

Оценивание параметров непрерывных распределений на основе метрологических характеристик

А. В. Степанов

ВНИИМ им. Д. И. Менделеева

stepanov17@yandex.ru

Аннотация. В метрологической практике нередко возникают ситуации, когда требуется построить модель случайной величины при отсутствии исходной выборки. Согласно Дополнению 1 к Руководству по выражению неопределенности измерения (GUM), решение задачи трансформирования распределений методом Монте-Карло требует задания законов распределения для всех входных величин; сгенерированные на их базе значения служат для последующего формирования распределения выходной величины. Согласно указанному Дополнению, входные распределения могут приписываться на основании любой доступной информации (среднего значения, стандартной неопределенности, квантилей, гистограмм или экспертных оценок). В работе обосновывается использование метода наименьших квадратов для оценивания параметров непрерывных распределений по известным метрологическим характеристикам: измеренному значению, стандартной неопределенности и границам интервала охвата, в условиях отсутствия исходных выборочных данных.

Ключевые слова: случайная величина, плотность вероятности, неопределенность, интервал охвата, оценивание параметров, метод наименьших квадратов, принцип максимума энтропии, численная оптимизация

1. ВВЕДЕНИЕ

В метрологических приложениях часто возникают ситуации, когда необходимо построить модель случайной величины на основе ограниченной информации, не включающей исходную выборку (что исключает возможность применения классических методов оценивания, таких как метод моментов или метод максимального правдоподобия). Например, в рамках Дополнения 1 к Руководству по выражению неопределенности измерения [1], при решении задачи трансформирования распределений методом Монте-Карло требуется задать распределения для всех входных величин модели измерения, для генерации входных данных и последующего получения распределения выходной величины модели. Согласно Дополнению 1, входные распределения могут быть установлены на основе любой доступной информации, в частности, среднего значения, стандартной неопределенности, квантилей, гистограмм или экспертных оценок. Когда входная величина не может быть адекватно описана общепринятым нормальным распределением (например, из-за существенной асимметрии или несоответствия бесконечного носителя нормального закона реальным физическим границам величины), представляется целесообразным использовать более гибкие параметрические семейства. Для подбора параметров этих распределений по заданным характеристикам в

данной работе предлагается использовать метод наименьших квадратов (МНК).

Исторически применение МНК для оценивания параметров измерений развивалось преимущественно в рамках анализа порядковых статистик [2, 3]. Метод квантильных наименьших квадратов традиционно применяется для подгонки теоретических функций распределения под выборочные данные [4]. При этом упомянутые работы, как правило, предполагают наличие исходной выборки и рассчитаны на использование однородной статистической информации (выборочные моменты или порядковые статистики). Предлагаемый подход предполагает объединение разнородной информации в условиях отсутствия исходных выборочных данных, позволяя не только подобрать параметры модели, но и количественно оценить согласованность предоставленных метрологических характеристик. Заметим, что в условиях возможной несогласованности или избыточности исходных данных МНК позволяет получить устойчивую аппроксимацию, сглаживающую их противоречия, в то время как классический принцип максимума энтропии потребовал бы строгого удовлетворения всем условиям, что привело бы к артефактам (потерям гладкости) в искомой функции плотности.

Далее будет дано описание соответствующего математического аппарата, включая выбор весов.

II. ПОДБОР ПАРАМЕТРОВ С ПОМОЩЬЮ МНК

Предположим, что имеется непрерывное распределение случайной величины X , принадлежащее некоторому параметрическому семейству $f(x, \theta)$ (далее также будем обозначать его $f(\theta)$), где θ – вектор параметров. Требуется оценить параметры, определяющие выбор конкретного распределения из данного семейства, основываясь на доступной информации о нем. В метрологических приложениях, где плотность f описывает информацию о вероятностной модели измеряемой величины X , в качестве такой информации могут выступать: измеренное значение величины, интерпретируемое как математическое ожидание $\mu = E X$; стандартная неопределенность u , соответствующая стандартному отклонению (СКО) $u = \sqrt{\text{Var}(X)}$; границы интервала охвата $c_{1,2}$ для заданного уровня вероятности P_0 . Заметим, что под интервалом охвата будем в дальнейшем понимать вероятностно-симметричный двусторонний интервал охвата $c_{1,2} = F^{-1}((1 \mp P_0)/2)$ (где F^{-1} – квантильная

функция), хотя, очевидно, ход рассуждений останется прежним и при рассмотрении произвольных интервалов охвата (например, односторонних).

Предоставленная информация (значения $\mu, u, c_{1,2}$) может быть использована для получения вектора оценок $\hat{\theta}$ параметров рассматриваемого семейства на основе взвешенного МНК, т. е. $\hat{\theta}$ представляет собой решение оптимизационной задачи:

$$\Phi(\theta) = w_{\mu}(\mu_{\theta} - \mu)^2 + w_u(u_{\theta} - u)^2 + w_c \left[(c_{1,\theta} - c_1)^2 + (c_{2,\theta} - c_2)^2 \right] \rightarrow \min_{\theta}$$

где веса w_{μ}, w_u, w_c предназначены для балансировки вкладов слагаемых Φ , а величины с индексом θ представляют собой «теоретические» значения рассматриваемых характеристик, как функций параметров θ . При подборе параметров симметричного распределения вместо последнего слагаемого, отвечающего границам интервала охвата $c_{1,2}$, можно рассмотреть слагаемое вида $w_K(K_{\theta} - K)^2$, где K – коэффициент охвата (для заданного уровня вероятности P_0). Функционал Φ может также включать оценки границ носителя распределения (если носитель является ограниченным)

В случае, когда число оцениваемых параметров равно числу наложенных ограничений, а минимум целевой функции может достигать нуля, МНК сводится, по сути, к численному решению системы нелинейных уравнений

$$\mu_{\theta} = \mu, \quad u_{\theta} = u, \quad c_{1,\theta} = c_1, \quad c_{2,\theta} = c_2$$

или, в случае симметричного распределения,

$$\mu_{\theta} = \mu, \quad u_{\theta} = u, \quad K_{\theta} = K.$$

Вопрос о решении такой системы для оценивания параметров двустороннего степенного распределения [5] рассматривался, например, в работе [6] (приведена программная реализация). В то же время применение МНК выглядит предпочтительным, поскольку предоставляемая информация (метрологические характеристики) может содержать неточности (как минимум, погрешности округления), приводящие к несовместности рассматриваемой системы уравнений.

Как отмечалось ранее, МНК является традиционным инструментом обработки избыточных данных, например, при необходимости согласования оценок одной и той же величины. В этом случае все оценки могут быть введены в целевой функционал Φ с различными весами (в случае их значительного расхождения следует убедиться в их статистической согласованности). Аналогично, могут использоваться границы интервала охвата для двух уровней вероятности, условно говоря, «стандартного» (например, 95%) и «критического» (например, 99.9%), и т. д.

Если имеется дополнительная информация о величине X , представленная в виде гистограммы (или нескольких), то ее также можно включить в итоговый функционал, например,

$$\Phi(\theta) = w_{\mu}(\mu_{\theta} - \mu)^2 + w_u(u_{\theta} - u)^2 + \dots + w_h \sum_{i=1}^{n_b} \frac{(n_i - E_i(\theta))^2}{E_i(\theta)},$$

где n_b – число бинов (корзин) в гистограмме, n_i – наблюдаемая частота в i -м бине; ожидаемое (теоретическое) количество попаданий случайной величины в i -й бин описывается формулой

$$E_i(\theta) = \left(\sum n_j \right) (F(b_i, \theta) - F(b_{i-1}, \theta)),$$

b_i – границы бинов; F – функция распределения семейства. В этом случае функционал Φ отображает в том числе и то, насколько подбираемая теоретическая плотность способна описать рельеф гистограммы. Включение последнего слагаемого в целевой функционал не меняет базовой логики метода, предназначенного для работы в условиях отсутствия выборочных данных. Здесь гистограмма – это не набор реализаций случайной величины, а агрегированная форма представления информации (взятая, например, из архивных отчетов), которая характеризуется дискретностью и потерей точности. Тем не менее, даже такие грубые данные позволяют учитывать характерные черты распределения (мультимодальность, существенная асимметрия, специфическая форма «хвостов»), которые не могут быть описаны через интегральные характеристики. При этом следует помнить, что случайные флуктуации частот вследствие недостаточного числа отсчетов могут заметно влиять на итоговую форму плотности вероятности. Поэтому вес w_h , не стоит брать слишком большим, а гистограмму, стоит рассматривать скорее как вспомогательный инструмент для уточнения формы распределения.

Значение суммарной остаточной невязки $\Phi(\hat{\theta})$ традиционно служит индикатором согласованности исходной измерительной информации с выбранной вероятностной моделью. Если оно велико, то это свидетельствует о внутреннем противоречии предоставленных характеристик либо об их несоответствии выбранной вероятностной модели. Способность величины $\Phi(\hat{\theta})$ выступать критерием качества аппроксимации может быть использована для выбора (в том числе, автоматизированного) семейства распределений из набора семейств-кандидатов. Сравнивая значения $\Phi_j(\hat{\theta}_j)$, полученные для различных семейств распределений $f_j(x, \theta)$, можно выбрать из них то, которое обеспечивает наименьшую невязку. Т. е. МНК здесь – не только инструмент оценивания параметров, но и критерий выбора модели, наиболее полно и непротиворечиво учитывающей доступную измерительную информацию.

III. ВЫБОР ВЕСОВЫХ КОЭФФИЦИЕНТОВ

Ключевым моментом при использовании МНК является обоснованный выбор весов w_{μ}, w_u, w_c, \dots , определяющих вклад невязок в целевой функционал. В простейшем случае можно считать их равными единице. Однако на практике невязки могут иметь различный порядок величин, а исходные характеристики –

неодинаковую степень достоверности и влияния на форму распределения. Например, оценки границ интервала охвата $c_{1,2}$ (или коэффициент охвата K) зачастую являются консервативными и округляются в большую сторону (что может приводить к их несовместимости с заданным значением стандартной неопределенности u для конкретной модели). Использование единичных весов также может привести к доминированию слагаемых с наибольшими абсолютными значениями (например, μ), что исказит итоговые оценки. Близость остаточной невязки к нулю при этом будет создавать иллюзию адекватности модели. Для обеспечения соизмеримости вкладов разных слагаемых можно, например, предложить использовать веса, обратно пропорциональные квадратам целевых значений параметров ($w_\mu = \mu^{-2}$, $w_u = u^{-2}$, ...), переходя к минимизации относительных невязок. Если известны неопределенности целевых характеристик, целесообразно выбрать веса, обратно пропорциональные их квадратам. При таком подходе (наиболее естественном в метрологической практике) вклад каждой характеристики в целевой функционал пропорционален ее точности. Если значение суммарной остаточной невязки при этом существенно превышает число избыточных условий, то это свидетельствует о несовместимости данных с выбранной моделью на заданном уровне точности.

В качестве альтернативы можно использовать приоритетное взвешивание, когда компоненты измерительной информации имеют разную достоверность (например, когда результату прямого измерения μ назначается больший вес, чем экспертным оценкам границ $c_{1,2}$). Другой вариант – нормировка на стандартную неопределенность (выбор весов, пропорциональных u^{-2}), которая, например, актуальна, когда значение μ близко к нулю, и использование относительного веса $w_\mu = \mu^{-2}$ некорректно.

Обоснованность выбора весов может подтверждаться проверкой качества подгонки на контрольных данных или методом кросс-валидации.

IV. НЕОПРЕДЕЛЕННОСТЬ ПОЛУЧЕННЫХ ОЦЕНОК

В соответствии с Руководством [7], для моделей, заданных неявно или через процедуры оптимизации, для оценивания неопределенности $\hat{\theta}$ следует применять стандартный аппарат ковариационного анализа. В рамках данного подхода ковариационная матрица оценок вычисляется по формуле

$$V(\hat{\theta}) = \hat{\sigma}^2 (J^T W J)^{-1},$$

где J – матрица Якоби (чувствительности), элементы которой представляют собой частные производные теоретических характеристик по параметрам θ ; W – диагональная матрица весов; $\hat{\sigma}^2 = \Phi(\hat{\theta})/d$, где d – число избыточных условий. При $d=0$ принимаем $\hat{\sigma}^2 = 1$, считая, что масштаб неопределенности уже учтен в весовой матрице.

Ввиду возможной существенной нелинейности квантильных функций или близости оценок к границам

области определения, стандартный линейризованный подход может давать смещенные результаты. В таких ситуациях неопределенность параметров целесообразно оценивать методом Монте-Карло в соответствии с [1]. Многократно решая задачу оптимизации для различных реализаций входных характеристик, можно получить эмпирические распределения оценок $\hat{\theta}$ и рассчитать их стандартные неопределенности, с учетом специфики модели.

V. ПРИМЕРЫ, ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

По ссылке [8] доступны примеры программной реализации МНК для оценивания параметров Бета-распределения, распределения Кумарасвами [9] и двустороннего степенного распределения [5] на языке Python.

В случае Бета-распределения, определенного на носителе $[a, b]$, $\theta = (a, b, \alpha, \beta)$, и

$$\mu_\theta = a + (b-a) \frac{\alpha}{\alpha + \beta}, \quad u_\theta = (b-a) \sqrt{\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}},$$

$$c_{1,\theta} = F^{-1}((1 - P_0)/2), \quad c_{2,\theta} = F^{-1}((1 + P_0)/2),$$

где $F^{-1}(p) = a + (b-a) Q(p, \alpha, \beta)$, Q – квантильная функция стандартного Бета-распределения. Опустим описание теоретических характеристик для остальных моделей, во избежание перегрузки текста второстепенными техническими деталями (их вид доступен в литературе). Заметим, что выражения для границ интервалов охвата для двустороннего степенного распределения и распределения Кумарасвами могут быть выражены в элементарных функциях (для двустороннего степенного распределения этот факт отражен в работе [6]), что облегчает решение оптимизационной задачи. Двустороннее степенное распределение позволяет естественным образом включить моду (наиболее вероятное значение) в целевой функционал.

Метод не зависит от типа распределения: расчетная процедура опирается на стандартные библиотеки оптимизации и остается неизменной для любого семейства. Такая алгоритмическая универсальность позволяет легко дополнять программный код при расширении набора типовых моделей-кандидатов.

VI. ЭНТРОПИЙНАЯ РЕГУЛЯРИЗАЦИЯ

Возвращаясь к упомянутому во введении принципу максимума энтропии (ПМЭ), отметим, что в метрологической практике он рассматривается как фундаментальный способ построения распределений при дефиците априорной информации. Плотность, соответствующая максимуму энтропии, является экспоненциальной [10] (при заданном математическом ожидании), при использовании дополнительной информации об интервалах охвата она становится кусочно-экспоненциальной [11]. В то время как ПМЭ формирует плотность с возможной потерей гладкости на границах интервалов охвата, МНК обеспечивает подбор параметров в рамках семейств с заданными свойствами (как аналитически гладких, так и моделей, негладкость которых продиктована их априорной структурой). Вопрос о предпочтительности той или иной формы является дискуссионным для конкретных задач, однако

использование гладких плотностей может лучше соответствовать непрерывному характеру физических величин и упрощать дальнейшие аналитические расчеты. Заметим повторно, что ПМЭ чувствителен к избыточности данных: даже при незначительной их рассогласованности исходная задача поиска плотности может не иметь решения.

Совмещение данных подходов возможно путем введения в функционал Φ дополнительного энтропийного слагаемого вида $-w_H H(f(\theta))$, $w_H \geq 0$, штрафующего модель за низкую информационную энтропию (при малых w_H приоритет отдается точности соответствия исходным характеристикам, а при больших – увеличению информационной неопределенности модели). Выбор компромиссного значения w_H может, например, осуществляться путем максимизации энтропии до тех пор, пока суммарная невязка МНК остается сопоставимой с погрешностями исходных характеристик. Данный подход помогает избежать неоправданно оптимистичных оценок точности и делает модель более устойчивой к незначительным погрешностям входных характеристик.

VII. ЗАКЛЮЧЕНИЕ

В работе предложен метод оценивания параметров вероятностных моделей на основе МНК в условиях отсутствия исходных выборочных данных. Подход не претендует на принципиальную новизну, тем не менее, в рассматриваемом контексте он доведен до уровня алгоритма, готового к непосредственному применению в метрологической практике. В отличие от подходов, ориентированных на порядковые статистики, данный алгоритм интегрирует разнородную информацию: моменты, интервалы охвата (и, при наличии, гистограммы). Использование взвешенного функционала обеспечивает корректный подбор параметров для широкого круга семейств, позволяя также автоматизировать выбор модели, что особенно важно при реализации процедур Монте-Карло согласно

Дополнению 1 к Руководству по выражению неопределенности [1]. При этом величина остаточной невязки метода служит индикатором непротиворечивости данных, позволяя выявлять скрытые ошибки в бюджете неопределенности. Для получения оценок точности параметров предлагается использовать стандартный аппарат ковариационного анализа (дополненный, в случае необходимости, методом Монте-Карло). В сравнении с ПМЭ, подход на основе МНК позволяет получать аналитически удобные гладкие плотности, что упрощает их дальнейшее использование.

СПИСОК ЛИТЕРАТУРЫ

- [1] CGM 101:2008. Evaluation of measurement data — Supplement 1 to the «Guide to the expression of uncertainty in measurement» — Propagation of distributions using a Monte Carlo method. BIPM, 2008. 90 p.
- [2] Lloyd E. H. Least-squares estimation of location and scale parameters using order statistics // *Biometrika*. 1952. Vol. 39, № 1/2. P. 88–95.
- [3] Contributions to Order Statistics / ed. by A.E. Sarhan, B.G. Greenberg. New York: John Wiley & Sons, 1962. 482 p.
- [4] Cox N.J. Quantile least squares as a method of fitting distributions // *The Stata Journal*. 2006. Vol. 6, № 4. P. 597–598.
- [5] Kotz S., Van Dorp J.R. Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications. Singapore: World Scientific Publishing, 2004. 308 p.
- [6] Степанов А.В., Чуновкина А.Г. Применение семейства двусторонних степенных распределений для обработки результатов измерений // *Измерительная техника*. 2027. Т. 75, № 1. С. 34–44.
- [7] JCGM 102:2008. Evaluation of measurement data — Guide to the expression of uncertainty in measurement. BIPM, 2008. 120 p.
- [8] Примеры, программная реализация: <https://github.com/stepanov17/scm2026>
- [9] Jones M.C. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages // *Statistical Methodology*. 2009. Vol. 6, № 1. P. 70–81.
- [10] Kullback S. Information Theory and Statistics. New York: John Wiley & Sons, 1959. 395 p.
- [11] Kapur J.N., Kesavan H.K. Entropy Optimization Principles with Applications. Academic Press, 1992. 408 p.