

КРИТИКА И БИБЛИОГРАФИЯ

В. Н. Перегудов. Метод наименьших квадратов и его применение в исследованиях. М., «Статистика», 1965

Рецензируемая книга представляет несомненный научный интерес; она — результат долголетнего плодотворного исследования автором темы, которая в последние десятилетия приобрела немалую важность для статистической науки.

Способ наименьших квадратов известен математикам, астрономам, геодезистам, статистикам уже более полутора веков, во всяком случае — из работ Лежандра (1806) и Гаусса (1809). Первоначальные задачи применения этого способа были достаточно ограничены: найти вероятнейшую числовую величину объекта по многократным не вполне совпадающим измерениям или вычислить несколько коэффициентов уравнения зависимости, если число наблюдений превышает число этих коэффициентов. Правда, и в этой роли способ наименьших квадратов приносил немалую пользу любым исследователям.

Главная задача современной прогрессивной статистики — исследование статистических закономерностей, в первую очередь закономерностей экономической жизни. Эти закономерности складываются под воздействием многих факторов и потому почти всегда достаточно сложны. Задача современного статистика — проанализировать связь явления с его факторами, умело отобрать главные факторы, построить нужную формулу уравнения связи, вычислить по имеющимся эмпирическим данным коэффициенты этого уравнения, статистически оценить степень существенности каждого фактора для исследуемой закономерности (связи). Немалое значение при этом имеет целесообразная, экономная схема вычислений — решение системы нормальных уравнений (иногда многих и сложных) при отыскании нужного уравнения связи.

Почти три четверти века исследователи пользовались для решения такой системы обычным способом уравнивания коэффициентов и последовательного исключения неизвестных (как поступал и Гаусс). Только в 1878 г. американский геодезист Дулитль придумал другой, бо-

лее экономный способ, освобождающий от громоздких вычислений при обычном способе. Однако «схема Дулитля» оставалась мало знакомой нашим статистикам.

В 30-х годах кафедра статистики Сельскохозяйственной академии им. Тимирязева (В. С. Немчинов в сотрудничестве с В. Н. Перегудовым) поставила широкие исследования в области сортоиспытания. Эти исследования приводили к отысканию сложных уравнений связи — уравнений множественной, часто криволинейной, корреляции. Одновременно возникла необходимость изобрести наиболее экономные схемы вычисления параметров таких уравнений. Схема Дулитля была постепенно изменена, усовершенствована. Новая схема была описана Немчиновым в 1945 г., в главе 22 его известного курса «Сельскохозяйственная статистика с основами общей теории». Она была тесно связана с применением интерполяционного ряда Чебышева (1855). «Эта схема, — писал В. С. Немчинов, — разработана на кафедре статистики ТСХА, причем порядок вычислений предложен В. Н. Перегудовым, а интерпретация схемы в полиномах Чебышева принадлежит автору книги» (стр. 330). Глава 22 (а вместе с нею и вся книга) заканчивалась следующим тезисом: «Основной путь дальнейшего развития методов статистического анализа данных хозяйственного опыта и результатов статистических наблюдений, несомненно, связан с теорией метода наименьших квадратов, трактуемого в полиномах Чебышева, и неразрывно с ним увязанным дисперсионным анализом» (стр. 346)\*.

Через год (1946) вышла следующая книга В. С. Немчинова: «Полиномы Чебышева и математическая статистика». В ней мы видим новый шаг вперед в раз-

\* В последнее двадцатилетие характерна (в связи с углублением теории) замена термина «способ наименьших квадратов» термином «метод наименьших квадратов».

работке метода наименьших квадратов. В главе 1 Немчинов подробно остановился на связи этого метода с ортогональными полиномами Чебышева. В главе 2, упомянув о схеме Дулитля и прежней схеме Немчинова — Перегудова, автор писал: «Однако в настоящее время могут быть предложены новые, более удобные схемы решения нормальных уравнений, основанные на иной системе полиномов Чебышева» (стр. 22). Далее достаточно подробно описывались: схема Дулитля (схема *D*, стр. 27—34); новая схема самого Немчинова (схема *N*, стр. 34—45); схема Перегудова (схема *P*, стр. 45—49). При любой из этих схем «полностью отпадает Гауссовский способ совместного решения нескольких уравнений, который до сих пор лежал в основе метода наименьших квадратов. Все эти преимущества имеют огромное практическое значение» (стр. 35). О схеме *P* Немчинов писал, что Перегудов исходил из других теоретических предположений, чем Чебышев; однако в своей окончательной форме его схема *P* вплотную подошла к методу Чебышева (стр. 49).

\* \* \*

После дополнительной переработки, совершенствования и практической проверки своей схемы на протяжении почти двух десятилетий, В. Н. Перегудов изложил ее принципиальные, теоретические основы в 1964 г. (статья «Метод наименьших квадратов в экономическом исследовании» в сборнике «Планирование и экономико-математические методы. К 70-летию со дня рождения академика В. С. Немчинова»), а годом позже дал подробное обоснование этой схемы и показал ее практическое применение в рецензируемой книге.

Взамен двух ранее известных принципов — именно, минимума остаточной суммы квадратов (у Гаусса) или максимума веса каждого коэффициента уравнения (у А. А. Маркова), — Перегудов выдвигает иной принцип, сводящийся к трем условиям: все наблюдения независимы; погрешности частных уравнений независимы от наблюдений  $y_i$ ; коэффициенты  $a_1, a_2, a_3, \dots$  искомого уравнения суть линейные функции наблюдений  $y_i$  (см. стр. 430—432 названной статьи).

В своей конечной разработке схема Перегудова дает возможность: а) осуществить идею Чебышева — постепенно развертывать полином (уравнение), вводить в него дополнительные члены без полного пересчета заново, путем некоторых поправок в вычислениях, — иначе говоря, последовательно уточнять искомое уравнение связи; б) в случае необходимости исключать любой член уравнения, оказавшийся несущественным; в) производить все вычисления наиболее экономным способом; г) контролировать каждый этап вычислений.

Книга В. Н. Перегудова состоит из пяти глав. В первой из них автор выясняет на конкретном примере сущность статистической связи и математической модели явления, излагает суть и главные положения метода наименьших квадратов, подводит читателя к анализу дисперсии. При этом подробно описывается специальная схема (АВ) вычисления коэффициентов искомого линейного уравнения регрессии, упрощенная «неполная» схема (В) для «некоторой научной реконструкции на более ранней стадии исследования» (стр. 56), даются числовые примеры расчетов по этим схемам, способы контроля каждой стадии вычислений и необходимая степень их точности. Во второй главе автор знакомит читателя с различными критериями существенности, умело связывая их воедино, и показывает приемы оценки степени существенности компонентов регрессии. При этом предлагается наиболее общая «полная» расчетная схема (АВС) и подробный пример вычислений по этой схеме. В третьей главе рассматриваются вопросы, связанные с отысканием параболических уравнений регрессии с одной независимой переменной: вычисление коэффициентов; оценка их существенности; исключение из уравнения одних и включение других степеней переменной; особо — оценка существенности свободного члена уравнения (с весьма интересными соображениями автора о его смысле и правомерности в уравнении, стр. 153—154); вопросы интерполяции и правомерности экстраполяции найденного уравнения; приемы сравнения двух регрессий. Все эти вопросы рассматриваются в непосредственной связи с числовыми величинами, находимыми при помощи вычислительных схем В. Н. Перегудова. В четвертой главе автор рассматривает аналогичный комплекс вопросов для уравнений множественной регрессии, линейной и параболической.

В первых четырех главах исходным числовым материалом для автора служили данные опытной агрономии и агрометеорологии, например: зависимость урожайности двух сортов подсолнечника от их средней урожайности (табл. 1, стр. 9); зависимость урожайности трех сортов пшеницы от четырех метеорологических факторов (табл. 3, стр. 24). В последней, наиболее обширной главе В. Н. Перегудов показывает применение метода наименьших квадратов в собственно экономических исследованиях на примере анализа производственной деятельности совхозности колхозов (89 колхозов). В этом примере за три признака-фактора взяты вся площадь пашни и удельные величины затрат труда и стоимости основных средств производства на 100 га угодий, а за три результивных признака — три удельные

величины валовой продукции сельского хозяйства колхозов: на 100 га угодий, на одного трудоспособного и на тысячу рублей основных средств производства (стр. 252). Построив исходную «математическую модель» исследуемого экономического явления в виде десятичного уравнения параболы второго порядка с тремя независимыми переменными (стр. 253), В. Н. Перегудов затем последовательно улучшает это уравнение (§ 28), связывает три результативных признака в один сводный (§ 29 и 30), ичисляет «сводный показатель потенциальных возможностей» производства (§ 31), включает в уравнение регрессии новый фактор с качественными градациями — почву (§ 32) и т. д. Числовой основой всех этих расчетов и анализа служат многочисленные и разнообразные величины, получаемые при помощи вычислительных схем автора.

\* \* \*

Рассматривая рецензируемое издание в целом как значительное явление в области статистики, мы не можем не высказать и некоторые замечания.

Обидно мало внимания уделяет В. Н. Перегудов истории развития способа метода наименьших квадратов, даже работам П. Л. Чебышева, А. А. Маркова, В. С. Немчинова. В сущности она исчерпывается несколькими фразами во «Введении». А так как В. Н. Перегудов на протяжении всей книги пользуется единственно термином «метод наименьших квадратов», то постепенная трансформация изначального *способа* в современный *метод* не получает четкого, выпуклого отражения.

Обычно историю способа наименьших квадратов ведут «от Гаусса» (1809) или «от Лежандра» (1806). Так поступает и В. Н. Перегудов (стр. 3). В связи с этим не лишена интереса новейшая работа О. Б. Шейнина по истории теории ошибок (неопубликованная рукопись, хранящаяся в ВИНТИ, 1963 г.)\*. Ее автор исследует: труды Дан. Бернулли (1777) и комментарии к ним Эйлера; работы некоторых последующих авторов, предлагавших

\* О. Б. Шейнин. К истории оценок непосредственных измерений и закона распределения случайных ошибок. Всес. ин-т науч. и техн. информ. М., 1963; 37 стр.— См. ее изложение в реферативных журналах: «РЖ Отдельный выпуск Геодезия», 1963, № 5, реф. 5.52.75; также «РЖ Математика», 1963, № 8, реф. 8.А.30.

оценки апостериорного характера; особенно же — результаты работ американца Роберта Эдрейна, который за год до Гаусса опубликовал два вывода закона нормального распределения случайных ошибок и применил его к выводу принципа арифметической средней и наименьших квадратов. О. Б. Шейнин отмечает логическую общность вывода нормального распределения и принципа наименьших квадратов у Эдрейна (Филадельфия, 1808) и у Гаусса (1809).

Вызывают возражения некоторые термины, применяемые В. Н. Перегудовым. «Уравнение регрессии  $y$  по  $x$ » звучит выразительнее, чем «... $y$  на  $x$ ». Нет нужды в термине «стандартное отклонение» (стр. 67 и далее) взамен общепринятого термина «среднее квадратическое (короче — квадратическое) отклонение».

Необходимо отметить и чисто издательские погрешности в рецензируемой книге. Крайне затрудняет ориентацию читателя в сложном содержании книги то, что оглавление (стр. 2) содержит наименования только пяти крупных глав, без заголовков тридцати пяти параграфов, их образующих; колонтитулы же в книге отсутствуют.

Важная для читателя библиография помещена после «Приложения» (таблицы значений критерия  $F$ ) и даже не упомянута в оглавлении. Перечень работ дан по алфавитному признаку, а не в хронологическом порядке, который показал бы развитие теории и практики способа (метода) наименьших квадратов от начала XIX в. до наших дней. Но всего хуже то, что напечатана эта библиография со многими ошибками. В заглавие латинского текста основополагающей работы Гаусса (1809) проникли две грубые ошибки. Из хорошо известного перевода этой работы на русский язык, сделанного в 1859 г. студентом Московского университета Догелем, издательство ухитрилось сотворить две отдельные оригинальные работы: Гаусса, с заглавием «Теория движения...», и якобы самого Догеля, с непостижимым заглавием «М., 1859!» Известная работа А. С. Чеботарева (1936) названа «Способы наименьших квадратов...». Библиографические описания сделаны с нарушением единых правил.

Наконец, нельзя обойти молчанием смехотворно малый тираж книги — 1500 экз. Уже через несколько дней по выходе в свет интересная книга стала библиографической редкостью.

Ф. Д. Лившиц