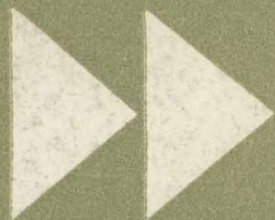


Академия
наук
СССР



Статистический
анализ
сельско-
хозяйственного
производства

Издательство «Наука»

Академия наук СССР
Центральный экономико-математический институт

**анализ
сельскохозяйственного
производства**

Ответственный редактор
Т. В. РЯБУШКИН



Издательство «Наука» Москва 1984

Ученые записки по статистике, т. 47.

В сборнике представлены статьи по актуальным проблемам статистического моделирования, анализа, прогнозирования, разработки показателей и статистических методов анализа эффективности сельскохозяйственного производства. Методические разработки проверяются на фактическом материале на уровне хозяйств областей, республик, отраслей.

Для статистиков, экономистов, занимающихся вопросами сельского хозяйства, плановиков.

Редакционная коллегия:

С. А. АЙВАЗЯН, А. Г. ВОЛКОВ,
В. М. СИМЧЕРА (зам. ответственного редактора),
Е. Г. ЯСИН

Научные редакторы тома:

М. М. ЮЗБАШЕВ, А. И. МАНЕЛЛЯ

Статистический анализ сельскохозяйственного производства

Утверждено к печати Центральным экономико-математическим институтом АН СССР

Редактор издательства В. С. Баковецкая
Художественный редактор И. Ю. Нестерова
Технический редактор М. Ю. Соловьева
Корректор Н. Г. Васильева

ИБ № 26593

Слано в набор 13 09 83 Подписано к печати 11 01 84 Т-05504 Формат 84×108^{1/32} Бумага типографская № 3 Гарнитура обыкновенная Печать высокая Усл. печ. л. 6,51 Усл. кр.-отт. 13,2 Уч.-изд. л. 13,9 Тираж 1450 экз. Тип. зак. № 809 Цена 2 р. 10 к.

Издательство «Наука» 117864 ГСП-7, Москва В-485, Профсоюзная ул., 90, 4-я типография издательства «Наука» 630077, Новосибирск, 77, Станиславского, 25

С 0701000000—065
042(02)—84 81-84-I

© Издательство «Наука», 1984 г.

ПРЕДИСЛОВИЕ

До настоящего времени специально ни один том «Ученых записок по статистике» не был посвящен статистическим методам исследования сельскохозяйственного производства. Между тем накоплен большой опыт применения математико-статистических методов в исследованиях экономики и технологии сельского хозяйства. С 1967 г. состоялось пять всесоюзных конференций по применению математической статистики в анализе экономики сельскохозяйственного производства. По этой проблеме публиковались отдельные статьи исследовательского характера в журналах «Экономика сельского хозяйства», «Экономика и математические методы», «Вестник статистики» и др. Однако обобщающих монографий или сборников, посвященных данной проблеме нет.

Выполнение Продовольственной программы, принятой майским (1982 г.) Пленумом ЦК КПСС, — основная политическая и экономическая задача народного хозяйства СССР на ближайшее десятилетие. Оперативное и качественное отражение хода выполнения Продовольственной программы, анализ процесса агропромышленной интеграции выходят за рамки внутриотраслевой сельскохозяйственной статистики, становятся на ближайшие годы главными проблемами советской статистической науки.

Июньский (1983 г.) Пленум ЦК КПСС поставил перед общественными науками задачу полнее раскрывать объективные законы общественного развития, своевременно выявлять назревающие тенденции.

Отдельным сторонам совершенствования статистического анализа и разработки предупреждающей информации посвящен предлагаемый том «Ученых записок по статистике».

Основные идеи, развиваемые авторами сборника, таковы.

Сельскохозяйственное производство с точки зрения экономического, статистического, кибернетического подхода — сложная открытая система, развивающаяся под влиянием как внутренних, управляемых факторов, так и неуправляемых факторов природной среды. В связи с

этим характеризующие его показатели следует рассматривать как величины ее случайной составляющей.

В анализе, планировании, управлении сельскохозяйственным производством необходимо учитывать стохастическую составляющую его показателей, применяя, изменяя, приспособлявая для этого вероятностные статистические и экономико-математические методы.

Огромную роль в совершенствовании анализа сельскохозяйственного производства имеет обеспечение достоверной и всесторонней информацией. Совершенствование системы показателей, методов их определения, автоматизация сбора, хранения, обработки, передачи информации имеет большое значение для успешного решения задач, поставленных перед сельским хозяйством партией и правительством.

Содержание статей, включенных в том, охватывает следующие проблемы: статистические методы анализа и моделирования результатов и эффективности сельскохозяйственного производства, прогнозирования его отдельных сторон, совершенствование системы показателей развития сельского хозяйства, создание подсистемы «Статистика сельского хозяйства» автоматизированной системы государственной статистики. Отдельные статьи имеют более широкое методологическое значение, выходящее за рамки одной отрасли.

Группа статей посвящена методам анализа взаимосвязей и моделированию сельскохозяйственного производства. О. П. Крастинь в статье «Проблемы исследования корреляционно связанных факторов» показывает, что мультиколлинеарность не является лишь ошибкой построения моделей, но с необходимостью вытекает из экономических законов сельского хозяйства. В статье исследуются возможности разложения множественных коэффициентов регрессии на составляющие на основе обратной матрицы коэффициентов системы нормальных уравнений и дается интерпретация этих составляющих. Методические положения иллюстрируются практически расчетами по Латвийской ССР.

М. И. Маршак в статье «Некоторые обобщения теории переменной корреляции», основываясь на известных примерах применения переменной корреляции в исследованиях С. П. Боброва, Н. С. Четверикова и Б. С. Ястремского, рассматривает более общий случай процессов с неустойчивыми параметрами. Анализируются две принципиально различающиеся задачи: проверка устойчивости

параметров корреляции и регрессии внутри выбранного временного интервала и использование рекуррентных формул для уточнения этих параметров при получении новых результатов наблюдений. Эти задачи решаются с использованием известных процедур проверки статистических гипотез и методов многомерной классификации объектов.

В статье И. М. Гурвич и В. П. Володина «Вопросы совершенствования анализа эффективности сельского хозяйства в условиях регистра» рассматриваются современные методы накопления, хранения, информационного поиска и выдачи справок по данным статистики сельского хозяйства с помощью ЭВМ. Разрабатываемая в ЦСУ СССР система хранения и обработки данных получила название регистра сельскохозяйственных предприятий. В статье освещается опыт создания и эксплуатации этой системы в Эстонской ССР.

В другой статье В. П. Володина «Методологические основы изучения эффективности сельского хозяйства на уровне отрасли» раскрываются основные направления отраслевого измерения и оценки эффективности производства.

В. А. Свободин и Н. Л. Латипов в статье «Моделирование эффективности растениеводства с учетом погодных условий» предлагают способы отражения в регрессионных моделях погодных условий для изучения влияния этих условий на продуктивность земледелия. Результаты расчетов представлены в виде номограмм, которые авторы рекомендуют использовать в анализе и прогнозировании показателей эффективности растениеводства.

Практические рекомендации по использованию параметров множественной регрессии в сравнительном анализе содержатся в статье И. Р. Годмане и др. «Оценка производственной деятельности сельскохозяйственных предприятий на основе регрессии». Результаты анализа выражаются в виде диаграммы, на которой оценка производственной деятельности отдельных хозяйств дается одновременно по субъективным факторам и объективным условиям производства.

В статье В. А. Колемаева «Система моделей развития агропромышленного комплекса» рассмотрена система моделей АПК, состоящая из центральной оптимизационной и периферийных статистических моделей, предназначенных для оценки структурных параметров центральной модели. Система позволяет в условиях заданных ограни-

чений на ресурсы получать оптимальные сбалансированные планы развития АПК.

В статье А. А. Френкеля и др. «Прогнозирование производительности труда в сельском хозяйстве РСФСР на основе динамической модели» изложена методика построения динамической модели, которую авторы предлагают использовать для прогнозирования производительности труда в сельском хозяйстве РСФСР на перспективу. При построении модели использованы методы кластерного анализа, регрессионного анализа, экспоненциального сглаживания.

А. Е. Шибалкин в статье «О зависимости динамики урожайности зерновых культур от метеорологических условий» рассматривает вопросы статистического анализа влияния агрометеорологических условий на урожайность, предлагает обобщенный показатель метеоусловий и проводит регрессионный анализ связи этого показателя с урожайностью.

Методика и результаты ретроспективной оценки прогнозов урожайности на пятилетие 1971—1975 гг., проведенной по 40 динамическим рядам урожайности различных культур по многим областям и республикам СССР, излагаются в статье М. Г. Каяйкиной, Р. П. Рудаковой «Методика и результаты ретроспективной проверки авторегрессионных прогнозов урожайности».

В статье Л. И. Ващукова и др. «Анализ условий текущего года и их влияния на ожидаемую урожайность» рассмотрена проблема прогнозирования урожайности текущего года, дан подробный анализ влияния факторов и описана модель постадийной оценки урожайности зерновых культур на уровне СССР.

В статье А. И. Манелли «Районы синхронных колебаний урожайности зерновых культур» излагается метод построения районов синхронных колебаний урожайности и взаимно компенсирующих районов, приводящий к однозначному разбиению административных единиц на группы. Результаты обобщаются в виде картограмм и таблиц, которые могут найти применение при прогнозировании валовых сборов зерна в текущем году, изучении связи колебаний урожайности с метеорологическими факторами, в работах по специализации и размещению сельскохозяйственного производства.

В статье А. М. Гатаулина «Основные тенденции динамики урожайности зерновых культур в Центральном районе Нечерноземной зоны» на базе рядов динамики

урожайности зерновых культур за два исторических периода (1883—1915 гг. и 1957—1978 гг.) исследуются закономерности динамики урожайности и рекомендуются методические подходы к дифференциации условий года на основе статистической обработки рядов динамики.

В. А. Тяпкин в статье «Применение экономико-статистических методов при сравнительной оценке качества земли» для разработки балльной оценки и сравнительного анализа качества почв использует метод аналитической группировки и дисперсионного анализа. Автор получает интересные результаты по связи балльных оценок почв Кустанайской и Владимирской областей с урожайностью, производительностью труда, себестоимостью и другими показателями эффективности земледелия.

В статье А. К. Алмаши, В. М. Симчеры, В. В. Шепы «Анализ эффективности интенсификации сельскохозяйственного производства в Венгрии» рассматриваются проблемы эффективности сельскохозяйственного производства. Дается анализ процесса интенсификации, проводится сравнение результатов производственной деятельности государственных и кооперативных хозяйств. В статье применен неизвестный в статистической практике в СССР методический принцип группировки хозяйств с равными удельными весами каждой группы в совокупном объеме конечной продукции.

Настоящий сборник представляет интерес для широкого круга специалистов, занимающихся вопросами применения экономико-статистических методов анализа сельскохозяйственного производства, совершенствования статистических показателей, автоматизированной разработкой статистической отчетности.

О. Д. Крастинь

ПРОБЛЕМЫ ИССЛЕДОВАНИЯ КОРРЕЛЯЦИОННО СВЯЗАННЫХ ФАКТОРОВ

Уравнение множественной регрессии в качестве экономико-математической модели имеет наиболее ясное содержание. Оно максимально устойчиво, если все факторы-аргументы корреляционно независимы друг от друга. В противоположном случае, если два или более отобранных факторов имеют функциональную связь между собой (при построении линейной модели — линейную связь), множественное уравнение регрессии вовсе не определимо; идентифицировать его параметры невозможно.

Положение, когда факторные признаки линейно связаны между собой, в литературе получило название «мультиколлинеарность». Мультиколлинеарность и ее последствия в разных аспектах рассмотрены в специальной литературе [1, с. 275—281; 2, с. 123—125; 3, с. 222—228; и др.], а также в прикладных исследованиях. Обычно рассматриваются два случая.

1. Факторы в реальной совокупности связаны между собой очень тесно, практически функционально, но в результате ошибок наблюдения появляется корреляционная связь. Множественное уравнение формально может быть определено, но оно крайне неустойчиво и не имеет реального смысла. Такая ситуация оценивается как ошибка моделирования: были использованы лишние факторы. Модель следует упростить.

2. Факторы в реальной совокупности не должны иметь корреляционную связь, но такая появилась в результате ошибок наблюдения или других случайных причин. В таком случае рекомендуется определение множественного уравнения безвзирая на некоторую корреляционную связь факторов.

Мультиколлинеарность в классическом смысле имеет место, например, в том случае, если с логической точки зрения в модель включен один и тот же реальный фактор несколько раз, но с использованием различных показателей и способов измерения (например, минеральные удоб-

рения в физическом весе, в пересчете на действующее вещество, в стоимостном выражении и т. д.). Сбор данных по другой, при желании значительно более крупной, совокупности, как правило, не может улучшить стабильность модели и обеспечить надежные показатели влияния на результативный признак каждого аргумента-фактора.

При исследовании мультиколлинеарности задача сводится к проверке нулевой гипотезы, в результате которой принимается одно из двух решений: «мультиколлинеарность, вероятно, есть» или «мультиколлинеарности, наверно, нет».

Таким образом, мультиколлинеарность рассматривается как альтернативное свойство: если оно присутствует, то в полной мере, если же не присутствует, то не присутствует вовсе.

Не отрицая целесообразности такой постановки в рамках многих задач, считаем, что имеются и такие задачи, для которых она мало соответствует реальному содержанию исследуемых связей и закономерностей. Например, при исследовании модели объема сельскохозяйственной продукции или доходов на 100 га земли в качестве факторов отобраны: 1) стоимость основных фондов сельскохозяйственного назначения на 100 га и 2) стоимость оборотных средств на 100 га. Известно, что производство невозможно как без основных фондов, так и без оборотных средств. По мере интенсификации производства и основные фонды, и оборотные средства в расчете на 100 га должны увеличиваться. Поэтому гипотеза о независимости этих факторов должна быть отвергнута уже на основе логической аргументации. Но с такой же уверенностью надо отвергнуть гипотезу о коллинеарности этих факторов.

Различная степень специализации, различная технология производства, уровень организаторской работы и сотни других обстоятельств, обоснованных и необоснованных решений дают возможность организовать производство при различных соотношениях основных фондов и оборотных средств. Поэтому целесообразно не только исследовать их самостоятельное влияние на результаты производства, но даже, если это в конкретной ситуации возможно, оптимальные соотношения этих факторов. В таких ситуациях корреляционную связь факторов нельзя приписывать ни ошибкам моделирования, ни ошибкам наблюдения. Она присуща исследуемому объекту и проблеме. Поэтому следует по возможности приспособить су-

ществующие методы статистического анализа и интерпретацию получаемых результатов к особенностям реальных связей.

Оценка корреляционной связи факторов должна дать ответ на следующим образом сформулированный вопрос. Имеется ли в рассматриваемой совокупности достаточно большая независимая вариация корреляционно связанных факторов, необходимая для установления более или менее надежных характеристик влияния каждого фактора на результативный признак, взятого в отдельности, или таких возможностей нет? Если таких возможностей нет, вместо двух взаимосвязанных аналитических факторов в модель включается один синтетический фактор, объединяющий два первых. Если и это невозможно, один из взаимосвязанных факторов из модели исключается как «статистически незначимый» и его «влияние» на результативный признак передается другому оставшемуся в уравнении ведущему фактору. При этом меняется степень условности коэффициента регрессии, если его интерпретируют как показатель среднего влияния факторного признака на результативный [4, с. 66—74].

Из сказанного вытекает вывод, что статистическую незначимость того или иного фактора нельзя интерпретировать как экономическую незначимость. Причинно (экономически) незначимые факторы вовсе не включаются в модель уже на стадии предварительного (логического) ее обоснования. Статистическая незначимость фактора означает всего лишь то, что по обработанным данным нельзя устанавливать среднее самостоятельное влияние «незначимого» фактора на результативный признак, которое было бы достаточно устойчиво при изменении модели и состава совокупности. Соответствующий коэффициент регрессии мало пригоден или вовсе не пригоден для решения задач анализа и прогнозирования. Но из этого никак не следует, что такой норматив не может быть установлен в других условиях, по данным другой, обычно более крупной совокупности.

В настоящее время достаточно подробно исследована экономическая интерпретация коэффициентов регрессии в различных уравнениях, содержащих различное количество факторов [4, с. 68—74]. Установлены причины числового несовпадения одноименных коэффициентов регрессии в различных моделях. Предложены понятия «условное, условно чистое и чистое влияние, или эффективность факторов». Коэффициенты регрессии,

соответствующие тому или другому содержательному понятию, правильно отражают действительность при наличии определенных содержательных предпосылок (условий).

Например, при изучении влияния удобрений на урожайность зерновых по данным 42 сельскохозяйственных предприятий Латвийской ССР были получены следующие результаты [5, с. 83]. При изучении парной связи

$$\tilde{x}_0 = 14,2 + 5,66x_1, \quad (1)$$

где x_0 — урожайность зерновых, ц/га; x_1 — количество внесенных удобрений, ц действующего вещества на 1 га посева.

Это означает, что каждый центнер удобрений при условии, что соответствующим образом будут изменены все другие факторы производства, дает увеличение урожайности зерновых в среднем на 5,66 ц/га (условное влияние).

При включении в анализ еще двух дополнительных факторов получаем [5, с. 162]

$$\tilde{x}_0 = -1,37 + 3,03x_1 + 4,67x_2 + 0,251x_3, \quad (2)$$

где x_2 — стоимость силовых и рабочих машин, используемых в полеводстве в сотнях рублей на 1 га; x_3 — качественная оценка пашни в баллах.

В новой постановке влияние удобрений на урожайность зерновых сократилась до 3,03 ц зерновых на 1 ц удобрений в действующем веществе. Это связано с тем, что зафиксированы на неизменном уровне два других фактора (стоимость машин и качество пашни). Тем самым степень условности этого показателя ниже. Он более конкретный. Его называем показателем условно чистого влияния фактора: он «очищен» от сопутствующего влияния двух других факторов. Но такая интерпретация еще не дает возможности указать, за счет взаимной корреляции каких факторов и насколько изменились коэффициенты регрессии при сопоставлении различных моделей.

Целью дальнейшего исследования является попытка количественного изучения различий парного и множественного коэффициентов регрессии, распределение их приращений по источникам возникновения.

Для удобства сначала рассматриваются стандартизованные коэффициенты регрессии, так называемые β -коэффициенты. Коэффициент парного линейного уравнения

регрессии в стандартизованном масштабе есть коэффициент корреляции.

Выше рассмотренные уравнения регрессии в стандартизованных единицах измерения имеют следующий вид:

$$\tilde{w}_0 = 0,633w_1 \quad (3)$$

и

$$\tilde{w}_0 = 0,338w_1 + 0,219w_2 + 0,499w_3, \quad (4)$$

где $w = x - \bar{x}/s_x$.

Таким образом, при рассмотрении условных связей при увеличении норм внесения удобрений на одно стандартное отклонение ожидается средний рост урожайности на 0,633 стандартных отклонения. При рассмотрении условно чистых связей увеличение норм внесения удобрений на одно стандартное отклонение обеспечивает средний рост урожайности только на 0,338 стандартных отклонений. Нам предстоит более подробно изучить разность этих показателей.

Для достижения поставленной цели необходимо изучить промежуточные результаты вычисления коэффициентов множественной регрессии. В этом случае наиболее удобным является метод обратной матрицы.

Коэффициенты множественного уравнения регрессии (β_{0j}) методом обратной матрицы вычисляются следующим образом [6, с. 1169—1182]:

$$(\beta_{0j}) = \|f_{ij}\| (r_{i0}), \quad (5)$$

где $\|f_{ij}\| = \|r_{ij}\|^{-1}$ — обратная корреляционная матрица ($r_{ij} = r_{ji}$); (r_{i0}) — вектор коэффициентов корреляции результативного признака (нулевого) со всеми факторными признаками ($r_{i0} = r_{0i}$).

Например, если у нас имеются три фактора ($j = 1, 2, 3$), то стандартизованный коэффициент регрессии для первого фактора $\beta_{01\ 23}$ равен:

$$\beta_{01\ 23} = r_{01}f_{11} + r_{02}f_{21} + r_{03}f_{31}. \quad (6)$$

Поскольку f_{11} (первый диагональный элемент обратной корреляционной матрицы) только в предельном случае может равняться единице, но, как правило, $f_{11} > 1$, то для более удобной интерпретации выражения (6) представим его в следующем виде:

$$\beta_{01\ 23} = r_{01} + r_{01}(f_{11} - 1) + r_{02}f_{21} + r_{03}f_{31}. \quad (7)$$

Стандартизованный коэффициент множественной регрессии $\beta_{01\ 23}$ выражен через сумму, первый из слагаемых

которой есть парный коэффициент корреляции. Решение выше поставленной задачи будет настолько успешным, насколько будет найдена убедительная интерпретация для остальных слагаемых выражения (7).

Выразив элементы обратной корреляционной матрицы через множественные и частные коэффициенты корреляции [6], выражение (7) представим следующим образом:

$$\begin{aligned} \beta_{01.23} = & r_{01} + r_{01} \frac{R_{1\ 23}^2}{1 - R_{1\ 23}^2} + \\ & + r_{02} \frac{r_{21.3}}{\sqrt{1 - R_{1\ 23}^2} \cdot \sqrt{1 - R_{2\ 13}^2}} + \\ & + r_{03} \frac{r_{31.2}}{\sqrt{1 - R_{1\ 23}^2} \cdot \sqrt{1 - R_{3\ 12}^2}}. \end{aligned} \quad (8)$$

Выражение (8) довольно сложное, чтобы сразу найти окончательную интерпретацию каждого слагаемого, поэтому наряду с обсуждением математического построения отдельных слагаемых используем результаты решения числового примера.

Система нормальных уравнений вышеприведенного примера, по которой можно вычислить параметры уравнения регрессии (4), следующая:

$$\begin{cases} \beta_1 + 0,34208\beta_2 + 0,44065\beta_3 = 0,63327; \\ 0,34208\beta_1 + \beta_2 + 0,18156\beta_3 = 0,42555; \\ 0,44065\beta_1 + 0,18156\beta_2 + \beta_3 = 0,68737. \end{cases} \quad (9)$$

Числовые константы в этой системе — коэффициенты парной корреляции. Они образуют так называемую корреляционную матрицу $\|r_{ij}\|$ и вектор свободных членов (r_{i0}). Обратная корреляционная матрица примера была найдена раньше [6]:

$$\|f_{ij}\| = \begin{vmatrix} 1,3609 & -0,36882 & -0,53271 \\ -0,36882 & 1,1340 & -0,043381 \\ -0,53271 & -0,043381 & 1,2426 \end{vmatrix}. \quad (10)$$

Используя эти данные, по формуле (5) вычисляем первый β -коэффициент, обращая внимание на отдельные слагаемые:

$$\begin{aligned} \beta_{01\ 23} = & 0,63327 + 0,63327 \cdot 0,3609 + 0,42555(-0,36882) + \\ & + 0,68737(-0,53271) = 0,63327 + 0,228555 - \\ & - 0,156951 - 0,366169. \end{aligned} \quad (11)$$

Первое слагаемое выражения (11) 0,63327, основа расчета, есть парный коэффициент корреляции. Из него вычитаются два числа — по количеству остальных факторов в модели. В первом приближении они могут быть интерпретированы следующим образом:

$$r_{02}f_{21} = r_{02} \frac{r_{21\ 3}}{\sqrt{1 - R_{1\ 23}^2} \cdot \sqrt{1 - R_{2\ 13}^2}} = -0,156951. \quad (12)$$

Это указывает, насколько условно чистая отдача удобрений ниже условной вследствие корреляции изучаемого фактора и сопутствующего фактора x_2 , с учетом всех косвенных связей. Этот показатель по абсолютной величине будет тем выше, чем выше частная корреляция x_1 и x_2 . Он повышается также по мере роста множественной корреляции рассматриваемых факторов с остальными; в этом случае знаменатель дроби уменьшается. Кроме того, разность условного и условно чистого влияния первого фактора будет больше, если результативный признак x_0 более тесно коррелирует с другим фактором x_2 , что отражает r_{02} .

Аналогично

$$r_{03}f_{31} = r_{03} \frac{r_{31\ 2}}{\sqrt{1 - R_{1\ 23}^2} \cdot \sqrt{1 - R_{3\ 12}^2}} = -0,366169 \quad (13)$$

указывает, насколько условно чистая эффективность удобрений ниже условной вследствие корреляции первого и третьего факторов (в стандартных отклонениях результативного признака на одно стандартное отклонение факторного) с учетом всех косвенных связей.

Общее сопутствующее влияние двух других факторов, которое было приписано удобрениям при парном анализе, может быть оценено как сумма только что рассмотренных величин. Обозначим ее буквой γ :

$$\begin{aligned} \gamma_{01\ 23} &= r_{02}f_{21} + r_{03}f_{31} = (-0,156951) + (-0,366169) = \\ &= -0,52312 \end{aligned} \quad (14)$$

(в стандартизованных единицах).

Суммируя (с учетом знака) величину (14) и парный коэффициент корреляции r_{01} , получаем ту часть β -коэффициента, которая характеризует независимую часть влияния удобрений на урожайность. В нее не включена перераспределенная часть совместного влияния комплек-

са факторов. Обозначим ее буквой λ

$$\lambda_{01\ 23} = r_{01} + \gamma_{01\ 23} = 0,63327 - 0,52312 = 0,11015. \quad (15)$$

Поскольку в линейной модели не предусмотрены члены, непосредственно выражающие совместное влияние факторов на результивный признак, методом наименьших квадратов это совместное влияние условно распределяется между факторами. Эта величина может быть оценена следующим слагаемым β -коэффициента, которую обозначим буквой ν :

$$\begin{aligned} \nu_{01\ 23} &= r_{01} (f_{11} - 1) = r_{01} \frac{R_{1\ 23}^2}{1 - R_{1\ 23}^2} = \\ &= 0,63327 \cdot 0,36090 = 0,22855. \end{aligned}$$

Таким образом, основные части разложения β -коэффициента следующие:

$$\beta_{01\ 23} = \lambda_{01\ 23} + \nu_{01\ 23} = 0,11015 + 0,22855 = 0,3387, \quad (16)$$

где λ — независимая часть β -коэффициента; ν — перераспределенная часть влияния комплекса факторов.

В рассматриваемом примере $\nu_{01\ 23} > \lambda_{01\ 23}$. Это свидетельствует о том, что показатель условно чистого влияния удобрений в рамках примера больше определяется перераспределенной частью совместного влияния факторов ν , нежели независимой частью λ . Удобрения, как и можно ожидать, свою эффективность проявляют в основном в комплексе с другим фактором производства.

Аналогично может быть проанализировано образование β -коэффициентов остальных факторов уравнения регрессии (4). Результаты удобно представить следующим образом.

Сопутствующие влияния, или смещения β -коэффициентов рассматриваемого примера, следующие:

$$\begin{aligned} \gamma_{01\ 23} &= r_{02}f_{12} + r_{03}f_{13} = -0,15695 - 0,36617 = -0,52312; \\ \gamma_{02\ 13} &= r_{01}f_{21} + r_{03}f_{23} = -0,23356 - 0,02982 = -0,26338; \end{aligned} \quad (17)$$

$$\gamma_{03\ 12} = r_{01}f_{31} + r_{02}f_{32} = -0,33735 - 0,01846 = -0,35581.$$

В рамках примера наиболее сильное смещение имеет β -коэффициент первого фактора (удобрений). Оно в основном вызвано корреляцией удобрений с качеством пашни:

$$|r_{03}f_{13} = -0,36617| > |r_{02}f_{12} = -0,15695|.$$

Независимые части β -коэффициентов:

$$\begin{aligned}\lambda_{01\ 23} &= r_{01} + \gamma_{01\ 23} = 0,63327 - 0,52312 = 0,11015; \\ \lambda_{02\ 13} &= r_{02} + \gamma_{02\ 13} = 0,42555 - 0,26338 = 0,16217; \\ \lambda_{03\ 12} &= r_{03} + \gamma_{03\ 12} = 0,68737 - 0,35581 = 0,33156.\end{aligned}\quad (18)$$

Наиболее значительную независимую часть имеет β -коэффициент третьего фактора (качество пашни). У остальных двух факторов она примерно равна и небольшая.

Перераспределенные влияния комплекса факторов следующие:

$$\begin{aligned}v_{01\ 23} &= r_{01} (f_{11} - 1) = 0,22855; \\ v_{02\ 13} &= r_{02} (f_{22} - 1) = 0,05702; \\ v_{03\ 12} &= r_{03} (f_{33} - 1) = 0,16676.\end{aligned}\quad (19)$$

Исходя из сложной корреляционной связи факторов, наибольшая часть совместного влияния факторов перераспределена на фактор x_1 (удобрение), значительно меньшая — на фактор x_3 (качество пашни), совсем малая — на фактор x_2 (машины).

Расчет самих β -коэффициентов может быть представлен так:

$$\begin{aligned}\beta_{01\ 23} &= \lambda_{01\ 23} + v_{01\ 23} = 0,11015 + 0,22855 = 0,3387; \\ \beta_{02\ 13} &= \lambda_{02\ 13} + v_{02\ 13} = 0,16217 + 0,05707 = 0,2192; \\ \beta_{03\ 12} &= \lambda_{03\ 12} + v_{03\ 12} = 0,33156 + 0,16676 = 0,4983.\end{aligned}\quad (20)$$

Разложение β -коэффициентов удобно представить в процентах:

$$\begin{aligned}\frac{\lambda_{01\ 23}}{\beta_{01\ 23}} + \frac{v_{01\ 23}}{\beta_{01\ 23}} &= 32,5\% + 67,5\%; \\ \frac{\lambda_{02\ 13}}{\beta_{02\ 13}} + \frac{v_{02\ 13}}{\beta_{02\ 13}} &= 74,0\% + 26,0\%; \\ \frac{\lambda_{03\ 12}}{\beta_{03\ 12}} + \frac{v_{03\ 12}}{\beta_{03\ 12}} &= 66,5\% + 33,5\%.\end{aligned}\quad (21)$$

Таким образом, β -коэффициенты второго и третьего факторов в основном состоят из независимой части влияния факторов. β -коэффициент первого фактора в основном образует перераспределенная часть влияния комплекса факторов.