МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
Учреждение образования
«Витебский государственный технологический университет»

Е. Ю. Вардомацкая

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ В ЛОГИСТИКЕ

Конспект лекций

для слушателей ФПК и ПК специальности
1-26 02 85 «Логистика»

Витебск
2020

УДК 004
ББК 32.97
В 18

Рецензенты:

кандидат физико-математических наук, доцент, заведующий кафедрой прикладного и системного программирования УО «Витебский государственный университет им. П.М. Машерова» Ермаченко С.А.;

главный бухгалтер ОАО «Витязь» Сванидзе Н.Н.

Рекомендовано к изданию редакционно-издательским советом УО «ВГТУ», протокол № 9 от 27.11.2019.

Вардомацкая, Е. Ю.
В 18 Экономико-математические методы и модели в логистике : конспект лекций / Е. Ю. Вардомацкая. – Витебск : УО «ВГТУ», 2020. – 76 с.
ISBN 978-985-481-631-9

Конспект лекций составлен в соответствии с действующей учебной программой дисциплины « Экономико-математические методы и модели в логистике» для слушателей специальности переподготовки руководящих работников и специалистов, имеющих высшее образование 1-26 02 85 «Логистика», и учитывает как требования к профессиональным компетенциям специалистов этого профиля, так и опыт преподавания данной дисциплины в УО «ВГТУ». В издании в определенной логической последовательности рассмотрены основные вопросы, обусловленные необходимостью практического использования экономико-математических методов и моделей в профессиональной деятельности специалистов по логистике. Значительное внимание уделено изучению возможностей использования пакетов прикладных программ для анализа и компьютерного моделирования логистических задач.

Адресуется слушателям переподготовки руководящих работников и специалистов, имеющих высшее образование, а также студентам экономических и технологических специальностей вузов, магистрантам и аспирантам, всем, интересующимся возможностью использования экономико-математических методов и моделей в профессиональной деятельности.

УДК 004
ББК 32.97

ISBN 978-985-481-631-9

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Целью изучения дисциплины «Экономико-математические методы и модели в логистике» является формирование у слушателей целостного и всестороннего представления об использовании экономико-математических методов и моделей при разработке и анализе эффективности логистических решений.

При изучении этого курса студенты должны получить теоретические знания по основам эконометрического моделирования, анализа и прогнозирования; изучить основные понятия и область применения экономико-математических методов и моделей в логистике.

Кроме теоретических знаний студенты должны получить и практические навыки в области применения современных концепций и компьютерных технологий построения логистических систем и цепей поставок; оценки, планирования, учета и анализа логистических затрат; оптимизации ресурсов в логистических системах на микро-, мезо- и макроуровнях и цепях поставок; моделирования логистических бизнес-процессов в цепях поставок; контроля результативности и эффективности логистических решений; поддержки логистических решений с помощью информационных систем и технологий.

Конспект лекций включает в себя как теоретический материал по ключевым темам дисциплины, так и практические примеры с пояснениями. Материал лекций четко структурирован, излишние подробности отсутствуют. Большое внимание уделяется возможностям использования компьютерных технологий для построения и анализа экономико-математических моделей, характеристике современных пакетов прикладных программ по эконометрике и экономико-математическому моделированию.

Для удобства работы и повышения степени усвоения теоретического материала к конспекту лекций прилагается диск с решением демонстрационных примеров по каждой теме. Примеры, записанные на диске, являются интерактивными, то есть пользователь имеет возможность адаптировать их под разный набор исходных данных и, как следствие, самостоятельно провести анализ полученных результатов.

Автор выражает благодарность студенткам группы Бу-25 Лысенковой Диане, Шаймардановой Лилии, Васеха Ольге, Семенченко Алине и Кузнечик Даше за помощь в отладке демонстрационных примеров и выражает надежду, что предлагаемый конспект лекций будет полезен не только студентам, но и магистрантам и аспирантам, изучающим эту дисциплину.

Лекция 1

ОСНОВНЫЕ ПОНЯТИЯ, ТЕОРЕТИЧЕСКИЕ ОСНОВЫ И ОБЛАСТЬ ПРИМЕНЕНИЯ ЭММ И М В ЛОГИСТИКЕ

План лекции

1.1 Предмет и содержание курса
1.2 Виды логистической деятельности
1.3 Понятие математической модели
1.4 Классификация экономико-математических моделей
1.5 Моделирование в логистике
1.6 Требования, предъявляемые к экономико-математическим моделям
1.7 Этапы построения экономико-математических моделей
1.8 Пакеты прикладных программ (ППП) для численного и символьного моделирования логистических задач

1.1 Предмет и содержание курса

Дисциплина «Экономико-математические методы и модели в логистике» (ЭММ и М) базируется на знаниях, полученных студентами при изучении таких дисциплин, как «Логистика», «Статистика», «Компьютерные информационные технологии».

Целью изучения дисциплины является формирование у слушателей целостного и всестороннего представления об использовании экономико-математических методов и моделей при разработке и анализе эффективности логистических решений.

Предмет изучения – логистическая деятельность предприятий и организаций.

Задачи изучения дисциплины

– применение современных концепций и компьютерных технологий построения логистических систем и цепей поставок;

– оценка, планирование, учет и анализ логистических затрат;

– оптимизация ресурсов в логистических системах на микро-, мезо- и макроуровнях и цепях поставок;

– моделирование логистических бизнес-процессов в цепях поставок;

– контроль результативности и эффективности логистических решений;

– поддержка логистических решений с помощью информационных систем и технологий.

В результате изучения учебной дисциплины студент должен:

знать:

– основы эконометрического моделирования, анализа и прогнозирования;

– основные понятия и область применения экономико-математических методов и моделей в логистике;

– современные пакеты прикладных программ по эконометрике и экономико-математическому моделированию;

уметь:

– использовать современные экономико-математические методы для прогнозирования, оптимального планирования и регулирования экономических процессов и явлений в логистике;

– уметь формулировать и ставить экономико-математические задачи в логистике;

– проводить идентификацию эконометрических моделей;

– моделировать экономические ситуации, связанные с оптимизацией исследуемых процессов;

– решать экономические задачи эконометрическими и оптимизационными методами;

владеть:

– основными приемами обработки статистических данных;

– методами аналитического и численного решения эконометрических и экономико-математических задач.

Таким образом, *основная задача этого курса* – систематизировать полученные ранее знания и дать студентам системные теоретические знания о применении математических методов и моделей в управлении логистической деятельностью предприятий и организаций любой формы собственности средствами компьютерных технологий.

Инструментарии изучения – пакеты прикладных программ общего и специального назначения (ТП MS Excel, СУП MS Project и некоторые другие).

1.2 Виды логистической деятельности

Логистика – это наука о *минимизации* затрат, человеческих и материальных ресурсов за счет оптимизации всех процессов. Профессиональная деятельность логиста направлена на выполнение только практических задач, поэтому каждая его ошибка может принести компании убыток.

Условно можно выделить следующие виды логистической деятельности:

1. *Военная логистика* – совместная и четко отлаженная работа военной промышленности, транспорта и тыловых служб. Именно военная логистика дала после окончания войны толчок к применению военного опыта в мирной экономике.

2. *Бизнес-логистика* (закупочная логистика, складская логистика, транспортная логистика, распределительная логистика, производственная логистика, таможенная логистика, логистика запасов, информационная логистика, комплексная логистика).

3. *Экологическая логистика*. Обеспечивает движение материала при любых производственных процессах вплоть до его превращения в товарный продукт и отходы с последующим проведением отходов до утилизации или безопасного хранения в окружающей среде.

4. *Бережливая логистика* (подробнее https://dic.academic.ru/dic.nsf/ruwiki/).

Несмотря на многозначность толкования, в русском языке термин *логистика* ассоциируется, прежде всего, с бизнес-логистикой. Охарактеризуем и рассмотрим подробнее те виды бизнес-логистики, которые чаще всего приходится обеспечивать специалисту-логисту.

Транспортная логистика – это система, по которой перевозчики осуществляют доставку товаров. Решение транспортных логистических задач условно можно разбить на такие этапы, как:

– составление маршрута;
– подбор транспортных средств;
– подбор рабочего персонала;
– финансовые расчёты и организация перевозки.

Несмотря на значительное развитие воздушного транспорта, более 40 процентов всего перевозимого товара все ещё транспортируется по воде. Просто это значительно дешевле.

Складская логистика. Обязанности складского логиста (товароведа) это:

– наблюдать за наличием товаров на складе и в случае необходимости сообщать о дефиците или избытке;
– следить за тем, чтобы продукция на складе имела товарный вид, не портилась, не слеживалась;
– организовывать поставки и отгрузки продукции, следить за правильностью выполнения процесса;
– вовремя проводить учёт.

Интересным является тот факт, что по оценкам независимых экспертов, товаровед – одна из профессий, в которых в ближайшие 20 лет компьютеры точно не заменят людей, настолько в этом виде деятельности важен именно человеческий фактор.

Ресурсная логистика. Логист-ресурсник – это логист-складовик, но на один уровень повыше. Обязанности логиста-ресурсника:

– составлять планы расходования материалов и производства товара;
– координация работы складов;

– распределение человеческих ресурсов.

По данным независимых исследований в настоящее время, логистика ресурсов входит в десятку самых стрессовых профессий.

Производственная логистика управляет материальным потоком во время прохождения им всех звеньев: от первичного источника до потребителя. Цель – снижение финансовых затрат.

Очевидно, что каждый из перечисленных видов логистической деятельности имеет специфические особенности и предназначен для решения определенных узконаправленных задач, практическая реализация которых невозможна без использования математико-логического аппарата или, иначе говоря, методов экономико-математического моделирования и компьютерных информационных технологий.

1.3 Понятие математической модели

Поведение любой системы (экономической, технологической, социальной и т. п.) зависит от множества факторов. Чем больше факторов, влияющих на поведение системы, выявлено, тем более точно (достоверно) можно регулировать и предсказывать ее поведение.

Процесс выявления факторов, влияющих на поведение системы, установления степени и характера взаимосвязи и влияния факторов друг на друга и на поведение системы в целом и есть *моделирование*. Таким образом, *модель* – это упрощенное описание поведения реальной системы.

Математическая модель – это система математических уравнений, неравенств, формул и различных математических выражений, описывающих реальный объект, составляющих его характеристики и взаимосвязи между ними.

Экономико-математическая модель (ЭММ) – математическое описание *экономического процесса* или явления, произведенное в целях его исследования.

Элементы модели – математические выражения, адекватно описывающие объект; *цель модели* – исследование и управление экономическим процессом.

Для построения и анализа любых видов математических и экономико-математических моделей нужно владеть определенными экономико-математическими методами. Термин *«экономико-математические методы»* понимается как обобщающее название комплекса экономических и математических научных дисциплин, объединенных для изучения социально-экономических систем и процессов. В этом смысле математические методы необходимы для обоснования математического обеспечения модели.

1.4 Классификация экономико-математических моделей

При классификации ЭММ могут использоваться разные признаки.

В основе классификации моделей *по содержательной проблеме* лежит объект моделирования. Например, односекторная комплексная макромодель описывает макроэкономические зависимости национальной экономики агрегированно, в то время как в многосекторной макромодели отражается движение потоков в разрезе секторов экономики (домашние хозяйства, сектор государственного управления, нефинансовые предприятия, банки, внешний мир) или в разрезе отраслей. Отраслевые модели и модели функциональных комплексов: модели промышленного производства, торговли, АПК, финансовые и другие – описывают особенности формирования показателей в указанных сферах. На уровне предприятий в зависимости от целей моделирования разрабатываются модели оптимизации производственной программы, транспортные модели, модели оптимизации раскроя материала, модели спроса и т. д.

При классификации моделей *по периодам прогнозирования* выделяют:

- *краткосрочные* (период прогнозирования до одного года);
- *среднесрочные* (до пяти лет);
- *долгосрочные* (свыше пяти лет) модели.

Как правило, построение данных моделей имеет различную методологическую базу. Например, при создании макромодели белорусской экономики на период прогнозирования до двух лет описано влияние показателей кредитно-денежной и налогово-бюджетной политики, в то время как алгоритм работы макромодели на период прогнозирования свыше двух лет строится в зависимости от динамики инвестиций: кредитно-денежный блок здесь имеет подчиненное значение. Более того, в краткосрочной перспективе влияние инфляционных процессов на динамику макропоказателей может оказать стимулирующее воздействие, в то время как в долгосрочной такое влияние почти всегда отрицательно. Модели должны адекватно отражать подобные процессы. При этом очень важным представляется согласование выявленных в исследуемой экономике зависимостей и положений экономической теории. Если обнаружено противоречие, то такую зависимость опасно закладывать в модель, которая строится для целей прогнозирования: имеет смысл рассмотреть связь прогнозируемого показателя с другими факторами.

По возможности учета *фактора неопределенности экономического процесса* различают модели:

- *детерминированные*, когда входные параметры задаются однозначно и выходные показатели определяются соответственно;
- *стохастические* – параметры модели, условия функционирования и характеристики объекта выражены случайными величинами и связаны стохастическими зависимостями, либо исходная информация также представлена случайными величинами. Последний класс моделей более

адекватно описывает экономические процессы, но при практическом использовании этих моделей возникают трудности информационного характера, поскольку исходы стратегий в различных экономических условиях оцениваются, как правило, экспертно.

По возможности учета *временных изменений* различают:

– *динамические модели* – модели, описывающие экономику в развитии. Как правило, в них имеется переменная, обеспечивающая связь последующего и предыдущего периодов, например инвестиции в основной капитал рассматриваются как функция прироста производства за период t_i;

– *статические модели* – это ЭММ, в которой все зависимости отнесены к одному моменту времени. Статические модели разрабатываются лишь для отдельно взятых периодов, в данном случае для t-1 или t, а развитие экономического процесса отображается рассчитанными показателями для периодов t-1 и t.

Несмотря на привлекательную особенность динамических моделей, их структура значительно сложнее статических, что не только усложняет информационное обеспечение модели, но и затрудняет экономическую интерпретацию расчетов, повышает степень неустойчивости модельных результатов.

По степени агрегирования объектов моделирования модели подразделяются на *макроэкономические* и *микроэкономические*. К первым относят модели, отражающие функционирование экономики как единого целого, в то время как микроэкономические модели связаны с такими звеньями экономики, как предприятия и фирмы.

По типу подхода к изучаемым социально-экономическим системам выделяют:

– *дескриптивные модели*, предназначенные для описания и объяснения фактически наблюдаемых явлений или для прогноза этих явлений. В качестве примера дескриптивных моделей можно привести балансовые, имитационные, эконометрические модели;

– *нормативные модели*, устанавливающие не то, каким образом устроена и развивается экономическая система, а как она должна быть устроена и как должна действовать при определенных критериях. В частности, все оптимизационные модели относятся к классу нормативных.

Наиболее распространенной является классификация *по используемому математическому аппарату*. При этом выделяют следующие классы моделей:

– *имитационные;*

– *эконометрические;*

– *оптимизационные;*

– *балансовые;*

– *модели исследования операций* (теория игр, модели теории массового обслуживания, управления запасами, сетевые модели и др.).

На практике часто используются комбинации этих моделей.

1.5 Экономико-математическое моделирование в логистике

Логистические потоковые процессы в форме системы товародвижения на практике образуют следующие блоки:

- закупки (снабжение);
- сбыт (продажи);
- перемещение (транспортировка);
- складирование (запасы).

Каждое предприятие, в силу универсальности логистики, в той или иной мере выполняет указанные блоки в своей производительно-коммерческой деятельности, вследствие чего эти блоки увязываются в единую систему с помощью управления (см. рис. 1.1).

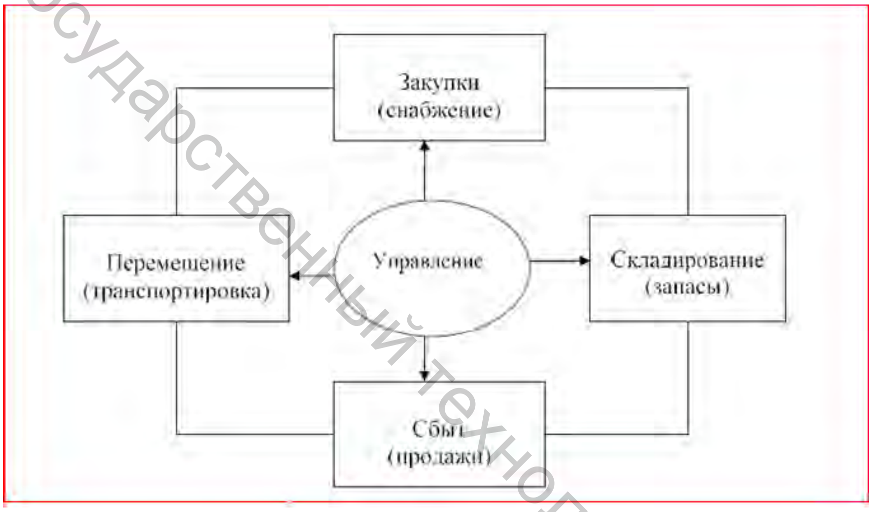


Рисунок 1.1 – Логистический функциональный блок

Среди экономико-математических методов, используемых для обоснования решений в области логистики, выделяют:

- функционально-стоимостной анализ;
- линейное программирование;
- сетевое планирование и управление;
- теорию игр;
- теорию вероятностей;
- теорию массового обслуживания;
- стохастическое (имитационное) моделирование;
- эвристические методы, экспертные методы.

Функционально-стоимостной анализ используется для комплексного решения задач, связанных с повышением качества продукции и одновременной экономией материальных и трудовых ресурсов предприятия.

Линейное программирование используется для выбора из ряда альтернативных решений наиболее благоприятного. В качестве критериев выбора могут служить минимальные издержки производства, максимальная прибыль, наименьшие затраты времени и (или) трудозатраты.

С помощью этого метода могут решаться следующие задачи:

1. Укрупненное планирование производства, т. е. составление графиков производства, целью которых является минимизация общих производственных издержек при заданных ограничениях по трудовым ресурсам и определенном уровне запасов.

2. Планирование ассортимента продукции, т. е. определение оптимального ее ассортимента, в котором каждому ее виду присущи свои издержки и потребности в ресурсах.

3. Управление технологическим процессом, т. е. сведение к минимуму отходов материалов, используемых при производстве продукции.

4. Регулирование запасов, т. е. определение рационального соотношения на складе готовой к отгрузке продукции, сырья и комплектующих.

5. Распределение рабочих по станкам и рабочим местам, ставящее целью минимизацию потерь рабочего времени и простоев оборудования и др.

Метод сетевого планирования и управления позволяет регулировать последовательность и взаимозависимость отдельных видов работ или операций в рамках какой-либо программы. Он позволяет четко фиксировать основные этапы работы, определять сроки их выполнения, разграничивать ответственность, экономить ресурсы, снижать затраты, предусматривать возможные мероприятия, снижающие риск. Достаточно эффективным является использование метода сетевого планирования при разработке программы производства нового товара и организации пробных продаж, подготовке и проведении сбытовых и рекламных кампаний.

Метод теории игр позволяет выработать решение в ситуациях, когда сталкиваются интересы двух или более сторон. В бизнесе теория игр используется для прогнозирования реакции конкурентов и потребителей на изменения цен, открытие новых рынков сбыта, предложение дополнительных услуг, модификацию и освоение новой продукции и т. д. Теория игр используется реже других аналитико-прогностических методов, так как ситуации реального мира зачастую очень сложны и настолько быстро изменяются, что невозможно точно спрогнозировать, как на них отреагируют субъекты рынка.

Методы теории вероятностей помогают принимать решения, которые сводятся к определению значения вероятностей наступления определенных событий и выбору из возможных действий наиболее предпочтительного. Например, производить базовый или модернизированный продукт, реорганизовывать или расширять производство, внедряться на новый географический рынок или нет.

Теория массового обслуживания – отрасль прикладной математики, использующая методы теории вероятностей, позволяющая изучить складывающиеся закономерности, связанные с потоком заявок на обслуживание и соблюсти необходимую очередность их выполнения. Ее предметом является количественная оценка качественной стороны процессов обслуживания клиентов. Теория массового обслуживания устанавливает зависимость между характером потока заявок, производительностью конкретного оборудования (поста обслуживания) и их числом, а так же эффективностью обслуживания. В ходе расчетов определяется оптимальное число каналов обслуживания по отношению к потребностям в них. В качестве критериев эффективности могут использоваться различные величины и функции (например: вероятность обслуживания каждой поступающей заявки; среднее число обслуживаемых заявок; среднее время простоя каналов обслуживания; пропускная способность системы; время ожидания заявок в очереди; совокупная стоимость простоя клиента и канала обслуживания и т. д.). Применение методов теории массового обслуживания позволяет установить, какие результаты могут быть достигнуты при использовании тех или иных форм организации процесса обслуживания клиентов. Различают следующие виды систем массового обслуживания: системы с потерями и без потерь; с ограниченным и неограниченным числом постов обслуживания; с ограниченным и неограниченным потоком клиентов; упорядоченные и неупорядоченные системы обслуживания. Правильный выбор параметров таких систем позволит избежать многих узких мест в производственных процессах предприятий и неполной загрузки отдельных его структурных подразделений, а так же обеспечит значительную экономию трудовых и материальных ресурсов.

К ситуациям, в которых эффективно использовать теорию массового обслуживания, можно отнести:

1. Проектирование телефонных линий для оказания информационных услуг (например, для справочного бюро или телефонной справки).
2. Организация процесса обслуживания клиентов на предприятии, в учреждении, в организации.
3. Оптимизация очереди транспортных средств под погрузку-разгрузку.
4. Расчет необходимой численности основных и вспомогательных рабочих предприятия и т. д.

Стохастическое (имитационное) моделирование – это способ изучения процессов или явлений реального мира на специально построенных моделях. Часто в качестве моделей применяются экономико-математические, а их реализация и изучение производится с помощью компьютерных технологий, т. е. виртуально. Экспериментируя на модели рассматриваемой системы, можно установить, как она будет реагировать на определенные изменения или события, в то время когда отсутствует возможность наблюдать эту систему в реальности.

Например, аэродинамическая труба – физическая модель процесса, используемая для проверки характеристик разрабатываемых самолетов и автомобилей.

Аналитики в области *производственно-финансовой деятельности* создают модели, позволяющие имитировать ожидаемый прирост производительности труда и прибылей в результате применения новых технологий и модернизации производства.

Аналитики в области *маркетинга* создают модели, имитирующие изменения ожидаемых объемов сбыта в связи с проведением рекламных кампаний, ростом или снижением цен на продукцию, изменением условий поставок сырья и материалов.

Эвристические методы (от греч. *heurisko* – отыскиваю, открываю) – специальные методы, базирующиеся на продуктивном творческом мышлении, используемые в процессе открытия нового. К основным эвристическим методам относят: метод фокусного объекта, метод «гирлянд», метод музейного экспоната, метод ТРИЗ, метод морфологического ящика.

Экспертные методы представляют собой процедуры принятия решений, основанные на использовании творческого потенциала специалистов. Как правило, специалисты с большим опытом (стажем) работы обладают собственными эвристико-интуитивными алгоритмами, накопленными в результате многолетней практической деятельности. И поэтому методы экспертных оценок позволяют достаточно быстро получить ответ о возможных процессах развития того или иного события на рынке, выявить сильные и слабые стороны предприятия, получить оценку эффективности тех или иных управленческих решений или мероприятий.

Применение экспертных методов в маркетинго-логистической деятельности целесообразно в силу объективных причин, позволяющих:

– учесть условия неопределенности исходных данных;
– рассмотреть сложную природу взаимосвязей процессов;
– сократить объем расчетных операций;
– учесть динамичность изменения данных;
– преодолеть недостаток необходимых знаний или времени, исключающих возможность проведения более детального анализа.

Правильное проведение экспертизы предполагает решение ряда вопросов, связанных с формированием экспертной группы, проведением процедуры экспертизы, выбором методов обработки результатов экспертных оценок. Основными требованиями к экспертам являются компетентность, профессионализм, авторитетность, беспристрастность. Процедура экспертизы предполагает коллективное генерирование идей на основе проведения дискуссий или опросов с применением анкет. Основными методами экспертных оценок являются ранговых корреляций, парных предпочтений, «Дельфи», «мозговой атаки».

В зависимости от направления и сути решаемых задач *при проведении логистического анализа* можно выделить следующие виды моделей:

– *эконометрические (регрессионные)*. В основе расчета таких моделей лежит теория вероятности и математическая статистика;
– *балансовые* – методы матричного исчисления;
– *финансовые* – методы финансового анализа;

- *оптимизационные* – методы математического программирования;
- *модели исследования операций* – теория вероятности;
- *имитационные* – теория вероятности и математическая статистика.

На практике, как уже говорилось выше, часто используются комбинации этих моделей: имитационные эконометрические, имитационные оптимизационные, имитационные балансовые и т. д.

Следует понимать, что любая модель всегда будет отличаться от реального поведения рассматриваемой системы. Это может быть связано:

- со спецификацией (видом) модели;
- с невозможностью учета всех влияющих факторов;
- с вмешательством случайных величин (рисков);
- с влиянием человеческого фактора.

В общем виде любая модель представляет собой зависимость вида

$$\mathbf{Y = f(x)},$$

где *x – набор независимых параметров*, оказывающих влияние на поведение модели; Y – выходной параметр (*зависимая переменная*), изменяющая свое значение при изменении зависимых параметров.

1.6 Требования, предъявляемые к экономико-математическим моделям

ЭММ модель должна:

- отвечать *целям* моделирования (*целевая функция*);
- отвечать *условиям* моделирования (*ограничения*);
- базироваться на достаточной и объективной информационной базе, включающей только значимые факторы;
- минимизировать участие человеческого фактора при исследовании;
- отвечать условиям информационно-целевой адекватности;
- быть эффективно реализуемой, то есть основываться на универсальных и эффективных методах решения.

1.7 Этапы построения экономико-математических моделей

Построение модели представляет собой итеративную процедуру, включающую следующие основные этапы: *экспериментальные модельные расчеты* и *оценка адекватности модели*. Более точно в качестве этапов экономико-математического моделирования можно выделить следующие:

1. *Постановка задачи.* Экономическое обоснование и методологическое и методическое обоснование модели предполагает изучение особенностей объекта моделирования и их отражение с помощью структуры разрабатываемой модели.

2. *Выбор математического аппарата и математический анализ модели.* Это этап формализации экономической проблемы, т. е. описания модели в виде конкретных математических зависимостей (функций, уравнений, неравенств и др.). Необходимо стремиться построить модель, относящуюся к хорошо изученному классу математических задач, что может потребовать некоторого упрощения исходных показателей модели. Однако, важно следить за тем, чтобы при этом не искажались основные черты моделируемого объекта. Возможна и такая ситуация, когда формализация проблемы приводит к неизвестной ранее математической структуре и ее практическая реализация может потребовать значительных затрат или быть неосуществимой.

3. *Информационное обеспечение модели.* Это, как правило, наиболее трудоемкий этап моделирования, так как требуется знание имеющейся статистической отчетности, сопоставимости отчетных данных во времени и по предприятиям. При этом необходимо принимать во внимание не только принципиальную возможность подготовки информации требуемого качества, но и затраты на подготовку информационных массивов.

4. *Программное обеспечение модели.* Этап включает подготовку программ на ПЭВМ с использованием стандартных ППП.

5. *Экспериментальные модельные расчеты и оценка адекватности модели.* На этом этапе проводятся многочисленные модельные эксперименты, изучается поведение модели при различных условиях и на этой основе оценивается адекватность модели. Оценка адекватности модели предполагает оценку соответствия модели моделируемому объекту или процессу. Адекватность – понятие условное, так как полного соответствия модели реальному объекту быть не может. Тем не менее существует ряд рассчитываемых статистических характеристик, которые позволяют оценить качество модели и рассчитать предполагаемые ошибки прогноза и модельные реакции.

Перечисленные этапы экономико-математического моделирования находятся в тесной связи. В частности, имеют место возвратные связи этапов. Так, на этапе построения модели может выявиться, что постановка задачи приводит к слишком сложной математической форме модели. В этом случае исходная постановка задачи должна быть скорректирована.

1.8 Пакеты прикладных программ (ППП) для моделирования логистических задач

Программная поддержка процессов моделирования, анализа и управления может осуществляться с использованием самых разнообразных программных средств. Назовем и охарактеризуем некоторые из них:

1. *Универсальные и специализированные языки программирования*.

2. *Стандартные офисные программные продукты*: Microsoft Office, Star Office, Lotus, Open Office и т. д.

3. *Системы компьютерной математики:* Gauss 6.0; Maple 10.03; Mathematica 5.2.0; MathLab 6.5; MuPad 4.0; SciLab 4.0; Maxima 5.9.1 и т. д.

4. *Системы управления проектами.* Primavera Project Planner (P3) (Primavera); Microsoft Project (Microsoft); Time Line (Time Line Solutions); Open Plan (Welcome Software); Artemis Views (Artemis Management Systems); CA-Super Project (Computer Associates International Inc.); Project Scheduler (Scitor Corp.); TurboProject (IMSI); Project Workbench (Applied Business Technology); Spider Project (Технологии управления Спайдер); PlanBee; Rillsoft Project и т. д.

5. *CASE-технологии*. К настоящему моменту наиболее интенсивное развитие получили два главных направления применения CASE-средств: *реорганизация* (перепроектирование) бизнес-процессов организации; *системный анализ и проектирование*, включающий функциональное, информационное и событийное моделирование как вновь создаваемой, так и существующей системы.

6. *Специализированные статистические пакеты*: SAS for Windows (SAS Institute Inc.); SPSS (SPSS Inc.); S-Plus (Mathworks) ; Systat (SPSS Inc.) ; NCSS (NCSS); STATA (Stata corp.); Statistica (Statsoft Inc.); Statgraphics Plus (Manguistics, Inc).

Лекция 2

РЕГРЕССИОННЫЙ АНАЛИЗ. СПЕЦИФИКАЦИЯ ЭКОНОМЕТРИЧЕСКОЙ МОДЕЛИ

План лекции

2.1 Особенности регрессионных моделей
2.2 Математическое обеспечение регрессионных моделей: общие статистические сведения
2.3 Модель парной и множественной регрессии
2.4 Возможности пакетов прикладных программ (ППП) для проведения регрессионного анализа

2.1 Особенности регрессионных моделей

Для решения задач экономического анализа и прогнозирования очень часто используются статистические, отчетные или наблюдаемые данные. При этом полагают, эти данные являются значениями случайной величины.

Случайной величиной называется переменная величина, которая в зависимости от случая принимает различные значения с некоторой вероятностью. Закон распределения случайной величины показывает частоту ее тех или иных значений в общей их совокупности.

При исследовании взаимосвязей между экономическими показателями на основе статистических данных, часто между ними наблюдается ***стохастическая зависимость.*** Она проявляется в том, что изменение закона распределения одной случайной величины происходит под влиянием изменения другой. Взаимосвязь между величинами может быть полной (функциональной) и неполной (искаженной другими факторами).

Пример функциональной зависимости – выпуск продукции и ее потребление в условиях дефицита.

Неполная зависимость наблюдается, например, между стажем рабочих и их производительностью труда. Обычно рабочие с большим стажем работы работают лучше молодых, но под влиянием дополнительных факторов – образование, здоровье и т. д. эта зависимость может быть искажена.

Раздел математической статистики, посвященный изучению взаимосвязей между случайными величинами, называется *корреляционным анализом*. Основная задача корреляционного анализа – это установление характера и тесноты связи между результативными (зависимыми) и факторными (независимыми) показателями (признаками) в данном явлении или процессе. Корреляционную связь можно обнаружить только при массовом сопоставлении фактов.

Характер связи между показателями определяется по корреляционному полю. Если Y – зависимый признак, а x – независимый, то, отметив каждый случай $X(i)$ с координатами x_i и y_i, получим корреляционное поле.

Теснота связи определяется с помощью коэффициента корреляции, который рассчитывается специальным образом и лежит в интервалах от минус единицы до плюс единицы. Если значение коэффициента корреляции лежит в интервале от *1* до *0,9* по модулю, то отмечается очень сильная корреляционная зависимость. В случае если значение коэффициента корреляции лежит в интервале от *0,9* до *0,6,* то говорят, что имеет место слабая корреляционная зависимость. Наконец, если значение коэффициента корреляции находится в интервале от *–0,6* до *0,6,* то говорят об очень слабой корреляционной зависимости или полной ее отсутствии.

Таким образом, корреляционный анализ применяется для нахождения характера и тесноты связи между случайными величинами.

Регрессионный анализ своей целью имеет вывод, определение (идентификацию) уравнения регрессии вида $Y = f(x)$, включая статистическую оценку его параметров. *Уравнение регрессии* позволяет найти значение зависимой переменной, если величина независимой или независимых переменных известна.

Практически речь идет о том, чтобы, анализируя множество точек на графике (т. е. множество статистических данных), найти линию, по возможности, точно отражающую заключенную в этом множестве закономерность (тренд, тенденцию) – линию регрессии.

По числу факторов различают одно-, двух- и многофакторные уравнения регрессии.

По характеру связи однофакторные уравнения регрессии подразделяются на:

а) *линейные:*

$$Y = a + bx\,;$$

где x – экзогенная (независимая) переменная; Y – эндогенная (зависимая, результативная) переменная; a, b – параметры уравнения регрессии;

б) *степенные*:

$$Y = a \cdot x^{b}\,;$$

в) *показательные:*

$$Y = a \cdot b^{x}\,;$$

г) *прочие*.

Обычно, для упрощения процедуры расчета и анализа, уравнение регрессии стараются привести к *линейному* виду.

2.2 Математическое обеспечение регрессионных моделей: общие статистические сведения

Математическое обеспечение расчета и анализа регрессионных моделей составляют методы математической статистики. Наиболее прост в использовании метод наименьших квадратов, который обеспечивает расчет и анализ уравнения регрессии линейного вида.

Пусть:

$x_1, x_2, \ldots, x_n$ – совокупность значений независимого, факторного признака;

$y_1, y_2 \ldots, y_n$ – совокупность соответствующих значений зависимого, результативного признака;

n – количество наблюдений.

Требуется рассчитать уравнение регрессии и оценить его адекватность, то есть соответствие определенным требованиям. Только в этом случае уравнение можно будет использовать для анализа экономических процессов и выполнения прогнозов.

Для того чтобы найти уравнение регрессии, т. е. зависимость вида Y = f(x), прежде всего, нужно исследовать тесноту связи между случайными величинами **x** и **Y**, т. е. корреляционную зависимость.

Для этого вычисляются следующие величины:

1. *Средние значения*

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$ для независимой переменной.

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$ для зависимой переменной.

2. *Отклонения от средних величин*

$$\Delta x_i = x_i - \bar{x}, \qquad \Delta y_i = y_i - \bar{y}\$$$

3. *Величины дисперсии и среднего квадратичного отклонения*

$$D_x = \frac{\sum_{i=1}^{n} \Delta x_i^2}{n - 1}, \quad D_y = \frac{\sum_{i=1}^{n} \Delta y_i^2}{n - 1}.$$

$$\sigma_x = \sqrt{D_x} \quad \sigma_y = \sqrt{D_y}$$

Величины дисперсии и среднего квадратичного отклонения характеризуют разброс наблюдаемых значений вокруг среднего значения. Чем больше дисперсия, тем больше разброс.

4. *Корреляционный момент (коэффициент ковариации):*

$$K_{x,y} = \frac{\Delta x_1 \cdot \Delta y_1 + \Delta x_2 \cdot \Delta y_2 + ... + \Delta x_n \cdot \Delta y_n}{n - 1} = \frac{\sum_{i=1}^{n} \Delta x_i \cdot \Delta y_i}{n - 1}.$$

Корреляционный момент отражает характер взаимосвязи между **x** и **y**. Если $K_{xy} > 0$, то взаимосвязь прямая. Если $K_{xy} < 0$, то взаимосвязь обратная.

5. *Коэффициент корреляции* **(**вычисляется по формуле)

$$R_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}.$$

Доказано, что коэффициент корреляции находится в интервале от минус единицы до плюс единицы ($-1 \le R_{xy} \le 1$). Коэффициент корреляции в квадрате (R_{xy}^2) называется коэффициентом детерминации. Если $R_{xy} \ge |0.8|$, то вычисления продолжаются.

6. *Вычисление параметров регрессионного уравнения.*

Коэффициент *b* находится по формуле

$$b = \frac{K_{xy}}{D_x};$$

После чего можно легко найти параметр *a:* $a = \overline{y} - b\overline{x}$

Коэффициенты *a* и *b* находятся *методом наименьших квадратов*, основная идея которого состоит в том, что за меру суммарной погрешности принимается сумма квадратов разности (остатков) между фактическими значениями результативного признака $\mathbf{y}_i$ и его расчетными значениями y_{ip}, полученными при помощи уравнения регрессии

$$y_{ip} = a + bx_i.$$

При этом величины остатков находятся по формуле

$$u_i = y_i - y_{ip},$$

где y_i – фактическое значение **y**; y_{ip} – расчетное значение **y**.

7. *Оценка величины погрешности линейного однофакторного уравнения*

Если разность между фактическим значением результативного признака и его расчетным значением обозначить как u_i:

$$\mathbf{u_i = y_i - y_{ip}},$$

где y_i – фактическое значение **y**; $\mathbf{y_{ip}}$ – расчетное значение **y**; u_i – разность между ними.

Тогда *суммарная погрешность* уравнения регрессии может быть определена по формуле

$$S = \frac{\sum_{n-2}^{n} u_i^2}{n-2}.$$

Поскольку $\bar{u}$ (среднее значение остатков) равно нулю, то суммарная погрешность равна остаточной дисперсии

$$D_u = \frac{\sum (u_i - \bar{u})^2}{n-2} = \frac{\sum u_i^2}{n-2} = S.$$

В экономических приложениях допустимая суммарная погрешность может составить не более 20 % от дисперсии результативного признака D_y.

8. *Стандартная ошибка* уравнения находится по формуле

$$\sigma_u = \sqrt{D_u},$$

где D_u – остаточная дисперсия.

9. *Относительная погрешность* уравнения регрессии вычисляется как

$$\vartheta = \frac{\sigma_u}{\bar{y}} \cdot 100\,\%,$$

где σ_u – стандартная ошибка; $\bar{y}$ – среднее значение результативного признака.

Если величина ϑ мала и отсутствует автокорреляция остатков, то прогнозные качества оцененного регрессионного уравнения высоки.

10. *Стандартная ошибка коэффициента b* вычисляется по формуле

$$S_b = \frac{\sigma_u}{\sqrt{nD_x}}$$

11. Для вычисления *стандартной ошибки коэффициента а* используется формула

$$S_a = \sigma_u \sqrt{\frac{D_x + \bar{x}^2}{n \cdot D_x}}$$

Стандартные ошибки коэффициентов используются для оценивания параметров уравнения регрессии.

Коэффициенты считаются значимыми, если $\frac{S_a}{|a|} < 0.5$; $\frac{S_b}{|b|} < 0.5$

Стандартные ошибки коэффициентов используются также для оценки статистической значимости коэффициентов при помощи *t – критерия Стьюдента.* Значения *t – критерия Стьюдента* содержатся в справочниках по математической статистике.

Если полученные результаты не являются значимыми и не могут быть использованы для прогнозных расчетов, то ситуацию можно поправить следующими способами:

а) увеличить число *n;*

б) увеличить количество факторов;

в) изменить форму уравнения.

2.3 Модель парной и множественной регрессии. Парная регрессия и корреляция

Парная регрессия представляет собой регрессию между двумя переменными – y и x, т. е. модель вида

$$\hat{y} = f(x),$$

где y – зависимая переменная (результативный признак); x – независимая, или объясняющая, переменная (признак-фактор). Знак «^» означает, что между переменными x и y нет строгой функциональной зависимости, поэтому практически в каждом отдельном случае величина y складывается из двух слагаемых

$$y = \hat{y}_x + \varepsilon,$$

где y – фактическое значение результативного признака; $\hat{y}_x$ – теоретическое значение результативного признака, найденное исходя из уравнения регрессии;

ε – случайная величина, характеризующая отклонения реального значения результативного признака от теоретического, найденного по уравнению регрессии.

Случайная величина ε называется также возмущением. Она включает влияние не учтенных в модели факторов, случайных ошибок и особенностей измерения. Ее присутствие в модели порождено тремя источниками: *спецификацией модели (видом функции f), выборочным характером исходных данных, особенностями измерения переменных*.

От правильно выбранной спецификации модели зависит величина случайных ошибок: они тем меньше, чем в большей мере теоретические значения результативного признака $\hat{y}_x$, подходят к фактическим данным y.

К ошибкам спецификации относятся

– неправильный выбор той или иной математической функции для $\hat{y}_x$;

– недоучет в уравнении регрессии какого-либо существенного фактора, т. е. использование парной регрессии вместо множественной.

Наряду с ошибками спецификации могут быть *ошибки выборки*, которые имеют место в силу неоднородности данных в исходной статистической совокупности, что, как правило, бывает при изучении экономических процессов. Если совокупность неоднородна, то уравнение регрессии не имеет практического смысла. Для получения хорошего результата обычно исключают из совокупности единицы с аномальными значениями исследуемых признаков. И в этом случае результаты регрессии представляют собой выборочные характеристики.

Использование временной информации также представляет собой выборку из всего множества хронологических дат. Изменив временной интервал, можно получить другие результаты регрессии.

Наибольшую опасность в практическом использовании методов регрессии представляют *ошибки измерения.* Если ошибки спецификации можно уменьшить, изменяя форму модели (вид математической формулы), а ошибки выборки – увеличивая объем исходных данных, то ошибки измерения практически сводят на нет все усилия по количественной оценке связи между признаками.

Особенно велика роль ошибок измерения при исследовании на макроуровне. Так, в исследованиях спроса и потребления в качестве объясняющей переменной широко используется «доход на душу населения». Вместе с тем статистическое измерение величины дохода сопряжено с рядом трудностей и не лишено возможных ошибок, например, в результате наличия скрытых доходов.

Предполагая, что ошибки измерения сведены к минимуму, основное внимание в эконометрических исследованиях уделяется ошибкам спецификации модели.

В парной регрессии выбор вида математической функции $\hat{y}_x = f(x)$ может быть осуществлен тремя методами:

1) *графическим*;
2) *аналитическим*, т. е. исходя из теории изучаемой взаимосвязи;
3) *экспериментальным*.

При изучении зависимости между двумя признаками графический метод подбора вида уравнения регрессии достаточно нагляден. Он основан на поле корреляции. Основные типы кривых, используемые при количественной оценке связей, представлены на рисунке 2.1:

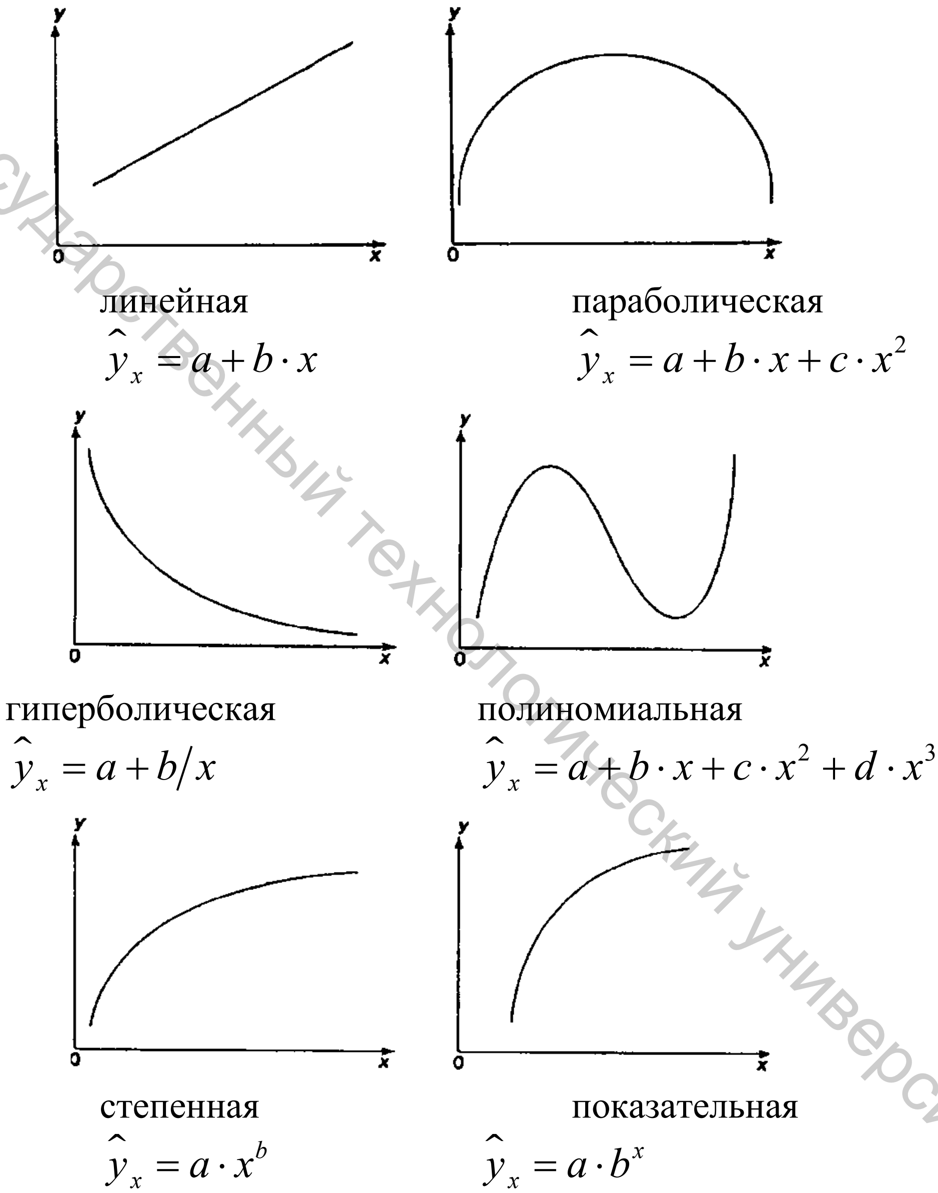


Рисунок 2.1 – Основные типы кривых, используемые при количественной оценке связей между двумя переменными

Если уравнение регрессии проходит через все точки корреляционного поля, что возможно только при функциональной связи, когда все точки лежат на линии регрессии $\hat{y}_x = f(x)$, то фактические значения результативного признака совпадают с теоретическими $y = \hat{y}_x$, т. е. они полностью обусловлены влиянием фактора x. В этом случае остаточная дисперсия

$$\sigma^2_{\text{ост}} = 0.$$

В практических исследованиях, как правило, имеет место некоторое рассеяние экспериментальных точек относительно линии регрессии. Оно обусловлено влиянием прочих, не учитываемых в уравнении регрессии, факторов. Иными словами, имеют место отклонения фактических данных от теоретических $\left(y - \hat{y}_x\right)$. Величина этих отклонений и лежит в основе расчета остаточной дисперсии

$$\sigma^2_{\text{ост}} = \frac{1}{n}\sum\left(y - \hat{y}_x\right)^2 \rightarrow \min.$$

Чем меньше величина остаточной дисперсии, тем меньше влияние не учитываемых в уравнении регрессии факторов и тем лучше уравнение регрессии подходит к исходным данным.

Считается, что число наблюдений должно в 7–8 раз превышать число рассчитываемых параметров при переменной x. Это означает, что искать линейную регрессию, имея менее 7 наблюдений, вообще не имеет смысла. Если вид функции усложняется, то требуется увеличение объема наблюдений, ибо каждый параметр при x должен рассчитываться хотя бы по 7 наблюдениям. Значит, если вид уравнения регрессии представлен параболой второй степени $\hat{y}_x = a + b \cdot x + c \cdot x^2$, то требуемый объем информации составит не менее 14 наблюдений.

Линейная модель парной регрессии и корреляции

Линейная регрессия представляет собой простейшую модель парной регрессии и находит широкое применение ввиду четкой экономической интерпретации ее параметров.

Линейная регрессия сводится к нахождению уравнения вида

$$\hat{y}_x = a + b \cdot x \text{ или } y = a + b \cdot x + \varepsilon. \quad (2.1)$$

Уравнение вида $\hat{y}_x = a + b \cdot x$ позволяет по заданным значениям фактора x находить теоретические значения результативного признака, подставляя в него фактические значения фактора x.

Построение линейной регрессии сводится к оценке ее параметров – a и b. Классический подход к оцениванию параметров линейной регрессии основан на методе наименьших квадратов (МНК). МНК позволяет получить такие оценки параметров a и b, при которых сумма квадратов отклонений фактических значений результативного признака y от теоретических $\hat{y}_x$ минимальна

$$\sum_{i=1}^{n}\left(y_i - \hat{y}_{x_i}\right)^2 = \sum_{i=1}^{n}\varepsilon_i^2 \quad \to \quad \min. \tag{2.2}$$

То есть из всего множества линий линия регрессии на графике выбирается так, чтобы сумма квадратов расстояний по вертикали между точками и этой линией была бы минимальной (рис. 2.2).

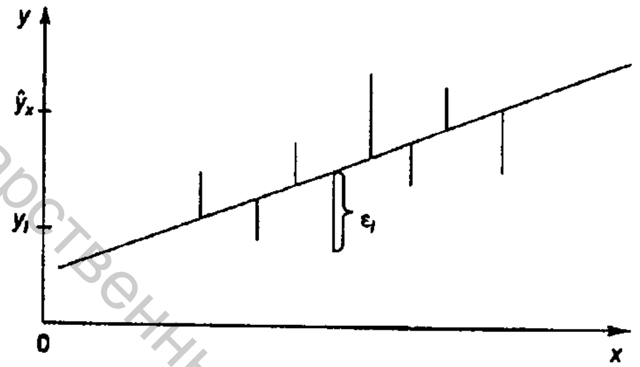


Рисунок 2.2 – Линия регрессии с минимальной дисперсией остатков

Как известно из курса математического анализа, чтобы найти минимум функции (2.2), надо вычислить частные производные по каждому из параметров a и b и приравнять их к нулю. Обозначим $\sum\limits_i \varepsilon_i^2$ через $S(a,\, b)$, тогда

$$S(a,\, b) = \sum\left(y - a - b\cdot x\right)^2.$$

$$\begin{cases} \dfrac{\partial S}{\partial a} = -2\sum\left(y - a - b\cdot x\right) = 0; \\ \dfrac{\partial S}{\partial b} = -2\sum x\left(y - a - b\cdot x\right) = 0. \end{cases} \tag{2.3}$$

После несложных преобразований, получим следующую систему линейных уравнений для оценки параметров a и b:

$$\begin{cases} a\cdot n \quad + b\cdot\sum x = \sum y; \\ a\cdot\sum x + b\cdot\sum x^2 = \sum x\cdot y. \end{cases} \tag{2.4}$$

Решая систему уравнений (2.4), найдем искомые оценки параметров a и b. Можно воспользоваться следующими готовыми формулами, которые следуют непосредственно из решения системы (2.4):

$$a = \overline{y} - b \cdot \overline{x}, \qquad b = \frac{\operatorname{cov}(x, y)}{\sigma_x^2}, \tag{2.5}$$

где $\operatorname{cov}(x,y) = \overline{y \cdot x} - \overline{y} \cdot \overline{x}$ – ковариация признаков x и y, $\sigma_x^2 = \overline{x^2} - \overline{x}^2$ – дисперсия признака x и

$$\overline{x} = \frac{1}{n}\sum x, \quad \overline{y} = \frac{1}{n}\sum y, \quad \overline{y \cdot x} = \frac{1}{n}\sum y \cdot x, \quad \overline{x^2} = \frac{1}{n}\sum x^2.$$

Ковариация – числовая характеристика совместного распределения двух случайных величин, равная математическому ожиданию произведения отклонений этих случайных величин от их математических ожиданий.

Дисперсия – характеристика случайной величины, определяемая как математическое ожидание квадрата отклонения случайной величины от ее математического ожидания. Математическое ожидание – сумма произведений значений случайной величины на соответствующие вероятности.

Параметр b называется коэффициентом регрессии. Его величина показывает среднее изменение результата с изменением фактора на одну единицу.

Возможность четкой экономической интерпретации коэффициента регрессии сделала линейное уравнение регрессии достаточно распространенным в эконометрических исследованиях.

Формально a – значение y при $x = 0$. Если признак-фактор x не может иметь нулевого значения, то вышеуказанная трактовка свободного члена a не имеет смысла, т. е. параметр a может не иметь экономического содержания.

Проверка адекватности модели

Уравнение регрессии всегда дополняется показателем тесноты связи. При использовании линейной регрессии в качестве такого показателя выступает *линейный коэффициент корреляции* r_{xy}, который можно рассчитать по следующей формуле:

$$r_{xy} = \frac{\overline{x \cdot y} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y},$$

где $\overline{xy}$ – среднее значение произведения величин используемых показателей; $\overline{x}$ – среднее значение показателя, рассматриваемого в качестве независимой переменной; $\overline{y}$ – среднее значение показателя, рассматриваемого в качестве зависимой переменной; σ_x – среднеквадратическое отклонение величины ***x***; σ_y – среднеквадратическое отклонение величины ***y***

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})}{n-1}}; \tag{2.6}$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})}{n-1}}, \tag{2.7}$$

где n – число значений переменных.

Линейный коэффициент корреляции находится в пределах $-1 \le r_{xy} \le 1$. Чем ближе абсолютное значение r_{xy} к единице, тем сильнее линейная связь между факторами (при $r_{xy} = \pm 1$ имеем строгую функциональную зависимость). Но следует иметь в виду, что близость абсолютной величины линейного коэффициента корреляции к нулю еще не означает отсутствия связи между признаками. При другой (нелинейной) спецификации модели связь между признаками может оказаться достаточно тесной.

Для оценки качества подбора линейной функции рассчитывается квадрат линейного коэффициента корреляции r_{xy}^2, называемый коэффициентом детерминации. *Коэффициент детерминации* характеризует долю дисперсии (изменения) результативного признака y, объясняемую регрессией, в общей дисперсии результативного признака.

Соответственно величина $1 - r_{xy}^2$ характеризует долю дисперсии (изменения) y, вызванную влиянием остальных, не учтенных в модели, факторов.

После того как найдено уравнение линейной регрессии, проводится оценка значимости как уравнения в целом, так и отдельных его параметров.

Проверить значимость уравнения регрессии – значит установить, соответствует ли математическая модель, выражающая зависимость между переменными, экспериментальным данным и достаточно ли включенных в уравнение объясняющих переменных (одной или нескольких) для описания зависимой переменной.

Чтобы иметь общее суждение о качестве модели из относительных отклонений по каждому наблюдению, определяют *среднюю ошибку аппроксимации:*

$$\bar{A} = \frac{1}{n}\sum\left|\frac{y - \hat{y}_x}{y}\right| \cdot 100\%. \tag{2.8}$$

Средняя ошибка аппроксимации не должна превышать *8–10 %.*

Оценка значимости уравнения регрессии в целом производится на основе F-критерия Фишера, которому предшествует дисперсионный анализ. Согласно основной идее дисперсионного анализа, необходимо сравнить *две суммы квадратов*: общая сумма квадратов отклонений переменной y от среднего значения $\overline{y}$ раскладывается на две части – *«объясненную регрессией» и «необъясненную регрессией»:*

$$\sum\left(y-\overline{y}\right)^2=\sum\left(\hat{y}_x-\overline{y}\right)^2+\sum\left(y-\hat{y}_x\right)^2,$$

где $\sum\left(y-\overline{y}\right)^2$ – общая сумма квадратов отклонений; $Q_{\mathbf{R}}=\sum\left(\hat{y}_x-\overline{y}\right)^2$ – сумма квадратов отклонений, объясненная регрессией; $Q_{\mathbf{E}}=\sum\left(y-\hat{y}_x\right)^2$ – остаточная сумма квадратов отклонений, характеризующая влияние неучтенных в модели факторов.

Схема дисперсионного анализа имеет вид, представленный в таблице 2.1 (где n – число наблюдений, m – число параметров при переменной x).

Таблица 2.1 – Схема дисперсионного анализа

Компоненты дисперсии	Сумма квадратов	Число степеней свободы	Дисперсия на одну степень свободы
Общая	$\sum\left(y-\overline{y}\right)^2$	$n-1$	$S^2_{\text{общ}}=\dfrac{\sum\left(y-\overline{y}\right)^2}{n-1}$
Факторная (регрессионная) Q_R	$\sum\left(\hat{y}_x-\overline{y}\right)^2$	m	$S^2_{\text{факт}}=\dfrac{\sum\left(\hat{y}_x-\overline{y}\right)^2}{m}$
Остаточная Q_E	$\sum\left(y-\hat{y}_x\right)^2$	$n-m-1$	$S^2_{\text{ост}}=\dfrac{\sum\left(y-\hat{y}_x\right)^2}{n-m-1}$

Сопоставляя факторную (регрессионную) и остаточную суммы квадратов (сумы дисперсии) в расчете на одну степень свободы, получим величину F-критерия Фишера

$$F=\frac{S^2_{\text{факт}}}{S^2_{\text{ост}}}. \tag{2.9}$$

Фактическое значение F-критерия Фишера (2.9) сравнивается с табличным значением $F_{\text{табл}}\left(\alpha;k_1;k_2\right)$ при уровне значимости α и степенях свободы $k_1=m$ и $k_2=n-m-1$. При этом, если фактическое значение F-критерия больше табличного, то признается статистическая значимость уравнения в целом.

Для парной линейной регрессии $m = 1$, поэтому

$$F = \frac{S^2_{\text{факт}}}{S^2_{\text{ост}}} = \frac{\sum\left(\hat{y}_x - \overline{y}\right)^2}{\sum\left(y - \hat{y}_x\right)^2} \cdot (n-2). \qquad (2.10)$$

Величина F-критерия связана с коэффициентом детерминации r^2_{xy}, и ее можно рассчитать по следующей формуле:

$$F = \frac{r^2_{xy}}{1 - r^2_{xy}} \cdot (n-2). \qquad (2.11)$$

Проверка значимости параметров модели

В парной линейной регрессии оценивается значимость не только уравнения в целом, но и отдельных его параметров. С этой целью по каждому из параметров определяется его стандартная ошибка: m_b и m_a.

Стандартная ошибка коэффициента регрессии определяется по формуле

$$m_b = \sqrt{\frac{S^2_{\text{ост}}}{\sum\left(x - \overline{x}\right)^2}} = \frac{S_{\text{ост}}}{\sigma_x \cdot \sqrt{n}}, \qquad (2.12)$$

где $S^2_{\text{ост}} = \dfrac{\sum\left(y - \hat{y}_x\right)^2}{n-2}$ – остаточная дисперсия на одну степень свободы.

Затем рассчитывается t-распределение Стьюдента $t_b = \dfrac{b}{m_b}$, которое затем сравнивается с табличным значением при определенном уровне значимости α и числе степеней свободы $(n-2)$.

Доверительный интервал для коэффициента регрессии определяется как $b \pm t_{\text{табл}} \cdot m_b$. Поскольку знак коэффициента регрессии указывает на рост результативного признака y при увеличении признака-фактора x ($b > 0$), уменьшение результативного признака при увеличении признака-фактора ($b < 0$) или его независимость от независимой переменной ($b = 0$) (см. рис. 1.3), то границы доверительного интервала для коэффициента регрессии не должны содержать противоречивых результатов, например, $-1{,}5 \le b \le 0{,}8$. Такого рода запись указывает, что истинное значение коэффициента регрессии одновременно содержит положительные и отрицательные величины и даже ноль, чего не может быть.

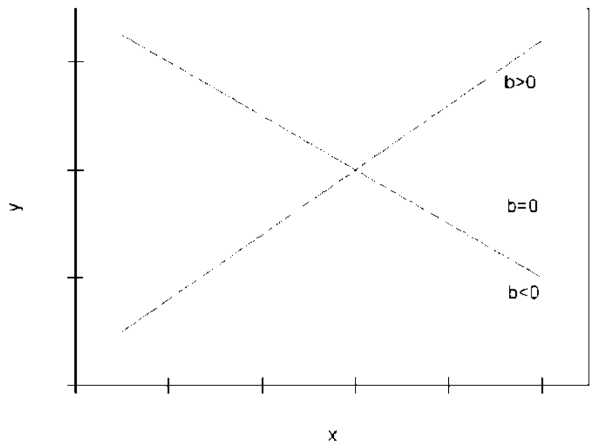


Рисунок 2.3 – Наклон линии регрессии в зависимости от значения параметра b

Стандартная ошибка параметра a определяется по формуле

$$m_a = \sqrt{S^2_{\text{ост}} \cdot \frac{\sum x^2}{n \cdot \sum \left(x - \bar{x} \right)^2}} = S_{\text{ост}} \cdot \frac{\sqrt{\sum x^2}}{\sigma_x \cdot n}. \tag{2.13}$$

Процедура оценивания существенности данного параметра не отличается от рассмотренной выше для коэффициента регрессии.

Вычисляется t-критерий $t_a = \dfrac{a}{m_a}$, его величина сравнивается с табличным значением при $n-2$ степенях свободы.

Значимость линейного коэффициента корреляции проверяется на основе величины ошибки коэффициента корреляции m_r

$$m_r = \sqrt{\frac{1-r^2}{n-2}}. \tag{2.14}$$

Фактическое значение t-критерия Стьюдента определяется как $t_r = \dfrac{r}{m_r}$.

Существует связь между t-критерием Стьюдента и F-критерием Фишера:

$$t_b = t_r = \sqrt{F}. \tag{2.15}$$

В прогнозных расчетах по уравнению регрессии определяется предсказываемое $\hat{y}_p$ значение как точечный прогноз $\hat{y}_x$ путем подстановки в уравнение регрессии $\widehat{y}_x = a + b \cdot x$ соответствующего значения x. Однако точечный прогноз явно не реален. Поэтому обычно он дополняется расчетом стандартной ошибки $\widehat{y}_p$, т. е. $m_{\hat{y}_p}$, и соответственно интервальной оценкой прогнозного значения $\widehat{y}_p$:

$$\widehat{y}_p - \Delta_{\hat{y}_p} \le \widehat{y}_p \le \widehat{y}_p + \Delta_{\hat{y}_p},$$

где $\Delta_{\hat{y}_p} = m_{\hat{y}_p} \cdot t_{\text{табл}}$, а $m_{\hat{y}_p}$ – средняя ошибка прогнозируемого индивидуального значения

$$m_{\hat{y}_p} = S_{\text{ост}} \cdot \sqrt{1 + \frac{1}{n} + \frac{\left(x_p - \overline{x}\right)^2}{n \cdot \sigma_x^2}}. \qquad (2.16)$$

Множественная регрессия и корреляция

Парная регрессия может дать хороший результат при моделировании, если влиянием других факторов, воздействующих на объект исследования, можно пренебречь. Если же этим влиянием пренебречь нельзя, то в этом случае следует попытаться выявить влияние других факторов, введя их в модель, т. е. построить уравнение *множественной регрессии*

$$y = f\left(x_1,\ x_2,\ \ldots,\ x_m\right),$$

где y – зависимая переменная (результативный признак); x_i – независимые, или объясняющие, переменные (признаки-факторы).

Множественная регрессия широко используется в решении проблем спроса, доходности акций, при изучении функции издержек производства, в макроэкономических расчетах и целом ряде других вопросов эконометрики. В настоящее время множественная регрессия – один из наиболее распространенных методов в эконометрике. Основная цель множественной регрессии – построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель.

Построение уравнения множественной регрессии начинается с решения вопроса о спецификации модели. Он включает в себя два круга вопросов: отбор факторов и выбор вида уравнения регрессии.

Факторы, включаемые во множественную регрессию, должны отвечать следующим требованиям.

1. Они должны быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количественного измерения, то ему нужно придать количественную определенность.

2. Факторы не должны быть интеркоррелированы и тем более находиться в точной функциональной связи.

Включение в модель факторов с высокой интеркорреляцией, может привести к нежелательным последствиям – система нормальных уравнений может оказаться плохо обусловленной и повлечь за собой неустойчивость и ненадежность оценок коэффициентов регрессии.

Если между факторами существует высокая корреляция, то нельзя определить их изолированное влияние на результативный показатель и параметры уравнения регрессии оказываются неинтерпретируемыми. Таким образом, хотя теоретически регрессионная модель позволяет учесть любое число факторов, практически в этом нет необходимости. Отбор факторов производится на основе качественного теоретико-экономического анализа в две стадии: на первой подбираются факторы исходя из сущности проблемы; на второй – на основе матрицы показателей корреляции определяют статистики для параметров регрессии.

Коэффициенты интеркорреляции (т. е. корреляции между объясняющими переменными) позволяют исключать из модели дублирующие факторы. Считается, что две переменные явно коллинеарны, т. е. находятся между собой в линейной зависимости, если $r_{x_i x_j} \geq 0,7$. Если факторы явно коллинеарны, то один из них рекомендуется исключить из регрессии. Предпочтение при этом отдается не фактору, более тесно связанному с результатом, а тому фактору, который при достаточно тесной связи с результатом имеет наименьшую тесноту связи с другими факторами.

Пусть, например, при изучении зависимости $y = f(x_1, x_2, x_3)$ матрица парных коэффициентов корреляции оказалась следующей (см. табл. 2.2).

Таблица 2.2 – Матрица парных коэффициентов корреляции

	y	x_1	x_2	x_3
y	1	0,8	0,7	0,6
x_1	0,8	1	0,8	0,5
x_2	0,7	0,8	1	0,2
x_3	0,6	0,5	0,2	1

Очевидно, что факторы x_1 и x_2 дублируют друг друга. В анализ целесообразно включить фактор x_2, а не x_1, хотя корреляция x_2 с результатом y слабее, чем корреляция фактора x_1 с y $\left(r_{yx_2} = 0,7 < r_{yx_1} = 0,8\right)$, но зато значительно слабее межфакторная корреляция $r_{x_2x_3} = 0,2 < r_{x_1x_3} = 0,5$. Поэтому в данном случае в уравнение множественной регрессии включаются факторы x_2, x_3.

По величине парных коэффициентов корреляции обнаруживается лишь явная коллинеарность факторов. Наибольшие трудности в использовании аппарата множественной регрессии возникают при наличии мультиколлинеарности факторов, когда более чем два фактора связаны между собой линейной зависимостью, т. е. имеет место совокупное воздействие факторов друг на друга. Наличие мультиколлинеарности факторов может означать, что некоторые факторы будут всегда действовать в унисон. В результате вариация в исходных данных перестает быть полностью независимой и нельзя оценить воздействие каждого фактора в отдельности.

2.4 Возможности пакетов прикладных программ (ППП) для проведения регрессионного анализа

Как говорилось выше, в общем виде любая модель представляет собой зависимость вида

$$Y = f(x),$$

где *x* – набор независимых параметров, оказывающих влияние на поведение модели; *Y* – выходной параметр, зависимая переменная, изменяющая свое значение при изменении зависимых параметров.

При построении (расчете) эконометрических (регрессионных) моделей задачей моделирования является:

- выявление и расчет тесноты и вида связи между зависимыми и независимыми параметрами;
- расчет уравнения регрессии;
- оценка адекватности модели;
- выполнение прогноза проведения экономической системы.

Расчет и анализ регрессионных моделей представляет собой достаточно долгий и трудоемкий процесс. Поэтому для моделирования и анализа задач экономики удобнее использовать пакеты прикладных программ, которые можно разбить на 2 типа:

Использующие численные методы – пакеты общего назначения (из семейства MS Office)

Использующие точные аналитические решения – аналитические пакеты специального назначения (SAS for Windows, Statistica, SPSS, NCSS, STATA и т. п.).

Наиболее доступным инструментарием для расчета эконометрических моделей является табличный процессор (ТП) MS Excel. В среде ТП MS Excel существует возможность построения регрессионных моделей и прогнозирования на базе следующих технологий:

– с использованием стандартных функций категории «Статистические» – *линейные* и *экспоненциальные* модели;

– графическим способом – *линейные, экспоненциальные, степенные, логарифмические* и т. д.

– с использованием специальной надстройки «Пакет анализа».

Таким образом, в ЭТ существует возможность расчета и анализа регрессионных моделей двух видов: *линейных*, которые получили наибольшее распространение в практической работе и *экспоненциальных*.

Некоторые другие виды моделей – степенная, логарифмическая, полиномиальная – можно получить графически.

Линейная однофакторная модель – это уравнение прямой на плоскости, которая может быть описана уравнением

$$Y = m * x + b \quad ,$$

где m – коэффициент уравнения регрессии, представляющий собой тангенс угла наклона прямой Y = f(x) к оси OX; b – свободный член, равный значению точки пересечения прямой Y = f(x) с осью OY.

Линейная многофакторная модель представляет собой уравнение прямой в многомерном пространстве и имеет вид

$$Y = b + m_1 \cdot x_1 + m_2 \cdot x_2 + \ldots + m_n \cdot x_n,$$

где $x_1, \ldots, x_n$ – независимые переменные (факторы); Y – зависимая переменная; $m_1, \ldots, m_n$ – коэффициенты уравнения регрессии.

Экспоненциальная однофакторная модель имеет вид

$$Y = b \cdot m^x.$$

Экспоненциальная многофакторная модель имеет вид

$$Y = b \cdot m_1^{x1} \cdot m_2^{x2} \ldots m_n^{xn}.$$

Для расчета *линейных моделей* Excel располагает рядом функций, работа которых основана на методе наименьших квадратов: ЛИНЕЙН, НАКЛОН, ОТРЕЗОК, ПРЕДСКАЗ, ТЕНДЕНЦИЯ. Регрессионный анализ предполагает не только расчет параметров модели, но и оценку ее адекватности по некоторым статистическим характеристикам. Часть этих характеристик можно получить как результат работы некоторых стандартных функций, *например*, ЛИНЕЙН, КОРРЕЛ, FРАСПОБР, СТЬЮДРАСПОБР. Для проведения полного корреляционно-регрессионного анализа Excel предлагает специальную надстройку *Пакет Анализа*. Рассмотрим поочередно перечисленные функции.

Функция ЛИНЕЙН рассчитывает статистику ряда с применением метода наименьших квадратов для вычисления уравнения прямой линии, которое наилучшим образом описывает исходные данные. Результатом работы функции является массив, который описывает полученную теоретическую прямую. Эта функция является поистине самой универсальной для расчета параметров линейных моделей. Во-первых, она может использоваться как для расчета однофакторных, так и многофакторных моделей (что определяется размером массива независимых переменных x_i), во-вторых, при желании в качестве результата, кроме коэффициентов уравнения регрессии, можно получить и статистические характеристики, характеризующие построенную модель.

Формат

ЛИНЕЙН (изв. значения Y, изв. значения X, константа; статистика)

Известные значения Y – аргумент обязательный для ввода (всегда только одномерный массив) – это известные значения Y, для которых параметры X по уравнению определены.

Известные значения X – аргумент не обязательный для ввода – это известные значения независимой переменной *X (может быть представлен как одномерным, так и* многомерным массивом). Массив *X* может быть опущен, тогда значения *X* устанавливаются автоматически как предварительный ряд чисел, начиная с 1. Но обязательно должно быть соответствие между размерностями массива *X* и *Y*, если массив *X* задан.

Константа – это логическое значение, которое указывает функции, каким образом должен быть определен коэффициент b. Если логическое значение – ИСТИНА или оно опущено, то b определяется в обычном порядке. Если константа равна ЛОЖЬ, то коэффициенты подбираются таким образом, чтобы выполнялось равенство $y = m \cdot x$ $(b = 0)$.

Статистика – аргумент не обязательный для ввода, логическое значение, которое может принимать значение ИСТИНА или ЛОЖЬ. Если статистика имеет значение ИСТИНА, то будет представлена дополнительная регрессионная статистика по регрессии, если ЛОЖЬ или опущено, то выходным массивом будет основная статистика, т. е. коэффициенты m_1, m_2, …, m_n и b.

В зависимости от количества независимых переменных уравнение получаемой теоретической прямой может иметь вид для однофакторной модели или для многофакторной модели.

В качестве результата функция ЛИНЕЙН возвращает массив коэффициентов уравнения регрессии и дополнительную статистику по регрессии, как показано в таблице 2.3.

Таблица 2.3 – Результаты, возвращаемые функцией ЛИНЕЙН

m_n	m_{n-1}	…	m_2	m_1	b
Se_n	Se_{n-1}	…	Se_2	Se_1	Se_b
r^2	Se_y				
F	d_f				
SS_{reg}	SS_{resid}				

Здесь m_1, $m_2…m_n$, b – коэффициенты уравнения регрессии.

Все остальное – дополнительная статистика по регрессии:

– Se_1, Se_2, Se_n – стандартные ошибки для коэффициентов m_1, $m_2…m_n$;

– Se_b – стандартная ошибка для свободного члена b;

– R^2 – коэффициент детерминированности, который показывает, как близко теоретическое уравнение описывает исходные данные;

– Se_y – стандартная ошибка для Y;

– F-критерий Фишера используется для определения того, является ли наблюдаемая взаимосвязь между зависимой и независимой переменными случайной или нет;

– Df – степень свободы системы (уравнение надежности);

– SSreg – регрессионная сумма квадратов;

– SSresid – остаточная сумма квадратов.

Для выполнения прогноза нового значения Y в полученное уравнение нужно в уравнение регрессии подставить новое значение x.

Функция ПРЕДСКАЗ на основании линейного тренда выполняет точечный прогноз значения зависимой переменной *Y*, соответствующее заданному *X*-значению, по существующим *X*- и *Y*-значениям.

Формат:

ПРЕДСКАЗ (X; известные значения Y; известные значения X)

X – это *точка* данных, для которой предсказывается значение.

Известные_значения_Y – это зависимый массив или интервал данных.

Известные_значения_X – это независимый массив или интервал данных.

Функция ТЕНДЕНЦИЯ – на основании линейного тренда вычисляет или предсказывает будущее значение зависимой переменной *Y*, соответствующее заданному массиву *X*-значений по существующим *X*- и *Y*-значениям.

В отличие от функции ПРЕДСКАЗ, функция ТЕНДЕНЦИЯ позволяет рассчитать теоретические значения *Y* для массива новых значений *X*.

Формат:

ТЕНДЕНЦИЯ (известные_значения_Y; известные_значения_X; новые_значения_X; константа)

Известные_значения_Y – это зависимый массив или интервал данных.

Новые_значения_X – массив (или интервал данных), который должен содержать столбец (или строку) для каждой независимой переменной, как и *Известные_значения_X*. Если аргумент *Новые_значения_X* опущен, то предполагается, что он совпадает с аргументом *Известные_значения_X*. *Константа* – логическое значение, которое указывает, требуется ли, чтобы константа b была равна 0.

Для построения моделей на основании *экспоненциальной зависимости* и дальнейшего расчета прогнозов Excel предлагает функции ЛГРФПРИБЛ и РОСТ, работа которых основана на вычислении экспоненциальной кривой, аппроксимирующей данные. Эти функции могут использоваться как для расчета однофакторных, так и многофакторных моделей.

Функция ЛГРФПРИБЛ в регрессионном анализе вычисляет экспоненциальную кривую, аппроксимирующую данные, и возвращает массив значений, описывающий эту кривую. Она является универсальной для расчета параметров экспоненциальных моделей, так как кроме коэффициентов уравнения регрессии может возвращать и дополнительную статистику по регрессии.

Формат:

ЛГРФПРИБЛ (известные значения_Y; известные значения_X; константа; статистика)

В отличие от функции ЛИНЕЙН аргумент *константа* – это логическое значение, которое указывает, требуется ли, чтобы константа b была равна 1.

Если константа имеет значение ИСТИНА или опущено, то b вычисляется обычным образом.

Если константа имеет значение ЛОЖЬ, то b полагается равным 1 и значения m подбираются так, чтобы удовлетворить соотношению $y = m^x$.

Способ использования функции ЛГРФПРИБЛ аналогичен функции ЛИНЕЙН, с той лишь разницей, что исходный набор данных аппроксимируется не прямой линией, а экспонентой. В соответствии с этим и количеством независимых переменных уравнение регрессии соответствует виду (10.5) или виду (10.6).

Функция РОСТ аналогична функции ТЕНДЕНЦИЯ и используется для расчета прогнозируемого экспоненциального роста на основании имеющихся данных. Она возвращает значения Y для последовательности новых значений X, задаваемых с помощью существующих X- и Y-значений.

Формат

РОСТ (известные значения_Y; известные значения_X; новые_значения_X; константа).

Пример. Линейная регрессия

В таблице 2.4 приведены показатели уровня жизни по территориям регионов республики Беларусь за 20XX г. Провести анализ зависимости среднедневной заработной платы, руб. (Y) от среднедушевого прожиточного минимума в день одного трудоспособного, руб. (x).

Таблица 2.4 – Показатели уровня жизни по территориям регионов республики Беларусь за 20XX г.

Номер региона	Среднедушевой прожиточный минимум в день одного трудоспособного, руб.	Среднедневная заработная плата, руб.
№	X	Y
1	25	48
2	24	45,5
3	20	40
4	21	42,5
...		
13	17	36
14	19	40
15	17	41
16	19	49

Решение

Разместим таблицу с исходными данными в ячейках A3:C25 рабочего листа Excel. В ячейки B24-B25 внесем значения среднедушевого прожиточного минимума, для которых требуется выполнить прогноз уровня среднедневной заработной платы (см. рис. 2.3).

Чтобы выполнить анализ зависимости среднедневной заработной платы, руб. (Y), от среднедушевого прожиточного минимума в день одного трудоспособного, руб. (X), следует построить однофакторную регрессионную модель вида $y = m \cdot x + b$.

Рассчитаем линейную регрессионную однофакторную модель (см. рис. 2.3 и рис. 2.4), для чего в ячейки E5:A9 введем функцию ЛИНЕЙН в формате = ЛИНЕЙН(C5:C23;B5:B23;1;1).

Результатом работы функции является массив значений:

– ячейки E5, F5 – коэффициенты уравнения регрессии m=1.719 и b=5,637;

– ячейка E7 – коэффициент детерминированности R2 = 0,8717;

– ячейка E8 – критерий Фишера F=115,529.

	A	B	C	D	E	F
1						
2						
3	Номер региона	Среднедушевой прожиточный минимум в день одного трудоспособного, руб.	Среднедневная заработная плата, руб.			
4		**X**	**Y**		**m**	**b**
5	1	25	48		**1,719349**	**5,637615**
6	2	24	45,5		0,159963	2,838818
7	3	20	40		**0,871726**	2,541449
8	4	21	42,5		**115,5291**	17
9	5	19	38,5		746,1977	109,8023
10	6	19	38			
11	7	16	35		**Функция Тенденция**	
12	8	17	29		***65,81***	
13	9	16	28		***91,61***	
14	10	13	27			
15	11	16	35			
16	12	15	34		**Функция ПРЕДСКАЗ**	
17	13	17	36		***65,81***	
18	14	19	40			
19	15	18	41		**Функция КОРРЕЛ**	
20	16	19	39		***0,933663***	
21	17	12	28			
22	18	13	26		**Функция F.ОБР.ПХ**	
23	19	11	24		***4,451322***	
24	***20***	***35***	***65,81***			
25	***21***	***50***	***91,61***			

Рисунок 2.3 – Расчет однофакторной регрессионной модели. Результаты

Таким образом, однофакторная регрессионная модель, оценивающая влияние среднедушевого прожиточного минимума в день одного трудоспособного на величину среднедневной заработной платы, имеет вид

$$\mathbf{Y = 1.719 \cdot x + 5,637.}$$

Вывод: поскольку коэффициент детерминированности $R^2 = 0,8717$ лежит в пределах 0,75–1, расчетное значение критерия Фишера F=115,529 больше табличного (F.ОБР.ПХ(0,05;1;F8)= 4,45), модель следует признать адекватной и использовать для прогнозирования.

В ячейке C24 рассчитаем прогнозное значение среднедневной заработной платы по формуле = E5*B22+F5. В ячейке C25 расчет аналогичен (см. рис. 2.4).

	A	B	C	D	E	F
1						
2						
3	Номер региона	Среднедушевой прожиточный минимум в день одного трудоспособного, руб.	Среднедневная заработная плата, руб.			
4		**X**	**Y**		**m**	**b**
5	1	25	48		**=ЛИНЕЙН(C5:C23;B5:B23;1;1)**	**=ЛИНЕЙН(C5:C23;B5**
6	2	24	45,5		=ЛИНЕЙН(C5:C23;B5:B23;1;1)	=ЛИНЕЙН(C5:C23;B5
7	3	20	40		**=ЛИНЕЙН(C5:C23;B5:B23;1;1)**	=ЛИНЕЙН(C5:C23;B5
8	4	21	42,5		**=ЛИНЕЙН(C5:C23;B5:B23;1;1)**	=ЛИНЕЙН(C5:C23;B5
9	5	19	38,5		=ЛИНЕЙН(C5:C23;B5:B23;1;1)	=ЛИНЕЙН(C5:C23;B5
10	6	19	38			
11	7	16	35		**Функция Тенденция**	
12	8	17	29		***=ТЕНДЕНЦИЯ(C5:C23;B5:B23;B24:B25;1)***	
13	9	16	28		***=ТЕНДЕНЦИЯ(C5:C23;B5:B23;B24:B25;1)***	
14	10	13	27			
15	11	16	35			
16	12	15	34		**Функция ПРЕДСКАЗ**	
17	13	17	36		***=ПРЕДСКАЗ(B24;C5:C23;B5:B23)***	
18	14	19	40			
19	15	18	41		**Функция КОРРЕЛ**	
20	16	19	39		***=КОРРЕЛ(B5:B23;C5:C23)***	
21	17	12	28			
22	18	13	26		**Функция F.ОБР.ПХ**	
23	19	11	24		*=F.ОБР.ПХ(0,05;1;F8)*	
24	***20***	***35***	*=$E$5*B24+$F$5*			
25	***21***	***50***	*=$E$5*B25+$F$5*			

Рисунок 2.4 – Расчет однофакторной регрессионной модели. Формулы

В ячейках Е12:Е13 рассчитаем прогнозное значение среднедневной заработной платы для среднедушевого прожиточного минимума, равного 35 руб. и 40 руб. (ячейки В22–В23) с использованием функции ТЕНДЕНЦИЯ:

=ТЕНДЕНЦИЯ(C5:C23;B5:B23;B24:B25;1)

В ячейке Е16 рассчитаем прогнозное значение среднедневной заработной платы с использованием функции ПРЕДСКАЗ:

=ПРЕДСКАЗ(B24;C5:C23;B5:B23).

В ячейке Е18 рассчитаем значение коэффициента корреляции R:

=КОРРЕЛ(C5:C23;B5:B23).

В ячейке Е12 рассчитаем табличное значение критерия Фишера:

=F.ОБР.ПХ(0,05;1;F8).

Полученные значения совпали с результатами, возвращенными функцией ЛИНЕЙН.

Аналогичным образом, используя функцию ЛГРФПРИБЛ, можно рассчитать экспоненциальную регрессионную модель, затем выполнить сравнительный анализ полученных результатов и сделать выбор. Некоторые другие виды регрессионных моделей в ТП MS Excel можно получить графически.

Лекция 3

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ ОПТИМАЛЬНОГО ПЛАНИРОВАНИЯ В ЛОГИСТИКЕ

План лекции

3.1 Понятие оптимизационной модели

Задачи оптимизации относятся к задачам математического программирования – это задачи определения наилучшего решения из множества допустимых.

В общем виде постановка задачи математического программирования состоит в определении значений переменных x1, x2, …, xn, при которых достигается максимум или минимум функции

$$f(x_1, x_2, \ldots, x_n) \to \max(\min) \tag{3.1}$$

при условиях:

$$\begin{cases} \varphi_1(x_1, x_2, .., x_n)\{\le, =, \ge\}b_1 \\ \varphi_2(x_1, x_2, .., x_n)\{\le, =, \ge\}b_2 \\ \qquad \ldots \\ \varphi_m(x_1, x_2, .., x_n)\{\le, =, \ge\}b_m \end{cases} \tag{3.2}$$

Функция (3.1) называется целевой функцией, а условия (3.2) – ограничениями данной задачи. Запись $\{\le, =, \ge\}$ в ограничениях означает, что возможен один из знаков ≤, = или ≥. В данной задаче *n* обозначает число переменных, а *m* – число ограничений.

Переменные задачи $x_1, x_2, ..., x_n$ могут иметь различный экономический смысл. Например, если предприятие выпускает три вида продукции, и нужно найти оптимальный план производства, то $x1, x2, x3$ – количество продукции каждого вида, которое необходимо производить. Если в задаче необходимо найти наилучшую комплектность груза для отправки потребителям, куда могут входить несколько компонентов (например, грузы разных наименований), то $x1$ и $x2$ – количество груза каждого вида, которое нужно включить в отправку.

Целевая функция в математическом виде выражает критерий оптимальности, т. е. служит для выбора наилучшего решения. Если используется максимизируемый критерий оптимальности (например, прибыль от производства продукции), то целевая функция стремится к максимуму. Если же в качестве критерия оптимальности выступают затраты (например, на транспортировку), то целевая функция стремится к минимуму.

Система ограничений (3.2) вытекает из ограниченности материальных, трудовых ресурсов, технологических требований или же из здравого смысла. Например, для задачи планирования производства продукции ограничения вытекают из ограниченности на предприятии материальных и трудовых ресурсов, используемых для производства этой продукции. Для задачи составления рациона ограничения заключаются в необходимости того, чтобы рацион был полноценным (содержал питательные вещества, витамины и микроэлементы, необходимые для жизнедеятельности коров).

В зависимости от характера целевой функции f и функций ограничений $\varphi_1, \varphi_2, .., \varphi_m$, говорят о различных видах задач математического программирования.

1. Если целевая функция задачи имеет линейный вид, а ограничения заданы в виде линейных уравнений или неравенств, то это задача линейного программирования. *Пример* линейного выражения: $3 \cdot x1 + 6 \cdot x2$.

2. Если целевая функция и/или ограничения содержат нелинейные функции, то это задача нелинейного программирования. *Пример* нелинейных функций: $x \cdot y$, x^2, $\sqrt{x}$, $\sin(x)$, $1/x$ и т. д.

3. Если содержательный смысл требует получения решения в целых числах, то такая задача является задачей целочисленного программирования. *Пример*: выпуск штучной продукции, назначение работников на работы (нельзя назначить на работу не целое число работников).

4. Если в задаче математического программирования необходимо учитывать фактор времени, то такая задача является задачей динамического программирования. Обычно решение задач динамического программирования может быть представлено как процесс пошагового принятия решений. На каждом шаге выбирается такое решение, которое не обязательно дает оптимальный результат на этом шаге, но обеспечивает наилучший исход всей операции в целом.

Наиболее разработанными являются методы решения задач линейного программирования. Начало линейной оптимизации было положено в 1939 г., когда вышла в свет работа профессора Ленинградского университета Л.В. Канторовича «Математические методы организации и планирования производства».

Модель задачи линейного программирования должна иметь вполне определенный вид: требуется найти оптимальное – максимум (минимум) значения целевой функции f при переменных x1, x2…, xn

$$f = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \to \min(\max) \tag{3.3}$$

при выполнении ограничений:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)\{\le, =, \ge\} b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n)\{\le, =, \ge\} b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n)\{\le, =, \ge\} b_m \\ \qquad x_i \ge 0, (i = \overline{1,n}) \end{cases} \tag{3.4}$$

где a_{ij}, b_i , c_j – заданные постоянные величины, m – число уравнений, n – число переменных.

Ограничения $x_i \ge 0, (i = \overline{1,n})$ с математической точки зрения являются необязательными, но в моделях экономических задач они, как правило, всегда присутствуют. Это связано с экономическим смыслом переменных *x1, x2, …, xn.* Например, если под x_i понимается количество продукции вида *i,* которое необходимо выпускать на предприятии, то очевидно, что оно не может быть отрицательным.

Ограничения (3.4) определяют область *допустимых* решений.

Под решением задач оптимизации понимается процесс выбора таких значений переменных *x,* принадлежащих допустимой области *D*, которые обеспечивают оптимальное значение некоторой функции *F(x),* называемой целевой.

Набор значений переменных *x1, x2,…,xn,* при котором выполняются все ограничения, называется допустимым решением или планом. Допустимое решение, при котором функция *f* принимает максимальное или минимальное значение, называется *оптимальным*.

Для нахождения оптимального решения следует иметь множество допустимых решений. Если число уравнений m в системе (3.4) равно числу переменных n, то такая система уравнений имеет только одно решение. В задачах линейного программирования число уравнений должно быть меньше числа переменных*: m<n.*

Все переменные, входящие в систему ограничений, должны быть и в целевой функции. Свободные члены *b1, b2…, bn* в системе ограничений должны быть (>=0) положительными или равны нулю.

Достаточно часто ограничения (3.4) задаются в виде системы неравенств.

К типовым оптимизационным задачам линейного программирования можно отнести:

– оптимизацию производственной программы;
– оптимизацию раскроя материалов;
– оптимизацию состава смеси;
– оптимизацию перевозок;
– оптимизацию финансовых показателей;
– оптимизацию штатного расписания и т. п.

Для решения задач линейного программирования необходимо составить математическую модель задачи.

3.2 Математическое обеспечение оптимизационных моделей

3.2.1 Оптимизация производственной программы

Под производственной программой понимается номенклатура продукции и объем выпуска каждого вида изделий, входящих в данную номенклатуру (ассортимент).

В качестве критерия эффективности производственной программы можно выбрать:

1. Max выпуск продукции в натуральном/стоимостном выражении

$$L(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} --> \max \qquad L(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} x_{ij} --> \max$$

при ограничениях на объем выпуска продукции:

– в натуральном выражении: $\sum_{j=1}^{n} x_j = B_i$

– в стоимостном выражении: $\sum_{j=1}^{n} c_j x_{ij} = A_i$

$x_{ij} >= 0$,

где x_{ij} – количество продукции i-го вида, вырабатываемое на одном станке j-го вида; c_{ij} – стоимость единицы i-й продукции, выпускаемой j-м способом, при ограничениях на объем выпуска продукции; *Bi* – плановый выпуск i-той продукции; *Ai* – плановый выпуск i-й продукции в стоимостном выражении.

2. Min себестоимости продукции

$$L(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} s_{ij} x_{ij} --> \min$$

где s_{ij} – себестоимость выпуска i-й продукции, выпускаемой j-м способом.

3. Min трудоемкости продукции

$$L(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} t_{ij} x_{ij} --> \min$$

при ограничениях на трудоемкость выпускаемых изделий

$$\sum_{i=1}^{m} t_{ij} x_{ij} \le T_{ji}$$

где t_{ij} – трудоемкость единицы i-й продукции, выпускаемой j-м способом; Tj – общая трудоемкость рабочих, работающих на j-м оборудовании.

4. *Max выручки от реализации*

$$L(x) = \sum_{j=1}^{n} p_j x_j --> \max$$

При ограничениях на запас i-го ресурса

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i$$

и неотрицательности переменных x_j>=0.

5. *Min затрат*

$$L(x) = \sum_{j=1}^{n} c_j x_j --> \min$$

При ограничении на объемы производства, например, на валовой выпуск продукции в целом (P): $\sum_{j=1}^{n} p_j x_j >= P$

где p_j – цена единицы продукции j-го вида; c_j – затраты ресурсов на выпуск единицы продукции j-го вида.

6. *Max загрузки промышленного оборудования.*

Если в задаче оптимизации производственной программы в качестве ресурсов выступает оборудование (станки различных видов), то ограничения должны описывать фонды времени работы оборудования соответствующего вида в станко-часах. В такой постановке задача 4 становится задачей загрузки взаимозаменяемого оборудования, в которой возможно использование различных критериев. Например, минимум неиспользованного остатка полезного времени работы i-го оборудования

$$L(x) = b_i - \sum_{i=1}^{m} a_{ij} x_j --> \min$$

при ограничениях на фонды времени оборудования и неотрицательности x_j.

После оптимизации производственной программы по различным критериям необходимо провести анализ результатов оптимизации для выбора стратегии развития производства: структуры выпуска продукции, рационального использования оборудования трудовых и сырьевых ресурсов и т. п.

3.2.2 Прикрепление потребителей к поставщикам

Открытая и закрытая транспортные модели

Сущность транспортной задачи линейного программирования состоит в наивыгоднейшем прикреплении поставщиков однородного продукта ко многим потребителям этого продукта. На практике постоянно возникает необходимость решения таких задач, особенно когда количество пунктов отправления и получения грузов увеличивается.

Условие транспортной задачи обычно записывается в виде матрицы, в которой потребители однородного груза размещаются по столбцам, а поставщики – по строкам (рис. 3.1).

Пункты отправления	Пункты назначения					Запасы
	B_1	…	B_j	…	B_n	
A_1	c_{11} x_{11}	… …	c_{1j} x_{1j}	… …	c_{1n} x_{1n}	a_1
…	… …	… …	… …	… …	… …	…
A_i	c_{i1} x_{i1}	… …	c_{ij} x_{ij}	… …	c_{in} x_{in}	a_i
…	… …	… …	… …	… …	… …	…
A_m	c_{m1} x_{m1}	… …	c_{mj} x_{mj}	… …	c_{mn} x_{mn}	a_m
Потребности	b_1	…	b_j	…	b_n	

Рисунок 3.1 – Условие транспортной задачи

В последнем столбце матрицы проставляют запас груза, имеющийся у каждого поставщика, а в последней строке – потребность в нем потребителей. На пересечении строк со столбцами (в клетках матрицы) записывают размер поставки, а также расстояние пробега по всем возможным маршрутам время доставки груза или затраты на перевозку единицы груза по этим маршрутам.

Математически *транспортная задача по критерию стоимости* формулируется следующим образом.

Имеется **n** потребителей и **m** поставщиков однородного груза.

Мощность *i -го* поставщика $(i = \overline{1,m})$ обозначим a_i,.

Спрос *j-го* потребителя $(j = \overline{1,n})$ b_j.

Затраты на перевозку одной тонны груза от *i-го* поставщика до *j-го* потребителя обозначим c_{ij}.

Размер поставки продукции поставщиком i потребителю j обозначим x_{ij}.

Общую сумму затрат на перевозку груза обозначим *через* L.

Запишем математическую модель задачи:

целевая функция – общая сумма затрат на перевозку должна быть минимальной

$$L = \sum_{1=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to min.$$

Ограничения:

– объем поставок *i-го* поставщика должен равняться количеству имеющегося у него груза

$$\sum_{j=1}^{n} x_{ij} = a_i \ (i = \overline{1,m});$$

– объем поставок *j-му* потребителю должен быть равен его спросу

$$\sum_{1=1}^{m} x_{ij} = b_j \ (j = \overline{1,n});$$

– запас груза у поставщиков должен равняться суммарному спросу потребителей

$$a_i = b_j;$$

– размер поставок должен выражаться неотрицательным числом:

$$x_{ij} \geq 0.$$

В случае выполнения равенства $\boldsymbol{a_i = b_j}$, когда запас груза у поставщиков равняется суммарному спросу потребителей, имеет место **закрытая транспортная модель.** Если это равенство не соблюдается, запас груза у поставщиков не равен суммарному спросу потребителей, имеет место **открытая транспортная модель.** В этом случае вводится фиктивный потребитель или фиктивный поставщик, и открытая модель приводится к закрытой.

В случае **превышения** запаса над потребностью, т. е. при

$$\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j$$

– вводится **фиктивный** *(n+1)-й* **пункт назначения** с потребностью

$$b_{n+1} = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j.$$

Соответствующие тарифы считаются равными нулю: $c_{i\,n+1}=0$ ($i=1,...,m$).

После этих преобразований получим закрытую модель транспортной задачи. Аналогично, в случае **превышения** потребности над запасом, т. е. при $\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j$ вводится фиктивный *(m+1)* пункт отправления с грузом $a_{m+1} = \sum_{j=1}^{n} b_j - \sum_{i=1}^{m} a_i$ а тарифы полагаются равными нулю: $c_{m+1\,j}=0$ ($j=1,...,n$).

После этих преобразований получим закрытую модель транспортной задачи.

К классическим экономико-математическим методам решения транспортных задач относится симплекс-метод, который является достаточно громоздким. Поэтому для решения транспортных задач были разработаны специальные методы, в соответствии с которыми решение транспортной задачи разбивается на 2 этапа. Сначала составляется **опорный план** решения. Для этих целей можно использовать:

– метод «северо-западного угла»;
– метод «наименьшего элемента»;
– метод двойного предпочтения и аппроксимации Фогеля.

После нахождения опорного плана перевозок, применяется один из **алгоритмов его улучшения**, приближения к **оптимальному**:

– метод падающего камня (нем.);
– метод потенциалов;
– метод дифференциальных рент;
– распределительный метод.

3.2.3 Транспортные модели с дополнительными ограничениями

Помимо рассмотренных выше, при составлении моделей транспортных задач могут использоваться следующие виды дополнительных ограничений.

1. Ограничения на пропускные способности:

– запрещена перевозка из пункта *Ai* в пункт *Bj*. В этом случае вводится очень высокий тариф *(M)* перевозки из пункта *Ai* в пункт *Bj* единицы продукции: $\boldsymbol{c_{ij} = M}$;

– из пункта *Ai* в пункт *Bj* требуется обязательно перевести точно *dij* единиц продукции. В клетку на пересечении *i-й* строки и *j-го* столбца вносится величина d_{ij}. Корректируется запас a_i и потребность bj: $a_i' = a_i - d_{ij}$ $b_j' = b_j - d_{ij}$. В дальнейшем эта клетка считается свободной, а для того, чтобы соответствующая переменная не попала в состав базисных переменных оптимального решения, этой клетке приписывается очень большой тариф: $\boldsymbol{c_{ij}=M}$;

– из пункта Ai в пункт Bj требуется перевести не менее α_{ij} единиц продукции. Считаем, что из пункта Ai в пункт Bj уже перевезено α_{ij} единиц продукции. Уменьшаем запас a_i и потребность b_j на величину $\alpha_{ij} : \alpha_i^{'} = \alpha_i - \alpha_{ij}$ $b_j^{'} = b_j - \alpha_{ij}$. Далее задача решается обычным методом, после чего корректируется полученное решение. Смысл этой корректировки заключается в следующем. Если в оптимальном решении переменная x_{ij} принимает значение x_{ij}, то в окончательном решении ей приписывается значение $\bar{x}_{ij} = \bar{x}_{ij} + \alpha_{ij}$;

– из пункта Ai в пункт Bj требуется перевести не более α_{ij} единиц продукции. В исходную таблицу вводится дополнительный столбец $b_j^{'}$.

2. *Ограничения на доплаты.*

Транспортной задачей с фиксированными доплатами называется транспортная задача, в которой дополнительно задана матрица доплат d_{ij} за проезд от i-го поставщика к j-му потребителю и требуется минимизировать суммарные расходы. Математическая модель задачи

$$z = \sum_{j=1}^{n} \sum_{i=1}^{m} \left(c_{ij} x_{ij} + d_{ij} y_{ij} \right) \to \min,$$

$$\begin{cases} \sum\limits_{j=1}^{n} x_{ij} = a_i, \\ \sum\limits_{i=1}^{m} x_{ij} = b_j, \\ x_{ij} \geq 0, \end{cases} \qquad y_{ij} = \begin{cases} 0, \text{ если } x_{ij} = 0 \\ 1, \text{ если } x_{ij} > 0 \end{cases}$$

К задачам транспортного типа относятся такие наиболее известные типы задач как:

- транспортная задача;
- задача о назначении;
- распределительная задача;
- задача целераспределения;
- задача перевозок с промежуточной обработкой;
- транспортная задача с дополнительными ограничениями;
- задача перевозок с резервированием;
- задача о максимальном потоке;
- задача о кратчайшем пути;
- многоиндексная транспортная задача.

Все варианты перечисленных задач могут иметь как линейные, так и нелинейные целевые функции.

3.3 Решение оптимизационных моделей средствами компьютерных технологий

Пример 1. Оптимизация плана производства.

Условие. Для производства трех видов изделий A, B и C используется четыре типа технологического оборудования. Время обработки единицы изделия каждого вида на каждом оборудовании и прибыль от реализации единицы изделия каждого вида представлены в таблице ниже.

Таблица 3.1 – Исходные данные к примеру 1

Вид продукции	Время обработки на типе оборудования, ч				Прибыль, у. е.
	I тип	II тип	III тип	IV тип	
A	1	3	1	2	3
B	6	1	3	3	6
C	3	3	2	4	4

На изготовление всех изделий администрация предприятия может предоставить оборудование первого типа не более чем на 84 часа, оборудование второго типа не более чем на 42 часа, оборудование третьего типа не более чем на 21 час и оборудование четвертого типа не более чем на 42 часа. Составить экономико-математическую модель задачи и рассчитать оптимальный план выпуска изделий A, B и C, обеспечивающий максимальную прибыль от их реализации. Сформировать отчет по результатам.

Решение.

1. ЭММ модель задачи.

Целевая функция, выражающая прибыль, получаемую от выпуска всей продукции:

L(x) = 3A+6B+4C--> max

Ограничения:

A + 6B + 3C <= 84
3A + B + 3C <= 42
A + 3B + 2C <= 21
2A + 3B + 4C <= 42
A, B, C >=0; A, B, C – целое.

После размещения исходных данных на рабочем листе ТП MS Excel и выполнения промежуточных расчетов в соответствии с экономико-математической моделью оптимальный план производства легко получить с помощью надстройки «Поиск Решения».

Пример 2. Оптимизация перевозок.

Имеются 5 пунктов производства и 4 пункта распределения продукции. Исходные данные к задаче: стоимость перевозки единицы продукции с *i-го* пункта производства в *j-й* центр распределения c_{ij} оформлены в виде таблицы и приведены на рисунке 3.2, где под строкой понимается пункт производства, а под столбцом – пункт распределения. Кроме того, в этой таблице в *i-той* строке указан объем производства в *i-м* пункте производства, а в *j-м* столбце указан спрос в *j-м* центре распределения. Необходимо составить план перевозок по доставке требуемой продукции в пункты распределения, *минимизирующий* суммарные транспортные расходы.

	Стоимость перевозки единицы продукции				Объемы производства
	5	9	3	10	10
	3	10	5	9	30
	7	2	3	8	20
	8	5	11	2	20
	5	9	10	5	20
Объемы потребления	50	10	30	10	

Рисунок 3.2 – Условие транспортной задачи

Решение

1. Математическая постановка задачи.

Обозначим через X_{ij} – объемы перевозок от *i-го* поставщика *j-му* потребителю. Тогда математическая модель задачи будет иметь вид:

– *Целевая функция* (стоимость перевозок):

$$F=5X_{11}+9X_{12}+3X_{13}+10X_{14}+3X_{21}+10X_{22}+5X_{23}+9X_{24}+7X_{31}+2X_{32}+3X_{33}+8X_{34}+8X_{41}+5X_{42}+11X_{43}+2X_{44}+5X_{51}+9X_{52}+10X_{53}+5X_{54} \rightarrow \min$$

– *Система ограничений на объемы производства*:

$$\begin{cases} x_{11}+x_{12}+x_{13}+x_{14}=10 \\ x_{21}+x_{22}+x_{23}+x_{24}=30 \\ x_{31}+x_{32}+x_{33}+x_{34}=20 \\ x_{41}+x_{42}+x_{43}+x_{44}=20 \\ x_{51}+x_{52}+x_{53}+x_{54}=20 \end{cases}$$

– *Система ограничений на объемы потребления*:

$$\begin{cases} x_{11}+x_{21}+x_{31}+x_{41}+x_{51}=50 \\ x_{12}+x_{22}+x_{32}+x_{42}+x_{52}=10 \\ x_{13}+x_{23}+x_{33}+x_{43}+x_{53}=30 \\ x_{14}+x_{24}+x_{34}+x_{44}+x_{54}=10 \end{cases}$$

– *Ограничения целочисленности и неотрицательности* переменных:

$x_{ij} >= 0$, *x_{ij} – целое.*

После размещения исходных данных на рабочем листе ТП MS Excel и выполнения промежуточных расчетов в соответствии с экономико-математической моделью оптимальный план перевозок легко получить с помощью надстройки «Поиск Решения».

Лекция 4

МОДЕЛИ СЕТЕВОГО И КАЛЕНДАРНОГО ПЛАНИРОВАНИЯ

План лекции

4.1 Сетевые методы планирования и управления, элементы сетевых графиков

4.2 Расчет параметров сетевых графиков

4.3 Построение календарных графиков комплекса взаимосвязанных работ

4.4 Использование сетевого и календарного планирования в логистике

4.5 Возможности ППП для расчета и анализа сетевых моделей

4.1 Сетевые методы планирования и управления, элементы сетевых графиков

Методология сетевого планирования была разработана в 1956 г. специалистами фирм «Дюпон» и «Ремингтон Ред» М. Уолкером и Д. Келли для проекта по модернизации заводов фирмы «Дюпон». Сетевое планирование и управление состоит из *структурного* и *календарного планирования* и *оперативного управления*[1].

Структурное планирование заключается в разбиении проекта на этапы и работы, оценки их длительности, определении последовательности их выполнения. Результатом структурного планирования является сетевой график работ, который используется для оптимизации проекта по длительности.

Метод сетевого планирования используется при планировании сложных комплексов взаимосвязанных работ. Основой этого метода является сетевой график.

Сетевой график (ориентированный граф) – это графическая модель некоторого комплекса взаимосвязанных работ (проекта или производственного процесса).

Работа – это процесс, приводящий к определенным результатам. Работа на графике изображается дугой (стрелкой). **Работа** имеет **продолжительность** и может требовать **ресурсов**. Над дугой может быть указана числовая характеристика работы (например, время выполнения).

[1] Государство, сохраняя за собой право собственности, передает имущество государственной организации в оперативное управление. Это право дает организации возможность владеть, пользоваться и в известной мере распоряжаться этим имуществом от своего имени. Право оперативного управления впервые получило законодательное закрепление в Основах гражданского законодательства СССР 1961г.

Вершинам графика соответствуют **события** (вершина изображается кружком или квадратиком).

Событие – факт окончания всех работ, в него входящих, и начала всех работ, из него исходящих. Пока не выполнены все работы, входящие в событие, не может свершиться само событие и, следовательно, не может быть начата ни одна из работ, выходящих из него. Событие не имеет продолжительности и не требует ресурсов.

Событие в сетевом графике имеет номер, а работа обозначается двумя номерами *(i,j),* где *i* – номер начального события работы, а *j* – номер конечного события работы (см. рис. 4.1). Продолжительность работы обозначается *t(i,j).*

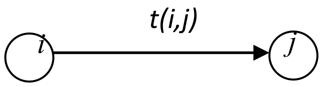


Рисунок 4.1 – Изображение работы на сетевом графике

Событие, с которого начинается выполнение проекта, называется **исходным** и обозначается **I**. Исходное событие не имеет предшествующих работ.

Событие, которое констатирует факт завершения проекта, называется **завершающим** и обозначается **S**. Завершающее событие не имеет последующих работ. В сетевом графике может быть только одно исходное и только одно завершающее событие.

Путь – это последовательность работ в сетевом графике. **Полный путь** – это цепочка следующих друг за другом работ, соединяющих исходное и завершающее событие.

На основе сетевого графика могут быть решены следующие задачи:

1. Анализ последовательности и взаимосвязи работ. Сам процесс построения сетевого графика дает возможность четко выявить взаимосвязь различных этапов проекта, условия начала тех или иных работ.
2. Определение срока выполнения проекта (критического срока).
3. Выявление возможностей задержки начала каждой работы или удлинения срока ее выполнения.
4. Оптимизация времени выполнения проекта или ресурсов, требуемых для его выполнения.

4.2 Расчет параметров сетевых графиков

Рассмотрим пример сетевого графика (рис. 4.2). Это график проекта некоторой логистической фирмы, включающий комплекс работ по подготовке к участию в тендере. Перечень работ приведен в таблице 4.1.

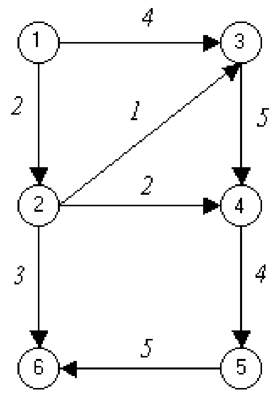


Рисунок 4.2 – Сетевой график примера

Таблица 4.1 – Перечень работ проекта по подготовке к участию в тендере

Содержание работы	Обозначение	Продолжительность работы, дней
Определение рекламной стратегии	(1,2)	2
Разработка дизайна проекта экспозиции	(1,3)	4
Определение количества и видов рекламно-информационных материалов	(2,3)	1
Заключение договора на участие и оплата аренды	(2,4)	2
Обучение и инструктаж персонала	(2,6)	3
Заказ оборудования и рекламных материалов, оплата счетов	(3,4)	5
Доставка оборудования, экспонатов и рекламных материалов	(4,5)	4
Техническое оформление стендов	(5,6)	5

Данный проект включает восемь работ и шесть событий. Сетевой график отражает взаимосвязь работ проекта.

Например, работа (2,3) имеет продолжительность 1 день. Она может быть начата только тогда, когда завершится работа (1,2).

Работа (3,4) имеет продолжительность 5 дней. Она может быть начата только тогда, когда завершатся обе работы, ей предшествующие: (1,3) и (2,3).

Событие 4 состоит в факте окончания обоих работ (2,4) и (3,4) и начала работы (4, 5). Событие 4 не наступит, если хотя бы одна из работ (2,4) или (3,4) не завершена. Аналогично можно объяснить смысл остальных событий.

В примере исходным является событие 1, а завершающим – событие 6.

Выделим следующие полные пути (они обозначаются номерами событий, через которые проходят):

μ_1=(1-2-3-4-5-6);

μ_2=(1-3-4-5-6);

μ_3=(1-2-4-5-6);

μ_4=(1-2-6).

Критическим *называется полный путь, имеющий наибольшую продолжительность во времени*. Критических путей на сетевом графике может быть несколько (при этом все они имеют одинаковую продолжительность).

Продолжительность критического пути определяет **критический срок проекта tкр.** Все остальные (некритические) полные пути выполняются параллельно с критическим путем (цепочкой работ) и завершаются раньше. **Критический срок, таким образом, показывает, за какое минимальное время может быть завершен весь проект**. Очевидно, что увеличение сроков выполнения проекта больше tкр невыгодно.

Работы, принадлежащие критическому пути, называются критическими. **Они не имеют резервов времени. Их несвоевременное выполнение ведет к срыву сроков всего проекта**.

В нашем примере определить критический путь легко: нужно перебрать все возможные полные пути, рассчитать продолжительность каждого из них и выбрать наибольший:

$t(\mu_1)$=2+1+5+4+5=17;

$t(\mu_2)$=4+5+4+5=**18**;

$t(\mu_3)$=2+2+4+5=13;

$t(\mu_4)$=2+3=5.

Критическим является полный путь **μ2**, т. к. он имеет наибольшую продолжительность. Критический путь принято выделять на графике жирной линией (рис. 4.3).

Однако, если сетевой график достаточно сложный, перебрать все возможные пути затруднительно. Поэтому используют более формальный подход:

– для каждого события рассчитывают ранний и поздний сроки свершения;

– на их основе определяют резервы времени всех событий и работ.

Проводят критический путь по тем работам и событиям, которые не имеют резерва времени.

Ранний срок свершения события – это самый ранний момент, к которому завершаются все работы, предшествующие этому событию.

Ранний срок свершения события рассчитывается последовательно для каждого события от исходного к завершающему по следующим формулам:

$t_p(I) = 0$, т. е. начало проекта принимается за нулевой момент времени;

$t_p(j) = t_p(i) + t(i,j)$, если событию *j* предшествует только одна работа;

$t_p(j) = \max\limits_{i \to j}\{t_p(i) + t(i,j)\}$, если событию предшествует несколько работ.

Здесь $i \to j$ – множество работ, заканчивающихся *j*-м событием (дуги, входящие в вершину *j*); $t_p(i)$ – ранний срок свершения события, с которого начинается работа *(i,j)*; $t(i,j)$ – продолжительность работы *(i,j)*.

Рассчитаем ранние сроки свершения событий для нашего примера. Результат расчетов для каждого события будем записывать возле соответствующей вершины графа на рисунке 2.3 (в скобках возле каждой вершины записываются поздние сроки свершения событий, которые мы рассчитаем далее).

tp(1)=0 (Расчет времени начинается с 0).

Событие 2 наступит тогда, когда закончится работа (1,2). Эта работа начнется в момент времени 0 и продлится 2 дня. Поэтому она закончится в 0+2=2 день:

tp(2)=tp(1)+t(1,2) =0+2=2.

В вершину 3 входят две стрелки, т. е. событие 3 наступит тогда, когда закончатся обе работы (1,3) и (2,3). Работа (1,3) начнется в момент времени 0 и продолжится 4 дня. То есть она закончится в 0+4=4 день. Аналогично работа (2,3) закончится в 2+1=3 день. Поскольку обе работы должны закончиться, чтобы наступило событие 4, нужно ориентироваться на самую позднюю из них, т. е. взять максимум по входящим в событие работам:

tp(3)=max{tp(1)+t(1,3), tp(2)+t(2,3)}=max{0+4,2+1}=4.

Так же находят ранние сроки остальных событий проекта:

tp(4)=max{ tp(2)+t(2,4),tp(3)+t(3,4)}=max{2+2,4+5}=9;

tp(5)= tp(4)+t(4,5)=9+4=13;

tp(6)= max{tp(2)+t(2,6), tp(5)+t(5,6)}=max{2+3,13+5}=18.

Критический срок проекта совпадает с ранним сроком свершения завершающего события проекта:

Ткр = tp(S).

Таким образом, рассчитав ранние сроки, мы узнали критический срок проекта нашего примера: $\boldsymbol{t_{кр} = t_p(6) = 18}$.

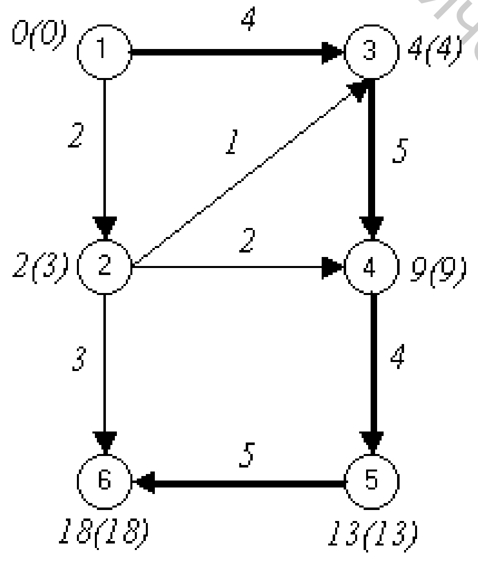


Рисунок 4.3 – Сетевой график примера с результатами расчетов

Поздний срок свершения события – это такой предельный момент, после которого остается ровно столько времени, сколько необходимо для выполнения всех работ, следующих за этим событием, к критическому сроку.

Поздние сроки свершения событий рассчитываются «**обратным ходом**» от завершающего события к исходному по следующим формулам:

$t_n(S) = t_p(S) = t_{кр}$, т. е. для завершающего события поздний срок свершения совпадает с критическим сроком;

$t_n(i) = t_n(j) - t(i,j)$, если событием i начинается одна работа;

$t_n(i) = \min\limits_{i \to j}\{t_n(j) - t(i,j)\}$, если событием i начинается несколько работ.

Здесь $i \to j$ – множество работ, начинающихся i-м событием (дуги, исходящие из вершины i); $t_n(j)$ – поздний срок свершения события, которым заканчивается работа (i,j); $t(i,j)$ – продолжительность работы (i,j).

Рассчитаем поздние сроки свершения событий для нашего примера и запишем их в скобках возле соответствующей вершины (рис. 4.3).

Для завершающего события*: tn(6)= tp(6)=18.*

Рассчитывая поздний срок свершения события 5, необходимо учитывать, что этим событием начинается работа (5,6), которая должна быть обязательно закончена к 18 дню. Она длится 5 дней, поэтому самый поздний момент, когда она должна начаться, это 18-5=13 день. Если вдруг событие 5 наступит, скажем, на 14 день, то работа (5,6) закончится на 14+5=19 день и срок выполнения всего проекта будет сорван. Поэтому можно записать для события 5:

tn(5)=tn(6)-t(5,6)= 18-5=13.

Событием 4 начинается одна работа (4,5). Она должна быть закончена к 13 дню для того, чтобы следующая за ней работа успела к критическому сроку. Поэтому работа (4,5) должна начаться не позже, чем на 13-4=9 день. Таким образом,

tn(4)=tn(5)-t(4,5)= 13-4=9.

Аналогично рассчитываем поздний срок свершения события 3:

tn(3)=tn(4)-t(3,4)= 9-5=4.

Событием 2 начинаются три работы: (2,3), (2,4) и (2,6). Все они должны успеть закончиться вовремя, т. е. работа (2,3) – к 4 дню, работа (2,4) – к 9 дню, а работа (2,6) – к 18 дню. Для этого работа (2,3) должна начаться не позже, чем на 4-1=3 день, работа (2,4) – на 9-2=7 день, а работа (2,6) должна начаться не позже, чем на 18-3=15 день. Чтобы успели все эти работы, нужно, чтобы успела та из них, которая начинается раньше. Поэтому нужно найти минимум по исходящим из события 2 работам:

tn(2)=min{tn(3)-t(2,3), tn(4)-t(2,4), tn(6)-t(2,6)}=min{4-1, 9-2, 18-3}=3.

Аналогично находится поздний срок свершения события 1, из которого выходят две работы:

tn(1)= min{tn(3)-t(1,3), tn(2)-t(1,2)}=min{4-4, 3-2}=0.

Резерв времени события показывает, на какой предельно допустимый срок может задержаться свершение события i без нарушения критического срока проекта:

$R(i) = t_п(i) - t_р(i)$

Рассчитаем резервы времени событий для нашего примера:

$R(1)=t_n(1)-t_p(1)=0-0=0;$
$R(2)= t_n(2)-t_p(2)=3-2=1;$
$R(3)= t_n(3)-t_p(3)=4-4=0;$
$R(4)= t_n(4)-t_p(4)=9-9=0;$
$R(5)= t_n(5)-t_p(5)=13-13=0;$
$R(6)= t_n(6)-t_p(6)=18-18=0.$

Таким образом, можно задержать свершение события 2 на 1 день. Остальные события 1, 3, 4, 5 и 6 не имеют резерва времени. Поэтому они принадлежат критическому пути. Если бы ранее мы не выделили критический путь на сетевом графике, то можно было бы провести его сейчас, после расчетов резервов времени событий, через события 1, 3, 4, 5 и 6. Для проверки следует сложить продолжительности работ этого полного пути, которые в сумме должны быть равны критическому сроку: *4+5+4+5=18= tкр.*

Резерв могут иметь не только события, но и работы проекта.

Полный резерв времени работы показывает, как можно увеличить время выполнения этой работы при условии, что срок выполнения всего комплекса работ не изменится.

Резервы работ определяются на основе параметров свершения событий по следующей формуле:

$$R(i,j) = t_n(j) - t_p(i) - t(i,j)$$

Рассчитаем резервы работ примера:

$R(1,2)=t_n(2)-t_p(1)-t(1,2)=3-0-2=1;$
$R(1,3)=t_n(3)-t_p(1)-t(1,3)=4-0-4=0;$
$R(2,3)=t_n(3)-t_p(2)-t(2,3)=4-2-1=1;$
$R(2,4)=t_n(4)-t_p(2)-t(2,4)=9-2-2=5;$
$R(2,6)=t_n(6)-t_p(2)-t(2,6)=18-2-3=13;$
$R(3,4)=t_n(4)-t_p(3)-t(3,4)=9-4-5=0;$
$R(4,5)=t_n(5)-t_p(4)-t(4,5)=13-9-4=0;$
$R(5,6)=t_n(6)-t_p(5)-t(5,6)=18-13-5=0.$

Критические работы резервов времени не имеют, т. е. еще раз убеждаемся в том, что критический путь мы выделили правильно.

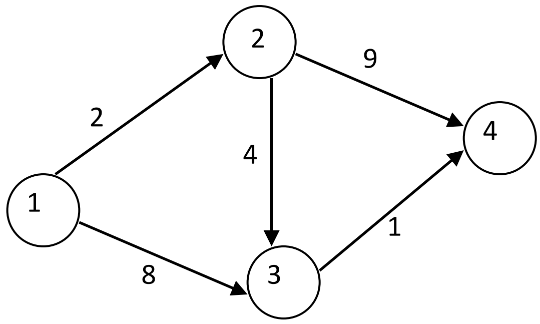
Резервы времени работ рассчитываются для организации контроля над выполнением проекта. Кроме того, зная эти резервы, можно оптимизировать срок выполнения проекта. Например, можно забрать ресурсы у тех работ, которые имеют резерв времени (снять часть рабочих с этих работ или урезать их финансирование) и передать их работам, лежащим на критическом пути. Тогда критические работы смогут быть выполнены раньше, что повлечет уменьшение критического срока всего проекта. Поскольку при таком перераспределении ресурсов критический путь может измениться, задача оптимизации критического срока является многоэтапной и может быть решена с использованием компьютера.

Пример 1. Для представленного на рисунке сетевого графика определите

a) критический срок;

b) ранний срок свершения события 3;

c) поздний срок свершения события 3.



Ответы:

a) перебираем все полные пути: $t(1\text{-}2\text{-}4)=2+9=\mathbf{11}$; $t(1\text{-}2\text{-}3\text{-}4)=2+4+1=\mathbf{7}$; $t(1\text{-}3\text{-}4)=8+1=\mathbf{9}$. Выбираем наибольшую продолжительность. $\mathbf{t_{кр}=11}$**;**

b) по определению $t_p(1)=0$. Ранний срок события 2 равен $t_p(2)=0+2=2$. В событие 3 входят две работы, поэтому находим максимум по этим работам: $t_p(3)=\max(0+8;2+4)=\mathbf{8}$**;**

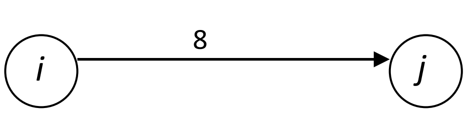
c) поздний срок события 4 по определению равен критическому сроку, т.е. $t_п(4)=11$. Из события 3 выходит одна работа, поэтому $t_п(3)=11-1=\mathbf{10}$**.**

Пример 2. Для работы (i,j), показанной на рисунке, рассчитаны параметры свершения событий. Определить:

a) полный резерв времени этой работы;

b) резерв события *i*;

c) резерв события *j*.



$t_p(i)=1;\ t_п(i)=2;\ t_p(j)=11;\ t_п(j)=14$

Ответы:

a) 5 ($R(i,j)=14-1-8=5$);

b) 1 ($R(i)=2-1=1$);

c) 3 ($R(j)=14-11$).

4.3 Календарное планирование

Календарное планирование заключается в составлении временной диаграммы работ и распределении между работами трудовых ресурсов (исполнителей). Результатом календарного планирования является диаграмма Ганта, графически отображающая периоды выполнения работ на оси времени. На этом этапе может выполняться оптимизация ресурсов и бюджета проекта.

Диаграмма Ганта отображает следующие параметры проекта:

1) структуру работ, полученную на основе сетевого графика;

2) состав используемых ресурсов и их распределение между работами;

3) календарные даты, к которым привязываются моменты начала и завершения работ.

Построение календарного графика можно проиллюстрировать на примере проекта «Подготовка к участию в тендере», (рис. 4.4–4.5).

	Название задачи	Длительность	Начало	Окончание	Предше
0	**Проект к Лк 4_2018**	**18 дней?**	**Вт 08.01.19**	**Чт 31.01.19**	
1	Начало работы	0 дней	Вт 08.01.19	Вт 08.01.19	
2	Определение рекламной стратегии	2 дней	Вт 08.01.19	Ср 09.01.19	1
3	Разработка дизайна проекта экспозиц	4 дней	Вт 08.01.19	Пт 11.01.19	1
4	Определение рекламно- информацион	1 день?	Чт 10.01.19	Чт 10.01.19	2
5	Заключение договора на участие и оп	2 дней	Чт 10.01.19	Пт 11.01.19	2
6	Обучение и инструктаж персонала	3 дней	Чт 10.01.19	Пн 14.01.19	2
7	Заказ оборудования и рекламных мат	5 дней	Пн 14.01.19	Пт 18.01.19	2;4;3
8	Доставка оборудования, экспонатов	4 дней	Пн 21.01.19	Чт 24.01.19	7;5
9	Техническое оформление стендов	5 дней	Пт 25.01.19	Чт 31.01.19	8

Рисунок 4.4 – Перечень работ проекта «Подготовка к участию в тендере»

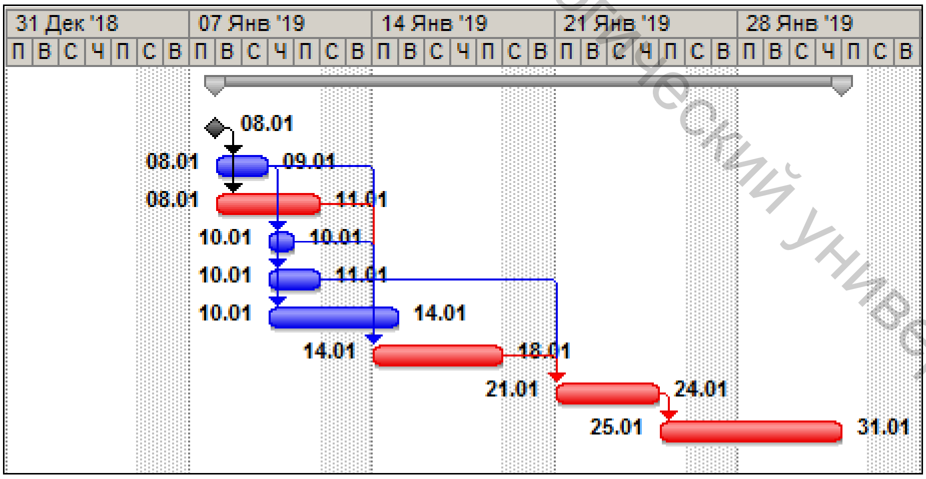


Рисунок 4.5 – Календарный график (Диаграмма Ганта)

4.4 Использование сетевого и календарного планирования в логистике

Сетевое планирование применяется для оптимизации планирования и управления сложными разветвленными комплексами работ, требующими участия большого числа исполнителей и затрат ограниченных ресурсов. Основная цель сетевого планирования – сокращение до минимума продолжительности проекта. Задача сетевого планирования состоит в том, чтобы графически, наглядно и системно отобразить и оптимизировать последовательность и взаимозависимость работ, действий или мероприятий, обеспечивающих своевременное и планомерное достижение конечных целей.

При реализации логистической деятельности методы и модели сетевого планирования могут использоваться при решении таких задач, как:

- логистика производства;
- логистика запасов;
- складская логистика;
- транспортная логистика;
- информационная логистика;
- распределительная логистика;
- логистика снабжения (закупок).

4.5 Возможности ППП для расчета и анализа сетевых моделей

Определенные возможности для расчета и оптимизации сетевых графиков предоставляет пользователю ТП MS Excel.

Специализированные системы управления проектами образуют отдельный сектор программного обеспечения, который достаточно широко представлен на рынке ПО. Появление подобных систем способствовало преобразованию искусства управления проектами в науку, в которой имеются четкие стандарты, методы и технологии:

1. Стандарт, разработанный Институтом управления проектами (Project Management Institute) принят в качестве национального стандарта в США (стандарт ANSI).

2. Стандарт по качеству в управлении проектами ISO 10006.

Применение этих технологий способствует своевременной реализации проектов в рамках выделенных бюджетов и с требуемым качеством.

Системы управления проектами используются для решения следующих основных задач.

1. Структуризация и описание состава и характеристик работ, ресурсов, затрат и доходов проекта.

2. Расчет расписания исполнения работ проекта с учетом всех имеющихся ограничений.

3. Определение критических операций и резервов времени для исполнения других операций проекта.

4. Расчет бюджета проекта и распределение запланированных затрат во времени.

5. Расчет распределения во времени потребности проекта в основных материалах и оборудовании.

6. Определение оптимального состава ресурсов проекта и распределения во времени их плановой загрузки.

7. Анализ рисков и определение необходимых резервов для надежной реализации проекта.

8. Определение вероятности успешного исполнения директивных показателей.

9. Ведение учета и анализ исполнения проекта.

10. Моделирование последствий управленческих воздействий с целью принятия оптимальных решений.

11. Ведение архивов проекта.

12. Получение необходимой отчетности.

На рынке (в том числе и русскоговорящем) в настоящее время наиболее популярными являются несколько систем управления проектами.

А. Microsoft Office Project – это комплексное решение корпорации Microsoft по управлению корпоративными проектами, которое позволяет управлять проектами любой сложности.

Б. Spider Project Professional (также существуют версии Desktop и Lite, разработчик «Технологии управления Спайдер») – пакет управления проектами, спроектированный и разработанный с учетом практического опыта, потребностей, особенностей и приоритетов Российского рынка. Этот пакет – единственная отечественная разработка в области СУП.

В. Primavera Project Planner Professional – профессиональная версия, предназначенная для автоматизации процессов управления проектами в соответствии с требованиями PMI (Project Management Institute) и стандартами ISO. В первую очередь этот пакет предназначен для использования в составе корпоративной информационной системы, хотя вполне может работать и автономно, помогая решать задачи календарно-сетевого планирования, определения критического пути, выравнивания ресурсов, и других задач моделирования проектов, групп проектов, портфелей и программ.

Г. Open Plan (разработчик Welcom Software Technology, сейчас Deltek) обеспечивает полномасштабное мультипроектное управление, планирование по методу критического пути и оптимизацию использования ресурсов в масштабах предприятия. Может эффективно использоваться на всех уровнях контроля и управления проектами – от высшего руководства и менеджеров проектов, до начальников функциональных подразделений и рядовых исполнителей.

Д. Rillsoft Project предназначен для расчета оптимального календарного плана работ, контроля, анализа и управления проектом.

Е. Merlin, OmniPlan Программные продукты по управлению проектами для операционной системы Mac OS.

Ж. OmniPlan. Программа для эффективного планирования бизнес-проектов. С помощью OmniPlan можно планировать дела, составлять диаграммы Ганта, отслеживать выполнение проектов, составлять отчеты и многое другое.

Лекция 5

МОДЕЛИ ПРИНЯТИЯ РЕШЕНИЙ

План лекции

5.1 Основные понятия теории игр

5.2 Статистические игры (игры с природой)

5.3 Критерии выбора наилучших стратегий в условиях неопределенности

5.4 Критерий выбора наилучших решений в условиях риска

5.1 Основные понятия теории игр

В экономике часто возникают ситуации, в которых интересы участвующих сторон противоположны. Такие ситуации называют конфликтными.

Математическую модель реальной конфликтной ситуации называют **игрой**.

В игре могут сталкиваться интересы двух (**игра парная**) или нескольких (**игра множественная**) противников.

Игра ведется по определенным правилам. Каждый участник игры имеет несколько вариантов возможных действий (**чистых стратегий**). Из них он выбирает такие варианты, которые, как он полагает, могут обеспечить ему наилучший результат (**исход игры**). При этом каждый игрок имеет лишь общее представление о множестве допустимых ответных действий партнера, но не о его конкретном решении. В связи с этим ни один из игроков не может контролировать положение, поэтому как одному, так и другому игроку решение приходится принимать в **условиях неопределенности**. Непременным остается только стремление игроков использовать любую ошибку партнера в своих интересах. Игры, в которых оба участника, действуя в строгом соответствии с правилами, в равной мере сознательно стремятся добиться наилучшего для себя результата, называются **стратегическими**.

В экономике модель поведения лиц в виде игры возникает, например, при попытке нескольких фирм завоевать наиболее выгодное место на конкурентном рынке, или, например, при желании нескольких компаний разделить некоторые ресурсы между собой так, чтобы каждому досталось как можно больше. Игроками в конфликтных экономических ситуациях являются фирмы, банки, отдельные лица и другие экономические агенты

5.2 Статистические игры (игры с природой)

В экономической практике нередко приходится формализовать (моделировать) ситуации, в которых один из участников безразличен к результату игры. Такие игры называют **статистическими** или **играми с природой.** Под термином «**природа**» понимают всю совокупность внешних обстоятельств, в которых сознательному игроку приходится принимать решение. Стороны, участвующие в подобной ситуации это – сознательно действующий игрок и «природа». В этом случае стратегиями природы будут ее возможные состояния. Сознательно действующий игрок может собрать дополнительную статистическую информацию о возможных состояниях природы. Потому **цель игры** – выбор наилучших стратегий (с точки зрения возможно большего выигрыша или возможно меньшего проигрыша сознательно действующего игрока).

Исход игры – это значение некоторой функции, называемой функцией выигрыша (платежной функцией). Платежная функция определяет для каждой совокупности выбранных игроками стратегий выигрыш каждой из сторон. Такая функция задается либо таблицей (платежная матрица), либо аналитическим выражением.

Пусть у игрока, действующего сознательно, есть m стратегий, которые мы обозначим A_1 , A_2 , …, A_m ,а у второго игрока (природы) – n стратегий, которые мы обозначим $П_1$, $П_2$, …, $П_n$. Прибыль или убыток (выигрыш или проигрыш) сознательного игрока, если он выберет стратегию A_i , а природа реализует стратегию $П_j$, обозначим a_{ij} ($i=\overline{1,m}$; ,$j=\overline{1,n}$) . Тогда можно построить таблицу выигрышей, которая называется платежной матрицей А:

$$A = \begin{array}{c} \\ \begin{matrix} A_1 \\ A_2 \\ \cdots \\ A_m \end{matrix} \end{array} \begin{array}{c} \begin{matrix} П_1 & П_2 & \cdots & П_n \end{matrix} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \end{array}.$$

Сознательно действующий игрок принимает решение о выборе той или иной стратегии, поэтому в дальнейшем будем называть его принимающим решение. При выборе наилучших стратегий различают две ситуации: ситуацию, в которой вероятности состояний природы неизвестны, и тогда говорят о принятии решений в **условиях неопределенности**, и ситуацию, в которой вероятности состояний природы известны, тогда говорят о принятии решений **в условиях риска**. Для каждой из ситуаций существуют свои критерии (принципы) выбора наилучших решений.

5.3 Критерии выбора наилучших стратегий в условиях неопределенности

1. Критерий Вальда. Данный критерий основан на принципе крайнего пессимизма. Принимающий решение считает, что какую бы стратегию он ни выбрал, природа реализует свое наихудшее состояние. В наихудших условиях принимающий решение находит наилучший выход.

Таким образом, принимающий решение для каждой стратегии A_i находит наименьший выигрыш

$$a_i = \max_{1\le i\le m} a1 = max_{1\le i\le m}\, min_{1\le j\le n}\, a_{ij}$$

Стратегия A_{io}, соответствующая a_{io} ,будет наилучшей по Вальду. Ее часто называют максиминной стратегией.

2. Критерий Сэвиджа. Этот критерий основан на принципе минимизации максимального риска. Риском r_{ij} ($i=\overline{1,m}$, $j=\overline{1,n}$) принимающего решение называют разницу между тем выигрышем, который он бы получил, если бы знал, какое состояние реализует природа, и его реальным выигрышем, т.е. $r_{ij} = \beta_{ij} - a_{ij}$, где $\beta_j = max_{1\le i\le m}\, a_{ij}$, $j=\overline{1,n}$.

Матрица рисков R имеет вид

$$R = \begin{array}{c|cccc} & \Pi_1 & \Pi_2 & \cdots & \Pi_n \\ \hline A_1 & & & & \\ A_2 & & \left(\begin{matrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{matrix}\right) & & \\ \cdots & & & & \\ A_m & & & & \end{array}.$$

Принимающий решение для каждой стратегии A_i находит максимальный риск r_i , $r_i = max_{1\le j\le n}\, r_{ij}$. Затем их максимальных рисков выбирает минимальный

$$r_{io} = min_{1\le i\le m}\, r_i = min_{1\le j\le n}\, r_{ij}\,.$$

Стратегия A_{io} , соответствующая минимальному из максимальных рисков r_{io} , будет наилучшей по Сэвиджу.

3. Критерий Гурвица. Это критерий пессимизма-оптимизма. Наилучшей по Гурвицу является стратегия A_{io} , соответствующая числу a_{i0} , которое рассчитывается по формуле

$$a_{io} = max_{1\le i\le m}\, a_i\,,\ a_i = \gamma\, min_{1\le j\le n}\, a_{ij} + (1-\gamma)\, max_{1\le j\le n}\, r_{ij}\,.$$

Значение параметра γ задает принимающее решение на основании своего опыта. Если γ =1, то критерий Гурвица преобразуется в критерий крайнего пессимизма:

$$a_{io} = max_{1\le i\le m} \left\{1 * min_{1\le j\le m}\, a_{ij} + max_{1\le j\le n}\, a_{ij}\right\} = max_{1\le i\le m}\, min_{1\le j\le n}\, a_{ij}\,.$$

Если γ=0, то получаем критерий пессимизма:

$a_{io} = max_{1\le i\le m}\ \{0 * mina_{1\le j\le n}\ a_{ij} + maxa_{1\le j\le n}\ a_{ij}\} = max_{1\le i\le m}\ min_{1\le j\le n}\ a_{ij}$

.

На практике выбирают $0<\gamma<1$.

5.4 Критерий выбора наилучших решений в условиях риска

Как уже было сказано ранее, в этой ситуации известны вероятности, с которыми реализуются состояния природы. Эти вероятности либо рассчитываются на основе статистических данных, либо определяются экспертным путем. Для принятия решений в условиях риска используется критерий Байеса.

Пусть принимающий решение имеет m стратегий, а природа – n, причем состояние природы $П_j$ реализуется с вероятностью p_j ($j=\overline{1,n}$). Для каждой стратегии A_i рассчитывается ожидаемый выигрыш $\bar{a}_i$:

$\bar{a}_i = \sum_{j=1}^{n} a_{ij}\ p_j$, $i=\overline{1,m}$. Наилучшей, по Байесу, будет стратегия $\bar{A}_i$, соответствующая наибольшему ожидаемому выигрышу $\bar{a}_{io}$

$\bar{a}_{io} = max_{1\le i\le m}\ \bar{a}_I = max_{1\le i\le m} \sum_{j=1}^{n} a_{ij}\ p_j$.

Пример 1.

Небольшое частное предприятие выпекает диетические хлебобулочные изделия. Оборудование позволяет выпекать 500, 600 и 700 кг изделий в день. Спрос на данный вид продукции также может составлять 500, 600 или 700 кг в день. Если хлебобулочные изделия не продаются в этот день, то они возвращаются на предприятие для переработки. Затраты на производство 1 кг изделий составляют 2 тыс. руб., а цена реализации – 3,5 тыс. руб. Дополнительные затраты в случае возврата составляют 1 тыс. руб. на 1 кг изделий.

Необходимо определить ежедневный объём выпечки диетических хлебобулочных изделий.

Решение. В этой ситуации можно выделить две стороны: менеджер предприятия, который должен принять решение об объёме производства, действующий сознательно, и спрос на хлебобулочные изделия, который не является сознательно действующим противником. Ситуацию можно назвать конфликтной, так как результаты действий одной стороны зависят от действия другой стороны, не всегда благоприятны для первой.

В данном примере один игрок – менеджер предприятия. Его возможные действия (стратегии): запланировать выпечку хлебобулочных изделий в объёме 500, 600 и 700 кг. Второй игрок – спрос на хлебобулочные изделия (природа). Его возможные действия: установить спрос на хлебобулочные изделия в объёме 500, 600 и 700 кг.

Рассчитаем платёжную матрицу. Она будет иметь размерность 3×3, так как игрок, принимающий решение, имеет три стратегии ($А_1$ – объём выпечки 500 кг, $А_2$ – 600 кг, $А_3$ – 700 кг), и второй игрок (природа) имеет три стратегии ($П_1$ – спрос составим 500 кг, $П_2$ – 600 кг, $П_3$ – 700 кг). Элементу платежной матрицы a_{11} соответствует стратегии $А_1$ и $П_1$. Это значит, что предприятие выпечет 500 кг хлебобулочных изделий, и спрос на них определится в объёме 500 кг, т. е. все изделия будут реализованы в тот же день. Тогда прибыль предприятия составит

$$(3{,}5 - 2)500 = 750 \text{ тыс. руб.},$$

т. е. a_{11} = 750 тыс. руб.

Далее рассчитаем элемент платёжной матрицы a_{12}. Ему соответствует $А_1$, $П_2$, т. е. предприятие выпечет 500 кг хлебобулочных изделий, а спрос на них определится в объёме 600 кг. Таким образом, все изделия проданы и прибыль предприятия составит

$$(3{,}5 - 2)500 = 750 \text{ тыс. руб.},$$

т. е. a_{12} = 750 тыс. руб.

Аналогично определим a_{13} = 750 тыс. руб.

Рассчитаем элемент платежной матрицы a_{21}. Предприятие выпечет 600 кг хлебобулочных изделий, а спрос на них определится в объеме 500 кг; 100 кг хлебобулочных изделий будет возращено на переработку. Тогда прибыль предприятия рассчитывается следующим образом:

$$(3{,}5 - 2)500 + (-2-1)100 = 750 - 300 = 450 \text{ тыс. руб.},$$

т. е. a_{21} = 450 тыс. руб.

Аналогично рассчитываются все остальные элемент платёжной матрицы.

В результате платёжная матрица имеет вид

$$А = \left(\begin{array}{c|ccc} & П_1 & П_2 & П_3 \\ \hline А_1 & 750 & 750 & 750 \\ А_2 & 450 & 900 & 900 \\ А_3 & 150 & 600 & 1050 \end{array}\right).$$

Так как в данном примере отсутствует информация о вероятностях, с которыми реализуются стратегии природы, то имеем ситуацию неопределённости. Для выбора наилучших стратегий воспользуемся приведенными выше критериями.

1. Критерий Вальда. Для каждой из стратегий выберем наименьший выигрыш. Для стратегии $А_1$ все состояния природы равнозначны, поэтому условно будем считать, что наименьшим выигрышем принимающего решение будет прибыль 750 тыс. руб., т. е., a_1 = 750 тыс. руб. Для стратегии $А_2$, наихудшим будет состояние природы $П_1$, а наименьшим выигрышем a_2 = 450 тыс. руб. Для стратегии $А_3$ наименьшим выигрышем будет a_3 = 150 тыс. руб. Запишем наименьшие выигрыши принимающего решения в дополнительный столбец платежной матрицы

$$A = \left(\begin{array}{lccc|c} & \text{П}_1 & \text{П}_2 & \text{П}_3 & \text{a}_i \\ \hline \text{А}_1 & 750 & 750 & 750 & \mathbf{750} \\ \hline \text{А}_2 & 450 & 900 & 900 & 450 \\ \text{А}_3 & 150 & 600 & 1050 & 150 \end{array}\right).$$

Далее из наименьших выигрышей принимающий решение выбирает наибольший, т. е. $\text{a}_1 = 750 = \max\{\text{a}_1, \text{a}_2, \text{a}_3\} = \max\{750, 450, 150\}$. Наибольший из наименьших выигрышей соответствует стратегии А_1. Это будет наилучшая стратегия по критерию Вальда.

Таким образом, если руководствоваться принципом крайнего пессимизма (критерием Вальда), то следует выпекать 500 кг хлебобулочных изделий диетических сортов в сутки. При этом прибыль предприятия будет не меньше 750 тыс. руб. при любом спросе.

2. Критерий Сэвиджа. Рассчитаем риск для каждой пары стратегии природы и принимающего решение. Если бы менеджер предприятия точно знал, что природа реализует своё состояние П_1, т. е. спрос составит 500 кг, то он бы выбрал стратегию А_1; при этом предприятие получило бы прибыль 750 тыс. руб. – наибольшую для состояния природы П_1, $\beta_1 = 750$. Для состояния природы П_2 наибольшая прибыль равна $\beta_2 = 900$, а для состояния природы П_3 – $\beta_3 = 1050$. По определению, для стратегии А_1 и состояния природы П_1 риск r_{11} составит $\beta_1 - \text{a}_{11} = 750 - 750 = 0$, для стратегии А_2 и состояния природы П_1 риск r_{21} равен $\beta_1 - \text{a}_{21} = 750 - 450 = 300$, и т. д.

Получаем матрицу рисков

$$R = \left(\begin{array}{l|ccc} & \text{П}_1 & \text{П}_2 & \text{П}_3 \\ \hline \text{А}_1 & r_{11} & r_{12} & r_{13} \\ \text{А}_2 & r_{21} & r_{22} & r_{23} \\ \text{А}_3 & r_{31} & r_{32} & r_{33} \end{array}\right) = \left(\begin{array}{l|ccc} & \text{П}_1 & \text{П}_2 & \text{П}_3 \\ \hline \text{А}_1 & 0 & 150 & 300 \\ \text{А}_2 & 300 & 0 & 150 \\ \text{А}_3 & 600 & 300 & 0 \end{array}\right).$$

Далее принимающий решение выбирает для каждой стратегии максимальный риск. Для стратегии А_1 максимальным будет риск, равный 300, т. е. $r_1 = 300$. Аналогично $r_2 = 300$; $r_3 = 600$. В матрицу рисков добавляем столбец, содержащий максимальный риск для каждой стратегии

$$R = \left(\begin{array}{lccc|c} & \text{П}_1 & \text{П}_2 & \text{П}_3 & r_i \\ \hline \text{А}_1 & 0 & 150 & 300 & 300 \\ \hline \text{А}_2 & 300 & 0 & 150 & 300 \\ \hline \text{А}_3 & 600 & 300 & 0 & 600 \end{array}\right).$$

Из максимальных рисков принимающий решение выбирает минимальный:
$r_1 = r_2 = \min\{r_1, r_2, r_3\} = \min\{300, 300, 600\}$, т. е. минимальному из максимальных рисков соответствуют и первая, и вторая стратегии. Наилучшими стратегиями по критерию Сэвиджа будут А_1 и А_2.

3. Критерий Гурвица. Пусть в данном примере принимающий решение в равной мере оптимист и пессимист и использует критерий Гурвица, в котором $\gamma = 1/5$.

Для каждой стратегии A_i рассчитаем число a_i ($i = \overline{1,3}$):

$$a_1 = 1/5 \cdot 750 + 4/5 \cdot 750 = 750;$$
$$a_2 = 1/5 \cdot 450 + 4/5 \cdot 900 = 810;$$
$$a_3 = 1/5 \cdot 150 + 4/5 \cdot 1050 = 870;$$
$$a_3 = \max\{a_1, a_2, a_3\} = \max\{750, 810, 870\} = 870.$$

Числу $a_3 = 870$ соответствует стратегия A_3, т. е. при таком выборе параметра γ наилучшим вариантом по Гурвицу является выпечка 700 кг хлебобулочных изделий.

Таким образом, в данном примере лучшей по всем критериям будет первая стратегия. Однако в некоторых задачах разные критерии могут рекомендовать различные стратегии. Это объяснятся неопределенностью ситуации. В этом случаи можно провести дополнительные исследования. И хотя использование игры с природой при принятии решений в условиях неопределенности не всегда даёт однозначный результат, принимающий решение упорядочивает данные, определяет состояние природы и свои возможные решения, оценивает потери и выигрыши для различных вариантов, что способствует повышению качества принимаемых решений.

Пример 2.

Пусть в примере 1 спрос на диетические хлебобулочные изделия в объеме 500 кг устанавливается с вероятностью $p_1 = 1/5$, в объёме 600 кг – с вероятностью $p_2 = 3/5$ и в объёме 700 кг – с вероятностью $p_3 = 1/5$.

Определить ежедневный объём выпечки хлеба.

Решение. Так как в данном примере известны вероятности стратегий природы, то для выбора наилучшей стратегии воспользуемся критерием Байеса.

Для каждой стратегии A_i рассчитаем ожидаемую прибыль $\overline{a_i}$ ($i = \overline{1,3}$):

$$\overline{a_1} = 1/5 \cdot 750 + 3/5 \cdot 750 + 1/5 \cdot 750 = 750;$$
$$\overline{a_2} = 1/5 \cdot 450 + 3/5 \cdot 900 + 1/5 \cdot 400 = 810;$$
$$\overline{a_3} = 1/5 \cdot 150 + 3/5 \cdot 600 + 1/5 \cdot 1050 = 600.$$

Согласно критерию Байеса, наилучшей будет стратегия, соответствующая наибольшему ожидаемому выигрышу:

$$\overline{a_2} = \max\{\overline{a_1}, \overline{a_2}, \overline{a_3}\} = \max\{750, 810, 600\} = 810.$$

При таких вероятностях спроса на диетические хлебобулочные изделия наилучшей, по Байесу, будет вторая стратегия: предприятию следует выпекать

600 кг хлебобулочных изделий в день, и тогда ожидаемая прибыль предприятия составит 810 тыс. руб.

В платёжную матрицу добавим столбец $\overline{a_i}$

$$A = \left(\begin{array}{c|ccc|c} & p_1 = 1/5 & p_2 = 3/5 & p_3 = 1/5 & \\ & \text{П}_1 & \text{П}_2 & \text{П}_3 & \overline{a_i} \\ \hline A_1 & 750 & 750 & 750 & 750 \\ A_2 & 450 & 900 & 900 & 450 \\ \hline A_3 & 150 & 600 & 1050 & 150 \end{array}\right).$$

В платёжной матрице подчеркнём строку, соответствующую наибольшей ожидаемой прибыли и наилучшей, по Байесу, стратегии.

Кроме ожидаемого выигрыша принимающий решение может рассчитать его вариацию для каждой стратегии. Обозначим вариацию выигрыша для стратегии A_i через V_i, $i = \overline{1,3}$.

Тогда

$V_1 = 1/5(750 - 750)^2 + 3/5(750 - 750)^2 + 1/5(750 - 750)^2 = 0;$

$V_2 = 1/5(810 - 450)^2 + 3/5(810 - 900)^2 + 1/5(810 - 900)^2 = 36\,400;$

$V_3 = 1/5(600 - 150)^2 + 3/5(600 - 600)^2 + 1/5(600 - 1050)^2 = 81\,000.$

Самую большую вариацию имеет третья стратегия, следовательно, она самая рискованная. Наименее рискованной является первая стратегия. Её риск равен 0. Но и ожидаемый выигрыш несколько меньше, чем у второй стратегии. Принимающему решение придется выбирать: либо стратегию A_1 с нулевым риском и несколько меньшим ожидаемым выигрышем, либо стратегию A_2, с несколько большим ожидаемым выигрышем и значительно большим риском. Выбор будет зависеть от склонности к риску принимающего решение.

ЛИТЕРАТУРА

1. Абчук, В. Н. Экономико-математическое моделирование / В. Н. Абчук. – СПб., 1999. – 310 с.

2. Балашевич, В. А. Экономико-математическое моделирование производственных систем / В. А. Балашевич, Н. М. Андронов. – Мн., 1999. – 210 с.

3. Вардомацкая, Е. Ю. Компьютерные информационные технологии. Конспект лекций. – Витебск : УО «ВГТУ», 2017. – 115 с.

4. Винстон, У. Л. Microsoft Excel. Анализ данных и построение бизнес-моделей / У. Л. Винстон; под общ. ред. Ю. П. Леоновой. – Москва: Русская редакция, 2005. – 576 с.

5. Мур, Дж. Экономическое моделирование в Microsoft Excel : пер. с англ. / Дж. Мур, Л. Уэдерфорд. – 6-е изд. – Москва : Издательский дом «Вильямс», 2004. – 1024 с.

6. Холод, Н. И. Экономико-математические методы и модели : учеб. пособие / Н. И. Холод [и др.] ; под. общ. ред. А. В. Кузнецова. – Минск : БГЭУ, 2000. – 412 с.

7. Шарстнев, В. Л. Компьютерные информационные технологии. Пакеты прикладных программ для моделирования и анализа задач экономики / В. Л. Шарстнев, Е. Ю. Вардомацкая. – Витебск, 2007.

8. Шарстнев, В. Л. Компьютерные информационные технологии: лабораторный практикум : пособие / В. Л. Шарстнев, Е. Ю. Вардомацкая. – Витебск : УО «ВГТУ», 2008. – 170 с.

9. Экономико-математические методы и модели : практикум / С. Ф. Миксюк [и др.]; под ред. С. Ф. Миксюк. – Минск : УО «БГЭУ», 2008. – 311 с.

10. Юферева, О. Д. Экономико-математические методы. – Минск : БГЭУ, 2002. – 103 с.

Учебное издание

Вардомацкая Елена Юрьевна

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ В ЛОГИСТИКЕ

Конспект лекций

Редактор *Т.А. Осипова*
Корректор *А.В. Пухальская*
Компьютерная верстка *Е.Ю. Вардомацкая*

Подписано к печати_______. Формат 60x90 1/16. Усл. печ. листов 4,9.
Уч.- изд. листов 5,9. Тираж 25 экз. Заказ № .

Учреждение образование «Витебский государственный технологический университет»
210038, Витебск, Московский пр-т, 72.
Отпечатано на ризографе учреждения образования
«Витебский государственный технологический университет».
Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий № 1/172 от 12 февраля 2014 г.
Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий № 3/1497 от 30 мая 2017 г.

Е. Ю. Вардомацкая

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ В ЛОГИСТИКЕ

Конспект лекций

Витебск
2020