

ОБ ОДНОЙ РЕАЛИЗАЦИИ МЕТОДА МОНТЕ-КАРЛО ДЛЯ ТРАНСФОРМИРОВАНИЯ TSP-РАСПРЕДЕЛЕНИЙ ПРИ АНАЛИТИЧЕСКИ ЗАДАННОЙ МОДЕЛИ ИЗМЕРЕНИЯ

А. В. Степанов^{1, *}

¹ Всероссийский научно-исследовательский институт метрологии
им. Д. И. Менделеева, Санкт-Петербург, Россия

Разработана программная реализация метода Монте-Карло для трансформирования распределений при аналитически заданной модели измерения и входных величинах, имеющих нормальное или TSP-распределение.

A software implementation of the Monte Carlo method has been developed for propagating distributions for an analytically specified measurement model and input quantities that have a normal or TSP distribution.

PACS: 06.20.Dk; 02.50.Ng; 07.05.Tr

ВВЕДЕНИЕ

Метод Монте-Карло широко применяется в метрологии [1, 2], особенно при решении задач трансформирования распределений для оценивания неопределенности измерений. Согласно приложению 1 к руководству по выражению неопределенности измерений (GUM) [3] данный метод является рекомендуемым инструментом для анализа моделей измерений, заданных математическими зависимостями. Он позволяет корректно оценивать неопределенность косвенных измерений, не требуя линейаризации модели и адекватно учитывая нестандартный вид распределений вероятности входных величин (например, их асимметрию). В отличие от классических подходов метод не требует априорных предположений о типе распределения выходной величины и позволяет оценивать полный набор статистических характеристик, а не только параметры рассеяния.

В данной работе представлена программная реализация [4] метода распространения неопределенности на языке программирования Python, включающая: использование семейств двусторонних степенных (Two-Sided Power, TSP) и нормальных распределений для описания входных

* E-mail: stepanov17@yandex.ru

величин, алгоритмы выбора параметров этих распределений на основе экспериментальных данных, механизм задания модели измерения с помощью аналитических выражений и комплексный расчет статистических показателей выходной величины. Также рассматривается вопрос распараллеливания выполняемых вычислений.

1. ВХОДНЫЕ ВЕЛИЧИНЫ; ИСПОЛЬЗОВАНИЕ TSP-РАСПРЕДЕЛЕНИЙ

Суть распространения неопределенности с использованием метода Монте-Карло согласно [3] заключается в трансформировании распределений посредством модели измерений $Y = F(X_1, X_2, \dots, X_n)$. Процедура включает выборку значений из распределений, приписанных входным величинам X_i , вычисление значений Y для каждого набора значений входных величин, построение распределения Y . Для описания распределений входных величин модели помимо семейства нормальных распределений предлагается использовать семейство TSP-распределений [5], сочетающее широкое разнообразие форм плотности с крайней простотой математического описания. В частности, для него могут быть получены аналитические выражения для границ интервала охвата (и коэффициента охвата в симметричном случае) [6], которые используются в рамках данного приложения. Плотность распределения TSP(a, m, b, p) имеет вид

$$f(x|a, m, b, p) = \begin{cases} \frac{p}{b-a} \left(\frac{x-a}{m-a} \right)^{p-1}, & a < x \leq m, \\ \frac{p}{b-a} \left(\frac{b-x}{b-m} \right)^{p-1}, & m \leq x < b \end{cases}$$

($p > 0$, $a < m < b$); семейство включает в себя равномерное и треугольное распределения, а также довольно близкое приближение к нормальному.

Задача трансформирования описывается входным JSON-файлом (пример можно найти по ссылке [4]), содержащим среди прочего параметры распределения входных величин. В случае TSP-распределения эти параметры могут быть заданы либо напрямую, либо получены из предоставленных пользователем показателей точности (неопределенности); эта возможность представляется весьма полезной с практической точки зрения. А именно, если входному значению X_i приписаны математическое ожидание μ , неопределенность u и интервал охвата $[c_1, c_2]$ для заданного уровня вероятности P_0 , то параметры соответствующего TSP(a, m, b, p)

будут рассчитаны на основе системы уравнений вида

$$\left\{ \begin{array}{l} \frac{a + (p-1)m + b}{p+1} - \mu = 0, \\ (b-a)^2 \frac{p-2(p-1)q(1-q)}{(p+1)^2(p+2)} - u^2 = 0, \\ a + (m-a) \sqrt[p]{\frac{1}{2} \frac{b-a}{m-a} (1-P_0)} - c_1 = 0, \\ b - (b-m) \sqrt[p]{\frac{1}{2} \frac{b-a}{b-m} (1-P_0)} - c_2 = 0 \end{array} \right.$$

($q = (m-a)/(b-a)$; считаем, что $\frac{1-P_0}{2} \leq q \leq \frac{1+P_0}{2}$, иначе вид двух последних уравнений будет отличаться). В случае симметричного TSP-распределения достаточно указать μ , u и коэффициент охвата K для уровня вероятности P_0 .

Также были реализованы методы оценки параметров распределений входных величин на основе выборок экспериментальных данных [7] и на основе аппроксимации гистограмм. Доступна генерация коррелированных выборок входных данных.

2. РЕАЛИЗАЦИЯ ВЫЧИСЛЕНИЙ

Для эффективного решения задачи вычисления математических выражений (F) над большими массивами данных были использованы библиотеки NumPy и NumExpr языка Python. Основные требования включали: поддержку аналитического описания модели измерения для произвольного числа входных величин, в том числе возможность задания подвыражений, использование встроенных математических функций и высокую производительность на больших массивах данных (использование многopotочности). Библиотека NumExpr использует векторизацию, значительно снижая накладные расходы на управление циклами, автоматически распределяет вычисления между доступными ядрами процессора и использует различные оптимизации, позволяющие значительно ускорить вычисления. Например, сравнение с прототипом программы, использовавшим встроенную функцию `eval()`, показало ускорение примерно в 20 раз при использовании ноутбука с процессором Intel Core i5-12450N (8 ядер/12 потоков, Hyper-Threading); сравнение проводилось на синтетических моделях с 10–60 входными величинами.

Для результирующего распределения Y , отражающего неопределенность измеряемой величины, вычисляются: оценка измеряемого значения (математического ожидания), стандартная неопределенность (стандартное отклонение), интервал охвата для заданной вероятности охва-

та P_0 (односторонний, двусторонний вероятностно симметричный или наименьший). Также вычисляются минимальное и максимальное значения Y , его медиана, строится гистограмма распределения. Выходные данные представлены в формате JSON; гистограмма (с указанием среднего значения и интервала охвата) дополнительно сохраняется в виде графика.

ЗАКЛЮЧЕНИЕ

Реализованное программное обеспечение может быть использовано при решении некоторых типовых метрологических задач. Применение TSP-распределений при описании входных величин модели позволяет перейти от метрологически обоснованных показателей точности к параметрам распределения, а также осуществлять выбор параметров распределения на основе имеющихся экспериментальных данных. Очевидно, это не единственное семейство распределений, которое может быть рассмотрено в данном контексте; библиотека может быть дополнена другими моделями распределений. В то же время данное семейство, несмотря на свою простоту, может быть использовано для аппроксимации различных непрерывных распределений (в том числе асимметричных), что становится особенно актуальным при небольшом объеме экспериментальных данных, когда классические критерии проверки гипотез о типе распределения оказываются неэффективными и речь не идет о «тонком» подборе подходящего распределения. TSP-семейство распределений имеет ограниченный носитель, что повышает актуальность его использования, когда множества допустимых значений входных величин ограничены.

СПИСОК ЛИТЕРАТУРЫ

1. *Harris P., Cox M.* On a Monte Carlo Method for Measurement Uncertainty Evaluation and Its Implementation // *Metrologia*. 2014. V. 51. P. S176–S182.
2. *van der Veen A., Cox M.* Getting Started with Uncertainty Evaluation Using the Monte Carlo Method in *R* // *Accreditation and Quality Assurance*. 2021. V. 26. P. 129–141.
3. JCGM Supplement 1 to the “Guide to the Expression of Uncertainty in Measurement” — Propagation of Distributions Using a Monte Carlo Method. Joint Committee for Guides in Metrology, 2008.
4. Код приложения. <https://github.com/stepanov17/grid2025> (дата обращения 17.09.2025).
5. *Kotz S., Van Dorp J.R.* Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications. World Sci. Publ., 2004.
6. *Stepanov A. V., Chunovkina A. G.* On Choosing Two-Sided Power Distribution for Measurement Data // *Adv. Math. Comput. Tools Metrology Testing XIII*. 2025. P. 268–277.
7. *Тырсин А. Н.* Метод подбора наилучшего закона распределения непрерывной случайной величины на основе обратного отображения // *Вестн. Южно-Урал. ун-та. Сер. «Математика. Механика. Физика»*. 2017. Т. 9, № 1. С. 31–38.