — вероятность того, что $|x^- - x^*| > \delta$ или $|x^+ - x^*| > \delta$.

Поскольку для любых событий A и B справедливо

$$\mathsf{P}(A \vee B) \le \mathsf{P}(A) + \mathsf{P}(B),$$

то можем оценить p как $p \le p^- + p^+$, где

$$p^- := \mathsf{P}(|x^- - x^*| > \delta) \qquad \text{и} \qquad p^+ := \mathsf{P}(|x^+ - x^*| > \delta).$$

Следовательно, если найдём δ, для которого

$$p^- \leq \tfrac{1}{2}(1-\alpha) \qquad \text{и} \qquad p^+ \leq \tfrac{1}{2}(1-\alpha),$$

то получим

$$p \leq p^- + p^+ \leq 1-\alpha,$$

и в целом будем иметь

$$P = 1 - p \geq \alpha.$$

Для найденного δ выполнится неравенство $\delta_{\text{инт}}(n) \leq \delta$.

Оценим сначала величину p^-.

Так как

$$\tilde{x}_i \le x^* + \varepsilon \qquad \text{для всех} \quad i,$$

то можем заключить, что $\max \tilde{x}_i \le x^* + \varepsilon$ и потому

$$x^- = \max \tilde{x}_i - \varepsilon \le x^*.$$

Следовательно, разность $x^- - x^*$ всегда неположительна, и неравенство $|x^- - x^*| \ge \delta$ равносильно $x^- - x^* \le -\delta$, т. е.

$$x^- \ge x^* - \delta.$$

По этой причине

$$p^- = \mathsf{P}\,(x^- < x^* - \delta).$$

Коль скоро $x^- = \max \tilde{x}_i - \varepsilon$, условие $x^- < x^* - \delta$

равносильно $\max \tilde{x}_i < x^* + \varepsilon - \delta$.

Наибольшее из n чисел $\tilde{x}_i$ меньше, чем $x^* + \varepsilon - \delta$ тогда и только тогда, когда все эти n чисел меньше его, и поэтому

$$p^- = \mathsf{P}\left((\tilde{x}_1 < x^* + \varepsilon - \delta)\ \&\ \ldots\ \&\ (\tilde{x}_n < x^* + \varepsilon - \delta)\right).$$

Поскольку переменные $\tilde{x}_i$ независимы друг от друга, то получаем

$$p^- = \mathsf{P}\left(\tilde{x}_1 < x^* + \varepsilon - \delta\right) \cdot \ldots \cdot \mathsf{P}\left(\tilde{x}_n < x^* + \varepsilon - \delta\right).$$

Наконец, $\tilde{x}_i$ одинаково распределены,

и поэтому $p^- = Q^n$, где обозначено

$$Q := \mathsf{P}\left(\tilde{x}_1 < x^* + \varepsilon - \delta\right).$$

Таким образом, для оценивания p^- необходимо оценить значение Q.

Как и ранее, удобно вычислять вероятность

$$q := \mathsf{P}\left(\tilde{x}_1 \geq x^* + \varepsilon - \delta\right)$$

противоположного события, так что $Q = 1 - q$.

Отметим, что

$$q = \mathsf{P}\left(\tilde{x}_1 \geq x^* + \varepsilon - \delta\right) = \mathsf{P}\left(x^* + \varepsilon - \delta \leq \tilde{x}_1 \leq x^* + \varepsilon\right),$$

так как оценка для $\tilde{x}_1$ сверху получена ранее.

Эту вероятность можно выразить через функцию плотности как

$$q = \int_{\varepsilon-\delta}^{\varepsilon} \rho(x)\ \mathrm{d}x \ = \ \delta\,\rho(\varepsilon) + \Delta q,$$

где $\Delta q = o(\delta)$ в силу свойств интеграла.

Тот факт, что $\Delta q = o(\delta)$ означает, что

$$\Delta q/\delta \to 0 \qquad \text{при } \delta \to 0,$$

т. е. для всякого $k > 0$ найдётся такое δ_0^-, что для всех $\delta \le \delta_0^-$ будем иметь $|\Delta q| \le k\delta$.

В частности, если возьмём $k = \rho(\varepsilon)/2$, можем заключить, что для всех $\delta \le \delta_0^-$ выполняется

$$q \ge \rho(\varepsilon)\,\delta - k\delta = \delta\rho(\varepsilon)/2.$$

Как следствие,

$$Q \;=\; 1 - q \;\le\; 1 - \delta\rho(\varepsilon)/2, \qquad (\star)$$

и можно использовать это неравенство для Q с целью оценивания p^-:

$$p^- = Q^n \le \big(1 - \delta\rho(\varepsilon)/2\big)^n.$$

Если хотим выбрать δ так, чтобы выполнялось $p^- \leq \frac{1}{2}(1-\alpha)$, то должно быть

$$\left(1 - \frac{\delta\rho(\varepsilon)}{2}\right)^n \leq \tfrac{1}{2}(1-\alpha).$$

Выполняя эквивалентные преобразования с этим неравенством, получим

$$\frac{\delta\rho(\varepsilon)}{2} \geq 1 - \left(\frac{1-\alpha}{2}\right)^{1/n},$$

и, наконец, $\delta \geq \delta_n^-$, где обозначено

$$\delta_n^- := \frac{2}{\rho(\varepsilon)}\left(1 - \left(\frac{1-\alpha}{2}\right)^{1/n}\right).$$

В целом, если $\delta_n^- \leq \delta \leq \delta_0^-$, то можем гарантировать, что

$$p^- \leq \tfrac{1}{2}(1-\alpha).$$

Оценим теперь асимптотическое поведение δ_n^- при $n \to \infty$.

Имеем

$$\left(\frac{1-\alpha}{2}\right)^{1/n} = \left(\exp\left(\ln\frac{1-\alpha}{2}\right)\right)^{1/n} = \exp\left(\frac{1}{n}\left(\ln\frac{1-\alpha}{2}\right)\right)$$

Выражение в аргументе экспоненты стремится к нулю при $n \to \infty$, так что применима приближённая формула $\exp(x) = 1 + x + o(x)$. Тогда

$$\left(\frac{1-\alpha}{2}\right)^{1/n} = 1 + \frac{1}{n}\ln\frac{1-\alpha}{2} + o\left(\frac{1}{n}\right),$$

откуда следует

$$1 - \left(\frac{1-\alpha}{2}\right)^{1/n} = -\frac{1}{n}\ln\frac{1-\alpha}{2} + o\left(\frac{1}{n}\right),$$

и

$$\delta_n^- = \frac{2}{\rho(\varepsilon)}\left(1 - \left(\frac{1-\alpha}{2}\right)^{1/n}\right) = -\frac{2}{\rho(\varepsilon)}\frac{1}{n}\ln\frac{1-\alpha}{2} + o\left(\frac{1}{n}\right).$$

Если обозначить

$$A^- := -\frac{2}{\rho(\varepsilon)} \ln \frac{1-\alpha}{2},$$

то получаем

$$\delta_n^- = \frac{A^-}{n} + o\left(\frac{1}{n}\right).$$

Совершенно аналогичным образом можно определить δ_n^+ и δ_0^+, для которых при $\delta_n^+ < \delta < \delta_0^+$ выполняется неравенство $p^+ \leq (1-\alpha)/2$ и доказать, что

$$\delta_n^+ = \frac{A^+}{n} + o\left(\frac{1}{n}\right)$$

для некоторой константы A^+.

Теперь мы готовы оценить $\delta_{\text{инт}}(n)$.

Если взять $\delta_n := \max\{\delta_n^-, \delta_n^+\}$, то

$$\delta_n = \frac{A}{n} + o\left(\frac{1}{n}\right),$$

где обозначено $A := \max\{A^-, A^+\}$.

При $n \to \infty$ выписанное выражение для δ_n стремится к нулю,
так что для достаточно больших n справедливо

$$\delta_n \leq \min\{\delta_0^-, \delta_0^+\}.$$

Следовательно, для таких n, взяв $\delta = \delta_n$, получаем

$$p^- \leq \tfrac{1}{2}(1-\alpha) \qquad \text{и} \qquad p^+ \leq \tfrac{1}{2}(1-\alpha),$$

так что в целом $P \geq \alpha$.

Поскольку мы определили $\delta_{\text{инт}}(n)$ как наименьшее число δ, удовлетворяющее этому условию, то можем заключить, что

$$\delta_{\text{инт}}(n) \leq \delta_n = \frac{A}{n} + o\left(\frac{1}{n}\right)$$

и

$$\delta_{\text{инт}}(n) = O\left(\frac{1}{n}\right).$$

Это и требовалось доказать.

Комментарий к доказательству

В нём, начиная с оценки ($\star$), используется тот факт, что $\rho(\varepsilon) > 0$.

Если $\rho(\varepsilon) = 0$, то дальнейшие построения невозможны
или их нужно сильно корректировать.

Из доказательства также следует увеличение точности интервальной оценки пересечения при росте значений $\rho(\varepsilon)$ и $\rho(-\varepsilon)$, т. е. вероятности наблюдения погрешностей, близких к концам интервала локализации.

Этот вывод находится в хорошем согласии с результатами вычислительных экспериментов из работы

> Walster G.W. *Philosophy and practicalities of interval arithmetic / Reliability in Computing. The Role of Interval Methods in Scientific Computing. Edited by: Ramon E. Moore. – Boston - San Diego - New York: Academic Press, 1988. – P. 309–323.*

Спасибо за внимание!