

УДК 519.24

DOI: 10.17223/00213411/63/4/40

В.П. ШУЛЕНИН

СВОЙСТВА РОБАСТНОСТИ МОДИФИЦИРОВАННЫХ ОЦЕНОК ХОДЖЕСА – ЛЕМАНА

В реальных данных физических экспериментов неизбежно появляются выпадающие наблюдения (выбросы), которые могут привести к существенным искажениям при статистической обработке данных. Поэтому важно использовать такие статистические процедуры, которые «защищены» от наличия выбросов в наблюдениях. К таким процедурам относится оценка параметра положения, предложенная Ходжесом и Леманом. Рассматриваются адаптивные оценки модифицированных вариантов оценки Ходжеса – Лемана. Исследуются свойства оценок в асимптотике и при конечных объемах выборки в рамках различных супермоделей, описывающих наличие выбросов и отклонения от гауссовской модели в сторону «утяжеления хвостов» распределений. Для построения адаптивных оценок используются выборочные оценки функционалов, описывающих степень «тяжести хвостов» распределений.

Ключевые слова: робастные оценки, выбросы, функция влияния, адаптивные оценки, метод статистических испытаний.

Введение

При решении прикладных задач, связанных с обработкой данных физических экспериментов, важно описать функцию распределения F_X наблюдаемой в эксперименте случайной величины X хотя бы в общих чертах посредством немногих простых параметров. Такая дескриптивная статистика связана с введением количеств (мер), характеризующих различные особенности распределений, и с нахождением оценок для них по исходным наблюдениям X_1, \dots, X_n над изучаемой случайной величиной X . Типичными количественными характеристиками являются математическое ожидание $M(X)$ и медиана $MED(X)$, которые определяют параметр положения θ случайной величины X . Для описания разброса случайной величины (параметра масштаба) обычно используют дисперсию $D(X)$ или стандартное отклонение $[D(X)]^{1/2}$, и интерквартильный размах. Общие требования для описания параметров положения и масштаба в виде функционалов $T(F_X)$ от распределения F_X наблюдений X_1, \dots, X_n над изучаемой случайной величиной X обсуждаются в работе [1, 2]. Для параметра масштаба робастные альтернативы стандартному отклонению, которые в отличие от него «защищены» от наличия выбросов в выборке, рассмотрены в [3]. К настоящему времени известно большое число различных оценок параметра положения (см., например, [4–6]). Большинство из них могут быть отнесены к общим классам оценок, либо к классу M , либо L , либо R [7–12]. Исторически первой и легко воспринимаемой оценкой параметра положения является выборочное среднее \bar{X} , которое является оценкой математического ожидания, построенной методом подстановки [13]. Стремление обосновать целесообразность этой оценки и наделить её некоторыми свойствами оптимальности даже послужило основным аргументом для введения К. Гауссом нормального распределения (см. цитату К. Гаусса, приведённую в работе [14], и см. также работы Тьюки [15, 16]). Автор работ [15, 16] отмечает, что: «предполагаемые достоинства средней арифметической были использованы для введения гауссовского распределения, а предполагаемая «истинность» гауссовского распределения была использована для доказательства оптимальности средней арифметической. Нет сомнений в том, что этот круговорот был очевиден для многих, работавших в этой области...». Далее Тьюки пишет: «к началу тридцатых годов, когда я впервые столкнулся с практическим использованием анализа данных, оно, по крайней мере, в наиболее искушенных руках, достигло стадии, когда получение итогового результата представляло собой исследование данных в том отношении, что явно «хорошие» данные представлялись средними значениями, а явно «плохие» наборы чисел, например, – медианами. Подобное разветвление вовсе не означало, что итоговый результат не может быть фиксированной функцией данных; оно означает только то, что эту фиксированную функцию не так-то просто записать... Прошло почти тридцать лет с тех пор, как я встретил специалиста по анализу данных, достаточно искушенного для того, чтобы избегать пользоваться средним арифметическим в большинстве случаев».

Уважаемые читатели!

Доступ к полнотекстовой версии журнала
«Известия высших учебных заведений. Физика»
осуществляется на платформе
Научной электронной библиотеки eLIBRARY.RU
на платной основе:

<https://www.elibrary.ru/contents.asp?titleid=7725>