

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций и стохастического анализа

**СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ДАННЫХ И ИХ
ПРИМЕНЕНИЕ НА ПРИМЕРЕ АНАЛИЗА ООО ВОЛГАТОРГ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

Студентки 2 курса 218 группы
направления 01.4.02 — Прикладная математика и информатика
механико-математического факультета
Горшеневой Марии Дмитриевны

Научный руководитель

доцент, к. ф.-м. н.

А. В. Шаталина

Заведующий кафедрой

д. ф.-м. н.

С. П. Сидоров

Саратов 2017

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Методы экономического анализа. Статистические методы анализа. Применение корреляционно-регрессионного метода для анализа де- ятельности сети универсамов «Семейный» — ООО Волгаторг	7
ЗАКЛЮЧЕНИЕ	12
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	13

ВВЕДЕНИЕ

Переход к рыночной экономике требует от предприятий повышения эффективности производства, конкурентоспособности продукции и услуг на основе внедрения достижений научно-технического прогресса, эффективных форм хозяйствования и управления производством, преодоления бесхозяйственности, активизации предпринимательства, инициативы и т.д.

Важная роль в реализации этой задачи отводится экономическому анализу деятельности субъектов хозяйствования. С его помощью вырабатываются стратегия и тактика развития предприятия, обосновываются планы и управленческие решения, осуществляется контроль за их выполнением, выявляются резервы повышения эффективности производства, оцениваются результаты деятельности предприятия, его подразделений и работников.

Современное состояние анализа можно охарактеризовать как довольно основательно разработанную в теоретическом плане науку. Ряд методик, созданных научными работниками, используются в управлении производством на разных уровнях. Вместе с тем наука находится в состоянии развития. Ведутся исследования в области более широкого применения математических методов, ЭВМ, позволяющих оптимизировать управленческие решения. Идет процесс внедрения теоретических достижений отечественной и зарубежной науки в практику. Перспективы развития анализ хозяйственной деятельности в теоретическом направлении тесно связаны с развитием в первую очередь математики, статистики, бухгалтерского учета и др.

Деятельность современных предприятий предопределяет необходимость проведения комплексного экономического анализа, связанного с обработкой большого объема информации. Комплексный экономический и финансовый анализ, предполагающий построение системы показателей, отражающей все существенные стороны деятельности организации, является базой для принятия решений, разработки финансовой политики любого экономического субъекта.

Для анализа используют множество различных методов. Статистические методы — основное средство исследования массовых повторяющихся явлений. Они применяются при возможности представления изменения анализируемых показателей как случайного процесса. Если связь между анализируемыми характеристиками не детерминированная, а стохастическая, то статистические и

вероятностные методы становятся практически единственным инструментом исследования.

В экономических исследованиях часто решают задачу выявления факторов, определяющих динамику экономического процесса. Такая задача чаще всего решается методами корреляционного и регрессионного анализа. Для достоверного отображения объективно существующих в экономике процессов необходимо выявить существенные взаимосвязи и не только выявить, но и дать им количественную оценку. Этот подход требует вскрытия причинных зависимостей. Под причинной зависимостью понимается такая связь между процессами, когда изменение одного из них является следствием изменения другого.

Ввиду того, что предприятие само по себе является сложной социально-экономической системой, функционирующей в условиях неопределенности, получение комплексных оценок деятельности предприятия является сложной задачей не только с экономической, но и с математической точки зрения. Как следствие, возрастает актуальность разработки экономико-математических моделей, которые могут позволить осуществлять анализ финансового состояния предприятия с достаточной точностью и надежностью.

Использование математических методов при анализе, оценке и прогнозировании предоставляет широкий спектр возможностей для организации. Актуальность темы обусловлена рядом преимуществ, которые получает организация при использовании статистических методов для анализа деятельности. При освоении методов организация имеет возможность применять их для повышения эффективности производства, конкурентоспособности продукции и услуг. Внедрение достижений научно-технического прогресса — всегда большой шаг к развитию.

Появление программных продуктов и пакетов, позволяющих, не обладая навыками программиста, с простотой выполнять сложные математические вычисления огромных массивов данных на ЭВМ, позволяют широко использовать статистические методы на любых предприятиях, особенно актуально для предприятий, где оборот данных велик. В данной работе рассматриваются способы использования для вычислений и исследований таких программных продуктов как Mathcad, Excel, Gretl. Освоение методов, позволяющих использовать подобные пакеты, автоматизирующие и в тысячи раз ускоряющие вы-

числения и анализ, даёт возможность широко пользоваться статистическими исследованиями. В перспективе внедрение таких систем и методов даёт большой толчок в развитии локального предприятия, а при их массовом использовании повышает экономические показатели страны.

В первой главе «Методы экономического анализа» выполнен обзор наиболее популярных, удобных и практически полезных современных методов анализа деятельности организации, осуществлена их классификация и упорядочивание. Большая часть исследования в первой главе посвящена статистическим методам анализа.

Вторая глава «Регрессионный и корреляционный анализ» описывает один из статистических методов обработки информации. Раздел рассматривает корреляционно-регрессионный анализ данных. Здесь приводятся формулы построения различных моделей, описываются коэффициенты и параметры для оценивания качества модели и приводится описание возможностей прогнозирования с использованием результатов корреляционно-регрессионного анализа.

В третьей главе «Применение корреляционно-регрессионного метода для анализа деятельности сети универсамов «Семейный» — ООО Волгаторг» рассматривается использование методов, описанных во второй главе, на практике. В разделе рассмотрен алгоритм применения модели корреляционно-регрессионного анализа для ООО Волгаторг. Проводится исследование взаимосвязей кадровых факторов и выручки подразделений методами программных продуктов Mathcad и Gretl, сравнение результатов исследований в данных пакетах.

Основная цель настоящей работы – систематизировать и структурировать информацию о методах статистического анализа данных, показать преимущества и недостатки регрессионного анализа применительно к оценке состояния и анализу показателей и взаимосвязей факторов для организации, предложить схему и алгоритм регрессионно-корреляционного исследования факторов, продемонстрировать процесс применения методов с использованием программных пакетов. При этом проводится сравнение результатов исследований в разных пакетах.

Информационную основу составили работы зарубежных и российских практиков в области моделирования, анализа и прогнозирования, представлен-

ные в научных монографиях, ведущих периодических изданиях по вопросам аналитики.

1 Методы экономического анализа. Статистические методы анализа. Применение корреляционно-регрессионного метода для анализа деятельности сети универсамов «Семейный» — ООО Волгаторг

В процессе экономического анализа, аналитической обработки экономической информации применяется ряд специальных способов и приемов. Существует различная классификация методов и приемов экономического анализа, зависящая от исследуемых признаков.

Необходимость применения статистических методов (методов многомерного статистического анализа) связана с существованием важной особенности реальных экономических систем, которая практически не учитывается в других системах оценки состояния предприятия. Выводы статистического анализа составляют существенный компонент системы поддержки принятия стратегических решения.

Особенностью методов статистического анализа является их комплексность, обусловленная многообразием форм статистических закономерностей, а также сложностью процесса статистических исследований.

Статистической наукой разработаны методы, с помощью которых можно измерить связь между явлениями, не используя при этом количественные значения признака, а значит, и параметры распределения. Статистические методы обеспечивают проведение глубокого и детализированного анализа.

Статистические методы — основное средство исследования массовых повторяющихся явлений. Они применяются при возможности представления изменения анализируемых показателей как случайного процесса. Если связь между анализируемыми характеристиками не детерминированная, а стохастическая, то статистические и вероятностные методы становятся практически единственным инструментом исследования.

В экономическом анализе наиболее известны методы множественного и парного корреляционного анализа. Для изучения одновременных статистических совокупностей служат закон распределения, вариационный ряд, выборочный метод.

Статистические методы включают использование средних и относительных величин, индексный метод, корреляционный и регрессионный анализ, построение вариационных рядов, статистическую группировку и др.

Метод корреляционного и регрессионного (стохастического) анализа ши-

роко используется для определения тесноты связи между показателями, не находящимися в детерминированной функциональной зависимости. Данный метод является пограничным между статистическими и математическими методами.

Корреляционный анализ тесно связан с регрессионным анализом (также часто встречается термин «корреляционно — регрессионный анализ», который является более общим статистическим понятием), с его помощью определяют необходимость включения тех или иных факторов в уравнение множественной регрессии, а также оценивают полученное уравнение регрессии на соответствие выявленным связям (используя коэффициент детерминации, определение которого будет введено позже).

Корреляционный анализ является статистическим методом, который решает следующие задачи:

1. Выявление связи между переменными путем точечной и интервальной оценки парных (частных) корреляций, вычисление и проверка значимости множественных коэффициентов корреляции и детерминации;
2. Отбор факторов, оказывающих наиболее существенное влияние на результативный признак, на основании измерения степени связи между ними.

При машинной обработке исходной информации на ЭВМ, оснащенных пакетами стандартных программ для ее анализа, вычисление параметров применяемых математических функций является быстро выполняемой счетной операцией.

Формально корреляционная модель взаимосвязи системы случайных величин $X = (X_1, X_2 \dots X_k)$ может быть представлена в следующем виде:

$$X = f(X, Z)$$

где Z – набор случайных величин, оказывающих влияние на изучаемые случайные величины.

Коэффициент корреляции служит для определения статистической взаимосвязи между двумя или несколькими переменными. Этот показатель варьируется от -1 до 1. Оценка коэффициента корреляции высчитывается по следующей формуле:

$$R(x, y) = \frac{cov(x, y)}{\sqrt{D_x D_y}},$$

$$cov(x, y) = \sum x_i y_i - \frac{1}{n} (\sum x_i) (\sum y_i),$$

$$D(x) = \sum x_t^2 - \frac{1}{n} (\sum x_t)^2,$$

Если $0,8 \leq |R| \leq 1$, то связь считается сильной.

Если $0,5 < |R| < 0,8$, то связь считается умеренной. Связь слабая или отсутствует в случае, если $|R| < 0,5$. Если выявленные линейные статистические зависимости окажутся сильными или умеренными, то целесообразно будет найти их математическое выражение в виде регрессионной модели и оценить статистическую значимость модели с помощью коэффициента детерминации R^2 .

Коэффициент детерминации R^2 — оценка качества уравнения регрессии:

$$R^2 = 1 - \frac{\sum_{i=0}^n (y_i - \hat{y}_i)}{\sum_{i=0}^n (y_i - y)}$$

где y_i — наблюдаемое значение зависимой переменной y , \hat{y}_i — значение зависимой переменной, предсказанное по уравнению регрессии, y — среднее арифметическое зависимой переменной.

В случае, если $R^2 > 0,8$, модель статистической связи, как правило, используется для прогнозирования изучаемого явления или показателя.

Корреляция (корреляционная зависимость) — статистическая взаимосвязь двух или нескольких случайных величин (либо величин, которые можно с некоторой допустимой степенью точности считать таковыми). При этом, изменения значений одной или нескольких из этих величин приводят к систематическому изменению значений другой или других величин.

Математической мерой корреляции двух случайных величин служит коэффициент корреляции. В случае, если изменение одной случайной величины не ведёт к закономерному изменению другой случайной величины, но приводит к изменению другой статистической характеристики данной случайной величины, то подобная связь не считается корреляционной, хотя и является статистической.

Линейная модель парной регрессии.

Пусть есть набор значений двух переменных $X_t, Y_t, t = 1, \dots, n$; ; можно отобразить пары (X_t, Y_t) точками на плоскости $X-Y$

Предположим, что шей задачей является подобрать функцию $Y = f(X)$ из параметрического семейства функций $f(X, B)$, «наилучшим» способом описывающую зависимость Y от X . Подобрать функцию в данном случае означает выбрать “ наилучшее» значение параметра B .

Теперь добавим к постановке задачи некоторые статистические свойства данных, так как для одного X мы можем наблюдать разные значения Y .

Запишем уравнение зависимости Y_t от X_t в виде $Y_t = a + bX_t + \epsilon_t$, $t = 1, \dots, n$, где X_t — неслучайная (детерминированная) величина, а Y_t — зависимая величина, ϵ_t — случайные величины; Y_t называется объясняемой (зависимой) переменной, а X_t — объясняющей (независимой) переменной или регрессором, а само уравнение называется регрессионным уравнением.

Коэффициент детерминации служит для оценки качества регрессии и находится в интервале от 0 до 1. Если $R^2 = 0$, то это означает, что регрессия ничего не дает, т.е. x_t не улучшает качество предсказания y_t по сравнению с тривиальным предсказанием $\hat{y}_t = y$. Другой крайний случай, когда $R = 1$, означает точную подгонку: все точки наблюдений лежат на регрессионной прямой (все $\epsilon_t = 0$). То есть, чем ближе к 1 значение R^2 , тем лучше качество “подгонки» прямой. Однако значимость коэффициента детерминации проявляется лишь на промежутке обучающей выборки и не в коем случае не гарантирует сохранения тенденции в будущем.

В работе рассматривается корреляционно-регрессионный анализ ООО Волгаторг. Для анализа используются средства программных продуктов Mathcad, Excel, Gretl.

ООО Волгаторг — это крупная торговая сеть, работающая на рынке Саратовской области с 2003 года. В настоящее время число универсамов, работающих под вывеской «Семейный» насчитывает более 170, а общий штат сотрудников превышает 3 000 человек.

В базах данных организации ООО Волгаторг хранятся огромные массивы информации. Для анализа была получена следующая информация:

- о сотрудниках организации всех подразделений за всё время существования организации

- о годовых продажах подразделений за последние 3 года.

Под подразделениями понимаются магазины торговой сети и отделы главного офиса организации

Была исследована зависимость стажа заместителя управляющего магазина от годовой выручки магазина. Эти данные позволят оценить взаимосвязь «текучки» кадров на должности от эффективности работы подразделения. Результаты анализа помогут предприятию:

- в составлении картины причинно-следственных связей роста и падения эффективности работы подразделений организации.
- корректировать кадровую политику организации.

По полученным данным в Mathcad был построен график зависимости выручки от стажа заместителя управляющего.

Для вычисления коэффициента корреляции имеется встроенная функция *corr*. Вычисление коэффициента корреляции показало значение 0,617. Коэффициент имеет значение выше уровня 0,5, это означает, что корреляция между признаками прослеживается.

Чтобы попробовать выявить возможную зависимость данных, с помощью стандартной функции Mathcad *line* проводится линейризация данных графика по стажу управляющего и выручке подразделения. Строя график по данным точкам получаем результат.

Для сравнения результатов анализа этот пример был так же проанализирован в эконометрической библиотеке Gretl.

Таким образом, используя данные из баз данных, организация имеет возможность рассматривать зависимости для поиска коррелирующих факторов и учета их в своей деятельности.

ЗАКЛЮЧЕНИЕ

Использование математических методов при анализе, оценке и прогнозировании предоставляет широкий спектр возможностей для организации. Из исследований, проведенных в работе, можно сделать вывод, что объем, точность и качественные характеристики исходной информации являются ключевыми факторами качественного регрессионного моделирования.

В данной работе были рассмотрены общие принципы анализа деятельности организации, построения математических моделей и оценки их параметров, представлены основные требования к исходной информации для построения моделей, рассмотрен алгоритм применения модели анализа на практике.

Были систематизированы современные методы анализа деятельности организации, осуществлена их классификация и упорядочивание, рассмотрены построения различных моделей, описываются коэффициенты и параметры для оценивания качества модели и приводится описание возможностей прогнозирования с использованием результатов корреляционно-регрессионного анализа.

Важным результатом является освещение использования методов, описанных во второй главе, на практике — рассмотрен алгоритм применения модели корреляционно-регрессионного анализа для ООО Волгаторг. Проведено исследование взаимосвязей кадровых факторов и выручки подразделений методами программных продуктов Mathcad и Gretl, сравнение результатов исследований в данных пакетах.

Таким образом, в настоящей работе были показаны преимущества и недостатки регрессионного анализа, применительно к оценке состояния и анализу показателей и взаимосвязей факторов для организации, проведено освоение методов, позволяющих использовать программные продукты, автоматизирующие и ускоряющие вычисления и анализ, что позволит организациям и предприятиям широко пользоваться статистическими исследованиями.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Абрютин, М.С., Анализ финансово-экономической деятельности предприятия / М.С. Абрютин, А.В. Грачев. – М.: Делон Сервис, 1998. – 264с.
2. Афанасьев, В.Н. Анализ временных рядов и прогнозирование: Учебник / В.Н. Афанасьев, М.М. Юзбашев. – М.: Финансы и статистика, 2001. – 228с.
3. Бенькович, Е.С. Практическое моделирование динамических систем / Е.С. Бенькович. – СПб, 2002. – 338с.
4. Васильков, Ю.В. Компьютерные технологии вычислений в математическом моделировании: Учеб. пособие. / Ю.В. Васильков, Н.Н. Василькова. – М.: Финансы и статистика, 2002. – 256с.
5. Герасимов, Б.И. Экономический анализ: Учеб. пособие / Б.И. Герасимов, Т.М. Коновалова, С.П. Спиридонов. – Тамбов: ТГТУ, 2006. – 64с.
6. Грищенко, О.В. Анализ и диагностика финансово-хозяйственной деятельности предприятия. – Режим доступа: <http://www.aup.ru/books/m67>
7. Дисперсионный анализ. – Режим доступа: <http://www.statsoft.ru/home/textbook/modules/stanman.html>
8. Елисеева, И.И. Общая теория статистики: Учебник / И.И. Елисеева, М.М. Юзбашев. – 3-е изд. – М.: Финансы и статистика, 1998. – 368с.
9. Ефимова, М.Р. Общая теория статистики / М.Р. Ефимова, Е.В. Петрова, В.Н. Румянцев. – М.:ИНФРА-М, 1996. – 377с.
10. Ипатова, Э. Р. Практикум по проектированию информационных систем: Учеб. пособие / Э. Р. Ипатова, Ю.В. Ипатов. – Магнитогорск: МаГУ, 2004. – 116с.
11. Ипатова, Э. Р. Введение в информационные системы: Учеб. пособие / Э. Р. Ипатова. – Магнитогорск: МаГУ, 2002. – 127с.
12. Ипатова, Э. Р. Проектирование информационных систем: Учеб. пособие / Э. Р. Ипатова, Ю.В. Ипатов. – Магнитогорск: МаГУ, 2005. – 187с.
13. Кленин, А.Н. Математическая статистика для экономистов-статистиков / А.Н. Кленин, К.К. Шевченко. – М.: 1990.
14. Уотшем, Дж. Количественные методы в финансах : Учеб. пособие / К. Паррамоу. – М.: ЮНИТИ, 1999;

15. Бендат, Дж. Прикладной анализ случайных данных / А. Пирсол. – М.: Мир, 1989. – 540 с.
16. Отнес, Р. Прикладной анализ временных рядов / Л. Эноксон – М.: Мир, 1982. – 428 с.
17. Гурский, Д.А. Вычисления в Mathcad 12 / Е.С Турбина – М.: Питер 2006. – 544с.
18. Брандт. З. Статистические методы анализа наблюдений / М.: «Мир 1975. – 313 с.
19. Берман Н.Д. Основы работы в MathCAD 15: Учебное пособие/ М.: Изд-во Тихоокеан. гос. ун-та, 2015. – 116 с.
20. Закс Л. Статистическое оценивание / М.: Статистика, 1976. – 598 с.