

Колонка редколлегии

МЕТОД СТАТИСТИЧЕСКИХ ИСПЫТАНИЙ — ИНСТРУМЕНТ ИССЛЕДОВАТЕЛЯ

© А. И. Орлов

Для решения конкретных прикладных задач исследователи постоянно разрабатывают новые методы обработки статистических данных — результатов измерений (наблюдений, испытаний, анализов, опытов) и экспертных оценок. Свойства нового метода необходимо изучить. Какие интеллектуальные инструменты можно применить для такого изучения?

Мощным инструментом исследователей в области математической статистики являются предельные теоремы теории вероятностей — закон больших чисел, центральная предельная теорема и т.п. Ориентированные на математику специалисты призывают ими и ограничиться. Однако для практического использования статистических методов предельных теорем недостаточно. Необходимо определить границу объема выборки, начиная с которой можно пользоваться результатами, полученными с помощью предельных теорем. При этом следует выяснить, как принимать решения, если объем имеющихся данных меньше этой границы.

С середины XX в. исследователю доступна универсальная «отмычка» — метод статистических испытаний (метод Монте-Карло), другими словами, имитационное моделирование. Он основан на использовании последовательности псевдослучайных чисел, свойства которых напоминают свойства рассматриваемых в теории вероятностей случайных величин. Основная идея состоит в последовательном выполнении следующих этапов: разработке вероятностно-статистической модели реального явления или процесса; планировании статистического испытания, в котором случайные величины заменяются псевдослучайными, полученными с помощью того или иного датчика; проведении большого числа испытаний (тысяч или миллионов); анализе полученных результатов расчетов.

С каждым этапом связаны соответствующие проблемы адекватности имитационного моделирования. Так, для предельных теорем обычно справедлив тот

или иной принцип инвариантности, т.е. в пределе исчезает зависимость от конкретного вида распределения. Однако при изучении скорости сходимости выбор этого конкретного вида весьма важен, поскольку от него зависит итоговый результат статистического моделирования: один для нормального распределения, другой — для логистического, третий — для распределения Коши.

Датчики псевдослучайных чисел лишь имитируют случайность. Алгоритмы получения псевдослучайных чисел имеют достаточно короткое описание, в то время как по определению А. Н. Колмогорова 60-х годов (в рамках теории информации) описание случайной последовательности должно расти пропорционально длине этой последовательности. Кроме этой глобальной причины методологической несостоятельности датчиков псевдослучайных чисел, есть и частные недостатки. Например, у некоторых популярных до настоящего времени датчиков три последовательных значения связаны линейной зависимостью.

Значения, рассчитанные с помощью метода Монте-Карло, имеют погрешности, определяемые конечностью числа испытаний. При оценивании вероятности события погрешность достигает величины $\sqrt{N}/2$, где N — число испытаний. Значит, для оценивания вероятности с точностью 10^{-6} необходимо $10^{12}/4$ испытаний. На практике провести такое количество испытаний невозможно.

Проблемы теории и практики статистических испытаний (Монте-Карло) заслуживают тщательного обсуждения. В разделе «Математические методы исследования» начинаем дискуссию о современном состоянии и перспективах развития статистического моделирования (имитационного моделирования). Предлагаем читателям принять участие в обсуждении математических методов исследования, использующих датчики псевдослучайных чисел. В нашем журнале предыдущая дискуссия о свойствах таких датчиков была проведена в 1985 – 1993 гг.