Math-Net.Ru

Общероссийский математический портал

Я. Ю. Никитин, Рецензия на книгу Болдин М. В., Симонova Г. И., Тюрин Ю. Н. “Знаковый статистический анализ линейных моделей”, *Теория вероятн. и ее примен.*, 1999, том 44, выпуск 3, 685–687

DOI: <https://doi.org/10.4213/tvp815>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.224.60.19

26 апреля 2023 г., 03:11:44

6. Кестен Г. Теория просачивания. М.: Мир, 1986.
7. Mahmoud H. Evolution of Random Search Trees. Chichester: Wiley, 1992, 324 p.
8. Hoare C. A. R. Quicksort. — Comput. J., 1962, v. 5, № 1, p. 10–15.
9. Кнут Д. Искусство программирования для ЭВМ. Т. 3: Сортировка и поиск. М.: Мир, 1978, 844 с.

6.01.1999 г.

Ватутин В. А.

Болдин М. В., Симонова Г. И., Тюрин Ю. Н. Знаковый статистический анализ линейных моделей. М.: Наука, 1997, 226 с.

Линейные модели, в том числе модели множественной регрессии и модели для различных планов дисперсионного и ковариационного анализа, играют, как известно, огромную роль в современных приложениях статистики. Классическая теория оценивания и проверки гипотез в таких моделях, берущая свое начало в знаменитом методе наименьших квадратов Гаусса–Лежандра, предполагает, что случайные составляющие (ошибки) наблюдений распределены по нормальному (гауссовскому) закону.

Постепенно статистики стали осознавать, что математические предположения, лежащие в основе гауссовской модели, на практике могут не выполняться. Появилась необходимость в построении непараметрической теории, пригодной для широкого класса распределений ошибок. На этом пути в первой половине XX века появились ранговые статистические процедуры. Они основаны на предположении, что случайные ошибки независимы и имеют одинаковое непрерывное распределение.

Несколько десятилетий интенсивного развития ранговых методов для анализа линейных моделей привели к богатой и далеко продвинутой теории, изложенной, например, в монографиях Пури и Сена [1] и Хеттманспергера [2], опубликованных в 80-х годах.

Между тем, возможен и другой вариант построения непараметрической теории, основанный не на рангах, а на знаках случайных ошибок или остатков. Сохраняя условие независимости ошибок, здесь можно отказаться от их одинаковой распределенности, предполагая лишь, что распределения ошибок обладают определенной симметрией, например, имеют нулевые медианы. Простейший представитель знаковых процедур — критерий знаков — считается самым старым статистическим критерием и по существу применялся Дж. Арбетнотом [3] еще в начале XVIII в.

Привлекательным свойством знаковых статистических методов является то, что при упомянутой медианной несмещенности ошибок знаковые тестовые статистики свободны от распределения исходных данных даже при конечных объемах выборки, а не только асимптотически. Опираясь на тестовые статистики, с помощью приема «обращения» оказывается возможным получать точные доверительные интервалы и строить точечные оценки для неизвестных параметров. Эти оценки, а также тестовые статистики являются асимптотически нормальными при минимальных предположениях о распределении наблюдений. Наконец, в отличие от классических процедур типа метода наименьших квадратов, знаковые процедуры устойчивы к выбросам и грубым ошибкам в данных.

Все эти несомненные достоинства знаковых методов, наряду с их привлекательной вычислительной простотой, поставили на повестку дня написание обобщающей монографии по знаковому статистическому анализу. Рецензируемое издание и является первой такой монографией в мировой литературе. Работы его авторов М. В. Болдина, Г. И. Симоновой и Ю. Н. Тюриня сыграли в последние годы наиболее важную роль в этой области статистики, поэтому неудивительно, что книга почти целиком построена на их собственных исследованиях предыдущих 10–15 лет.

Книга состоит из введения, в котором прекрасно изложены как цели, так и основная идеология исследования, и двух частей. Первая часть посвящена регрессионному

и дисперсионному анализу; точнее, анализу факторных экспериментов, поскольку дисперсии в этом не участвуют — ни выборочные, ни теоретические. Существование последних даже не предполагается. Во второй части рассматриваются знаковые методы для линейных временных рядов. Подробно изучены модели авторегрессии. Впрочем, без серьезных нововведений разработанные авторами методы могут действовать и для комбинированных моделей авторегрессии — скользящего среднего. Для каждой из рассмотренных моделей авторы дают решения основных статистических задач: оценивания (точечного и интервального), проверки гипотез (как точечных, так и линейных) и задачи асимптотического исследования предложенных правил, в том числе их устойчивости.

Метод асимптотического исследования основан на изучении и использовании остаточных эмпирических, а также случайно-взвешенных эмпирических процессов и полей. Эти объекты естественно появляются сначала как статистики локально-оптимальных знаковых тестов для проверки точечных гипотез. Затем эти статистики служат для оценивания неизвестных параметров. Эта линия — через тесты и доверительные множества к точечным оценкам — составляет существенную особенность рецензируемой книги.

Разделы книги, посвященные асимптотическому анализу, технически наиболее трудны. Здесь тоже есть своя особенность: все асимптотические выводы авторы основывают на теоремах о равномерной асимптотической линейности знаковых статистик на расширяющихся компактах.

В процессе работы над книгой М. В. Болдину, Г. И. Симоновой и Ю. Н. Тюри-ну удалось доказать ряд новых интересных результатов и упростить доказательства старых, заполнить ряд пробелов в теории, собрать обширный материал воедино и привести его в логичную и стройную систему. Можно согласиться с авторами, утверждающими на с. 7, что «по своим возможностям и кругу доступных им задач знаковые методы оказываются вполне сопоставимыми с ранговыми». Это стало особенно ясно после создания их монографии.

Достоинством рецензируемого издания является также большое количество содержательных и со вкусом подобранных примеров, вычисления в которых проведены до конца, с помощью специально разработанных вычислительных процедур (пакета программ «Sign»). В тексте книги множество таблиц и графиков, которые облегчают понимание и использование разработанных методов.

Книга написана в ясном и четком стиле и вполне может служить основой для специального курса на математических факультетах университетов. Не случайно она немедленно появилась в английском переводе [4].

Основным недостатком непараметрических и, в частности, знаковых методов, считается их недостаточно высокая эффективность или мощность. Это естественная плата за использование не всей информации, содержащейся в наблюдениях, а только ее части, отвечающей вектору знаков. Действительно, в гауссовской регрессии эффективность знаковой оценки в сравнении с оптимальной оценкой равна, как известно, $2/\pi$. Однако для ошибок с «тяжелыми» хвостами, которые часто встречаются в современных моделях естествознания и экономики, эффективность знаковых процедур относительно процедур метода наименьших квадратов может быть и больше единицы и вообще сколь угодно большой. Некоторая потеря в эффективности компенсируется выигрышем в большей робастности.

Вообще же здесь уместно привести известное в статистическом фольклоре высказывание Ч. Эйзенхарта [5]: «“Практическая мощность” теста определяется произведением математической мощности на вероятность того, что тест будет использован». Представляется, что рецензируемая монография ярко демонстрирует привлекательность знаковых процедур в линейных моделях и тем самым существенно повышает их «практическую мощность».

В монографии принята следующая система ссылок: все заимствованные резуль-

таты снабжены ссылками, а отсутствие ссылки означает, что ранее этот факт в литературе не обсуждался. Поскольку большая часть приводимых результатов принадлежит самим авторам, читателю ввиду отсутствия ссылки подчас трудно понять, кто из них и когда установил тот или иной результат. В книге явно не хватает историко-библиографических замечаний в конце текста. Такие замечания, снабженные дополнительной современной библиографией, несомненно, пошли бы на пользу читателю.

Спорной представляется мне и идея поместить классические монографии Андерсона, Лемана, Гаека и Шидака, Хеннана и т.д. в списке литературы дважды: в русской и английской части. На мой взгляд, это создает определенную путаницу и не дает никаких преимуществ. Впрочем, эти недостатки легко устранить при последующих изданиях.

Монография М. В. Болдина, Г. И. Симоновой и Ю. Н. Тюрина — заметный и весьма существенный вклад в статистическую литературу. Она безусловно окажет стимулирующее влияние на теоретиков и будет полезна всем тем, кто применяет непараметрические методы в прикладных статистических задачах.

СПИСОК ЛИТЕРАТУРЫ

1. *Puri M. L., Sen P. K.* Nonparametric Methods in General Linear Models. New York: Wiley, 1985, 399 p.
2. *Hettmansperger T. P.* Statistical Inference Based on Ranks. New York: Wiley, 1984, 323 p.
3. *Arbuthnott J.* An argument for divine providence taken from the constant regularity observed in the birth of two sexes. — *Philos. Trans.*, 1710, v. 27, p. 186–190.
4. *Boldin M. V., Simonova G. I., Tyurin Yu. N.* Sign-Based Methods in Linear Statistical Models. Providence, RI: Amer. Math. Soc., 1997.
5. *Nelson L. S.* A sign test for correlation. — *J. Qual. Technol.*, 1983, v. 15, № 4, p. 199–200.

24.XII.1998

Никитин Я. Ю.

Goldreich O ed. **Modern Cryptography, Probabilistic Proofs and Pseudorandomness**. Berlin–Heidelberg–New York: Springer-Verlag, 1999, xv + 182 p.

Три темы, составившие название книги и соответствующие названиям трех ее глав, несмотря на кажущуюся разнородность, тесно связаны друг с другом и с принципиальными основами теории вероятностей.

Слово «криптография» долгое время обозначало методы защиты государственных и военных тайн и неспециалистами произносилось с опаской. Однако с начала 80-х годов за рубежом стали открыто проводиться конференции по криптографии, и это слово стало привычным в западной научной литературе, а после изменения общественного строя — и в нашей стране. Причинами таких изменений стали по меньшей мере два обстоятельства: быстрое развитие электронных средств связи (удобных, но, как правило, полностью открытых для постороннего вмешательства) и необходимость надежной защиты переписки и конфиденциальной информации большого количества фирм, не доверяющих государственным органам, поскольку они могут находиться под влиянием конкурентов и транснациональных корпораций. Чтобы удовлетворить запросы расширившегося круга потребителей методов защиты информации, были использованы качественно новые идеи, предложенные криптографами, математиками и специалистами по вычислительным наукам. В результате «современная криптография» стала математически строгой наукой, тесно связанной с теорией вероятностных доказательств, и стимулировала разработку новых подходов к разработке понятия псевдослучайности.