

Н. Б. ОСИПЕНКО, А. А. СЛУКА  
Математический факультет,  
кафедра математических проблем управления

## **ПОВЫШЕНИЕ КАЧЕСТВА УСВОЕНИЯ КОРРЕЛЯЦИОННО-РЕГРЕССИОННОГО АНАЛИЗА ЗА СЧЕТ ЕГО ВЫПОЛНЕНИЯ В РАЗНЫХ ПРОГРАММНЫХ СРЕДАХ**

Обработка статистических данных уже давно применяется в самых разнообразных видах человеческой деятельности: экономике, медицине, социологии и т. д. Всесторонний и глубокий анализ информации, так называемых статистических данных, предполагает использование различных специальных методов, важное место среди которых занимает корреляционный и регрессионный анализы обработки статистических данных [1].

Основными задачами корреляционного анализа являются проверка статистических гипотез о наличии и силе корреляционной связи и оценка силы связи. Не все факторы, влияющие на процессы, являются случайными величинами, поэтому при анализе явлений обычно рассматриваются связи между случайными и неслучайными величинами. Такие связи называются регрессионными, а метод математической статистики, их изучающий, называется регрессионным анализом.

Использование возможностей современной вычислительной техники, оснащенной пакетами программ машинной обработки статистической информации на ЭВМ, делает практически осуществимым оперативное решение задач изучения взаимосвязи показателей методами корреляционно-регрессионного анализа. При машинной обработке исходной информации на ЭВМ, оснащенных пакетами стандартных программ ведения анализов, вычисление параметров применяемых математических функций является быстро выполняемой счетной операцией.

Однако, опыт преподавания и усвоения вопросов, связанных с возможностями обработки статистических данных и в частности корреляционно-регрессионного анализа показывает, что студенты сталкиваются с многочисленными теоретическими трудностями. Вторая часть проблем связана с необходимостью изучения различных программных сред для решения задачи корреляционно-регрессионного анализа и приводит к практическим проблемам. Особо стоит вопрос об отсутствии описания ограничений и ориентиров различных методов анализа данных в программных системах, что привело к уменьшению эффективности машинных методов анализа данных.

Целью данной работы явилось сравнение возможностей проведения корреляционного и регрессионного анализов в разных программных средах (Excel, MathCad, Statistica, Maple) и разработка электронного пособия по спецкурсу «Программные средства статистической обработки экспериментальных данных».

Приведем краткие результаты проведенного исследования.

**Excel** – программный продукт с богатыми возможностями, который можно использовать на многих уровнях. Электронная таблица Excel [2], вероятно, самая простая в использовании программа. Многие ее возможности были специально разработаны для того, чтобы и начинающие, и профессиональные пользователи могли выполнять повседневные задачи просто и быстро. При выполнении определенных процедур программа ведет пользователя шаг за шагом, основные команды редактирования и форматирования просты и понятны.

В Excel можно использовать более 400 математических, статистических, финансовых и других специализированных функций, связывать различные таблицы между собой, выбирать произвольные форматы представления данных, создавать иерархические структуры.

С помощью пакета Excel пользователь сможет быстро найти коэффициент корреляции, построить корреляционное поле и регрессионную прямую, а также проверить гипотезу о том, что простая линейная регрессия  $Y$  по  $X$  отсутствует. С помощью линейного регрессионного

анализа *Пакета анализа Excel* можно быстро найти коэффициенты уравнения регрессии.

Программа *Excel*, являясь лидером на рынке программ обработки электронных таблиц, определяет тенденции развития в этой области.

**MathCad** – универсальный математический пакет для выполнения инженерных и научных расчетов. Основное его преимущество – естественный математический язык, на котором формируются решаемые задачи. Он имеет удобный для описания аналитических моделей интерфейс, ориентированный на классическую форму языка математики [3]. Программа достаточно проста в освоении для того, чтобы ее применяли не только ученые, инженеры или студенты вуза, но и школьники старших классов.

Так же *MathCad* имеет встроенный язык программирования. С одной стороны это вынужденная мера, поскольку разработчики не могут предвосхитить и обеспечить все потенциальные требования пользователей, а необходимость освоения языка является недостатком программы. С другой стороны, синтаксис языка довольно простой, его использование может существенно расширить круг задач, решаемых исследователями, при этом вспомогательные инструменты, в частности построение графиков всегда у него под рукой.

*MathCad* завоевал популярность во всем мире. Им пользуются в работе свыше 5 млн. человек [3]. Ежегодно выпускаются новые версии *MathCad*, улучшается его интерфейс, расширяются возможности отдельных функций.

**Statistica** – это универсальная интегрированная система, предназначенная для статистического анализа и визуализации данных, содержащая широкий набор процедур анализа для применения в научных исследованиях различных направлений, технике, бизнесе, учебном процессе.

Преимущества *Statistica* перед другими статистическими пакетами, подчеркивающие целесообразность его использования в образовательном процессе следующие [4]: с помощью реализованных в ней языков программирования (*SCL*, *Statistica Basic*), снабженных специальными средствами поддержки, легко создаются законченные пользовательские решения и встраиваются в различные другие приложения или вычислительные среды; переведен на русский язык и можно приобрести лицензионную, русифицированную версию, издано большое число книг с подробным описанием системы *Statistica*; особую актуальность приобретает при изучении студентами трудоемких, сложных с математической точки зрения, и громоздких в реализации методов многомерного анализа;

При выполнении корреляционного и регрессионного анализов в Statistica сложностей у пользователей, как правило, не возникает.

Таким образом, Statistica является одной из наиболее простых для неподготовленного пользователя систем, с наименьшим периодом овладения ее возможностями и удачным набором графических возможностей.

**Maple** – это мощная вычислительная система, предназначенная для выполнения сложных вычислений как аналитическими, так и численными методами. Maple содержит проверенные, надежные и эффективные символьные и численные алгоритмы для решения огромного спектра математических задач [5], умеет выполнять сложные алгебраические преобразования и упрощения.

При проведении корреляционного и регрессионного анализов пакет Maple достаточно сложен в использовании. Для работы с ним необходимо использовать дополнительную литературу, так как Maple имеет большое количество встроенных функций и математических пакетов.

На основе проведенных исследований разработано электронное пособие, которое состоит из трёх блоков: первый блок содержит теоретический материал о корреляционно-регрессионном анализе, второй – практическую часть с описанием возможности проведения корреляционно-регрессионного анализа, третий – образцы примеров и практические задания по лабораторным занятиям. Электронное пособие ориентировано на студентов математического факультета, изучающих спецкурс «Программные средства статистической обработки экспериментальных данных».

В заключение остановимся на основной причине повышения качества преподавания корреляционно-регрессионного анализа за счет использования разных программных сред. При таком подходе, имея дело с одной статистической задачей и используя при этом несколько языков обработки, у студента формируется осознание некоего теоретического инварианта (содержательный смысл корреляционно-регрессионного анализа) за пределами технического манипулирования программным инструментом.

### Литература

- 1 Айвазян, С. А. Прикладная статистика: исследование зависимостей : учебное пособие / С. А. Айвазян. – М. : Финансы и статистика, 1985. – 488 с.
- 2 Фрай, К. Microsoft Office Excel : учебное пособие / К. Фрай. – Питер : ЭКОМ Паблишерз, 2009. – 480с.
- 3 Алябьева, С. В. Mathcad для студентов: справочное издание / С. В. Алябьева [и др.]. – Петрозаводский государственный университет, 2007. – 154 с.

4 Боровиков, В. П. Популярное введение в Statistica: справочное издание / В. П. Боровиков. – М.: КомпьютерПресс, 1998. – 267 с.

5 Прохоров, Г. В. Пакет символьных вычислений Maple : справочное издание / Г. В. Прохоров, М. А. Леденев, Колбеев В. В. – М.: Компания Петит, 1997. – 203 с.