Кафедра Экономической Информатики и Автоматизации Управления.

Ростовский Государственный Экономический Университет

(РИНХ)

**Реферат на**

**тему: «Отечественные статистические пакеты».**

Выполнила Рыбакова О.А.

Группа № 427

Проверила Нельзина О.Г.

Номер зачетки 03146

Специальность Актуарий для банков и страховых компаний

г. Ростов-на-Дону 2004г.

**С О Д Е Р Ж А Н И Е**

**1. Введение**

**2. Общие сведения об СПП.**

**3. Легкость использования пакета и его освоения**

**4. Виды статистических пакетов**

**5. Требования к статистическим пакетам общего назначения**

**6. Пакет STADIA**

**7. История создания системы ЭВРИСТА.**

**8. ППП «ОЛИМП»**

**9. ППП «МЕЗОЗАВР»**

**10. Заключение**

**11. Список литературы**

**В В Е Д Е Н И Е**

## Статистический пакет -*программный продукт*, предназначенный для статистической обработки *данных*.

Являются надежным инструментом повышения качества принимаемых решений. В пакет, как правило, входит: деловая графика, дисперсионный анализ, регрессионный анализ, анализ временных рядов и пр.

Для успешной работы любой организации в условиях рынка и конкуренции, конечно, необходим тщательный анализ имеющейся информации о создании продукции, её сбыте, эксплуатации, а также анализ информации о конкурентах и т. п. Конечно можно положиться на интуицию, но, скорее всего, правильное решение будет принято, только если у вас имеется огромный многолетний опыт в этой сфере деятельности. Но можно (и даже нужно, несмотря на весь опыт) пойти другим путём — воспользоваться специальными средствами для обработки информации. И вот здесь-то и потребуется профессионализм — необходимо правильно выбрать программное обеспечение, корректно ввести информацию, которую вы хотите проанализировать, выбрать методы и алгоритмы для решения именно этой задачи и многое другое.

Можно выделить **2 вида** статистических пакетов.

**Из зарубежных пакетов это** STATGRAPHICS, SPSS, SYSTAT, BMDP,SAS, CSS, STATISTICA, S-plus, и др.

**Из отечественных** можно назвать такие пакеты, как STADIA, ЭВРИСТА, МЕЗОЗАВР, ОЛИМП: Стат-Эксперт, Статистик-Консультант, САНИ, КЛАСС-МАСТЕР и др.

Отечественные статистические пакеты, которые устойчиво представлены на рынке в течение последних лет, в значительной степени лишенытаких **недостатков**, которые есть у западных продуктов. Они предполагают наличие широкого первоначального статистического образования, доступной литературы и консультационных служб. Поэтому они содержат мало экранных подсказок и требуют внимательного изучения документации на английском языке.

**Общие сведения об СПП.**

Здесь излагается методика сравнительного анализа однотипных отечественных пакетов, позволяющая такой продукт правильно выбрать и в дальнейшем эффективно использовать.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Статистическая система** | **Версия** | **Цена** | **Минимальные системные требования** | | | | **Продавец** | **Категория пользователя** |
| **ОС** | **Процессор** | **Емкость НЖМД** | **Объем ОЗУ** |
| Класс-Мастер | 2.0 | 320 | W | 386 | 4 | 4 | Изд-во "ТВП" | Ср |
| Класс-Мастер | 1.0 | 300 | D | 286 | 2 | 1 | Изд-во "ТВП" | Ср |
| Мезозавр | 2.0 | 520 | W | 386 | 4 | 4 | Изд-во "ТВП" | Ср |
| САНИ | 1.0 | 300 | D | 286 | 1 | 1 | Изд-во "ТВП" | В |
| Stadia | 6.0 | 500 | W | 386 | 5 | 4 | Информатика Компьютеры | В-Ср |
| Stadia | 5.0 | 250 | D | 386 | 0,4 | 1 | Информатика&Компьютеры | В-Ср |
| Stat-Media | 1.1 | 250 | D | 286 | 1 | 1 | Полихимэкс | Ср-Н |
| Starc | 4.0 | н/д | D | н/д | н/д | н/д | ТОО Data-Center | В-Ср |
| Квазар | 1.0 | н/д | D | н/д | н/д | н/д | ИММ УрО РАН | В |
| Palmoda | 1.0 | 300 | D | 286 | 2 | 1 | ВЦ РАН | В |
| Лорег | 1.0 | 350 | W | 386 | 10 | 4 | н/д | В |
| PolyAnalyst | 2.0 | 1850 | OS/2 | 486 | н/д | 8 | ЗАО Megaputer | Н |
| Олимп | 1.21 | 500 | W | 386 | н/д | 4 | ЗАО CPS | Ср-Н |
| Олимп | н/д | 350 | D | 286 | н/д | 1 | ЗАО CPS | Ср-Н |
| Статистик-Консультант | 3.0 | н/д | W | 386 | 3 | 4 | ТОО "Тандем" | Ср-Н |
| Ростан | 1.0 | н/д | D | 286 | 2 | 1 | БГУ | В |
| Cтатэкс | 1.0 | н/д | D | 286 | н/д | 1 | РМ и ПК Казахстан | Н |

#### Легкость использования пакета и его освоения

Этот раздел методики построения рейтинга предлагает учитывать качество документации по наличию "Руководства пользователя" и учебных пособий по методам, реализованным в пакете. Кроме того, оценивается удобство интерфейса "человек - компьютер", качество встроенной подсистемы помощи, степень удобства управления данными и работы с графикой и таблицами.

Степень интеллектуальности пакета в первую очередь предполагает организацию такого режима работы СПП, при котором пользователь имеет достаточно квалифицированное *статистическое ассистирование* в ходе всего процесса статистического анализа, т. е. при выяснении природы (генезиса) анализируемых данных, при выборе подходящих моделей и методов, их увязывании в технологическую цепочку, при интерпретации результатов и т. п. При этом основные показатели вовсе не обязательно связаны с наличием в пакете подходящей экспертной системы. Речь идет о развитой системе компьютерной консультационной поддержки (по статистике), охватывающей различные **стадии решения задачи**:

 ориентирование пользователя в существующих литературных источниках по применяемым статистическим методам, а также обеспечение его подсказками по используемой терминологии, понятиям, существующим решениям аналогичных задач;

 помощь в постановке задачи, подробный предварительный анализ исходных данных c акцентированием внимания пользователя на их генезисе и особенностях;

 подбор подходящего вида модели и технологической цепочки обрабатывающих модулей;

 описание набора типичных статистических "ловушек" и способов, как их избежать;

 помощь в интерпретации промежуточных и финальных результатов статистического анализа;

 предложение направлений дальнейшего исследования.

Необходимым условием обеспечения достаточной степени интеллектуальности СПП мы считаем привлечение к разработке пакета (а при оценке готового - учет научного уровня) специалистов по теории и методам статистического анализа данных. Кроме того, представляется важным, как используются интерактивные и графические возможности современного ПК, а также средства, связанные с формализацией и предоставлением опыта и профессиональных знаний специалистов по прикладной статистике.

**Виды статистических пакетов**

Основную часть имеющихся пакетов составляют специализированные пакетов и пакеты общего назначения.

***Специализированные пакеты*** обычно содержат методы из одного - двух разделов статистики или методы, используемые в конкретной предметной области (контроль качества промышленной продукции, расчет страховых сумм и т.д.). Чаще всего встречаются пакеты для анализа временных рядов (например, ЭВРИСТА, МИЗОЗАВР, ОЛИМП: Стат-Эксперт), регрессионного и факторного анализа. Обычно эти пакеты содержат весьма полный набор традиционных методов в своей области, а иногда включают также и оригинальные методы и алгоритмы, созданные разработчиками пакета. Как правило, пакет и его документация ориентированы на специалистов, хорошо знакомых с соответствующими методами.

***Пакеты общего назначения*.** Особое место на рынке занимают так называемые статистические пакеты общего назначения. Широкий диапазон статистических методов, дружелюбный интерфейс пользователя привлекает в них не только начинающих пользователей, но и специалистов. Универсальность этих пакетов особенно полезна:

* + на начальных этапах обработки, когда речь идет о подборе статистической модели или метода анализа данных;
  + когда поведение статистических данных выходит за рамки использовавшейся ранее модели;
  + в процессе обучения основам статистики.

Именно пакеты общего назначения составляют большинство продаваемых на рынке статистических программ. К таким пакетам относятся системы STADIA и SPSS, а также пакеты STATGRAPHICS, STATISTICA, S-plus, и др.

**Требования к статистическим пакетам общего назначения**

Для того чтобы статистический пакет общего назначения был удобен и эффективен в работе, он должен удовлетворять многочисленным и весьма жестким требованиям. В частности, необходимо, чтобы он:

* содержал достаточно полный набор стандартных статистических методов;
* был достаточно прост для быстрого освоения и использования;
* отвечал высоким требованиям к вводу, преобразованиям и организации хранения данных;
* имел широкий набор средств графического представления данных и результатов обработки;
* предоставлял удобные возможности для включения в отчеты таблиц исходных данных, графиков, промежуточных и окончательных результатов обработки;
* имел подробную документацию, доступную для начинающих и информативную для специалистов-статистиков.

Наконец, немаловажное значение имеет цена пакета.

Пакеты, рассчитанные на массового пользователя, стоят дешевле, чем западные – обычно 500-1500 долларов. Эти пакеты отличаются от профессиональных, прежде всего ориентацией на индивидуального пользователя: преимущественно диалоговым режимом работы, наличием ограничений по объему обрабатываемых данных и т.д.

**Отечественные** статистические пакеты стоят существенно дешевле, как правило, их цена составляет от 50 до 300 долларов.

**Пакет STADIA**

***Пакет STADIA*** разработан и поддерживается НПО “Информатика и компьютеры” при активном участи ведущих специалистов МГУ им. М.В.Ломоносова. Пакет содержит широкий набор методов анализа данных из всех областей статистики и доступен широкому кругу прикладных специалистов, менеджеров и студентов. Сейчас распространяется версия 6.2 для среды Windows. Пакет может появляться в трех вариантах: study, basе и рrоf, различающихся лишь объемами обрабатываемых массивов и ценой. Самый дешевый вариант study имеет максимальный объем матрицы данных в 400 чисел. Он предназначен главным образом для учебных заведений и задач с небольшими объемами данных. Самая дорогая версия STADIA 6.2 рrоf. имеет максимальный объем матрицы данных 20000 чисел и расширенные возможности статистических процедур для их обработки по сравнению с базовыми версиями. У пакета имеется бесплатная учебно-демонстрационная версия, позволяющая обрабатывать большое количество демонстрационных примеров из всех разделов статистического анализа. Эта версия также допускает ввод с клавиатуры и полную обработку данных пользователей. Однако при этом существуют ограничения на объемы вводимых данных, и отсутствует возможность сохранения введенных данных в файле. Документация пакета является одновременно детальным справочником по использованию статистических методов и может быть приобретена отдельно от пакета.

**История создания системы ЭВРИСТА.**

Идея создания специализированного статистического пакета по анализу и прогнозированию временных рядов возникла вначале 80-х годов на кафедре математической статистики Московского государственного университета. Главным идеологом будущей программной системы выступил старший научный сотрудник кафедры, к.ф.-м.н. Ю.Г.Баласанов. Первая версия системы ЭВРИСТА была реализована на языке ФОРТРАН для ЭВМ БЭСМ-6 и с 1984 года началось и использование системы в учебном процессе факультета.

Первая коммерческая версия системы ЭВРИСТА для персонального компьютера появилась 1987 году и ее первым покупателем стало объединение КАМАЗ (г. Набережные Челны). Несмотря на то, что первые персональные компьютеры имели слабые (особенно с нынешних позиций) графические возможности, разработчики по максимуму старались их использовать, и в результате ЭВРИСТА, одна из немногих программных систем того времени, уже имела полностью графический многооконный интерфейс.

В 1988 году Ю.Г.Баласанов и его коллеги создают специализированную статистическую группу высокопрофессиональных статистиков-программистов в СП ДИАЛОГ.

В 1991 году выходит вторая версия системы для IBM-совместимых компьютеров.

В январе 1993 года коллектив разработчиков системы ЭВРИСТА в полном составе переходит в ТОО “Центр Статистических Исследований” и все последующие версии системы выходят в рамках этой организации.

В настоящее время система ЭВРИСТА имеет более 500 зарегистрированных пользователей и по праву заслужила звание популярнейшей отечественной специализированной статистической системы по анализу и прогнозированию временных рядов. Среди наших пользователей: Центральный Банк России, Московский Сбербанк, АКБ "Гута-банк", Bank of America, Институт проблем переходного периода, Институт социологии парламентаризма, МГУ им. М.В.Ломоносова и многие другие.

С 1997 “ЭВРИСТА” - зарегистрированная торговая марка ООО “Центр Статистических Исследований” (свидетельство N 148880 комитета Российской федерации по патентам и товарным знакам (РосПатент)). Статистическая система “ЭВРИСТА” зарегистрирована как самостоятельное электронное издание ООО “Центр Статистических Исследований” (свидетельство N0114-97.1.0.RUS Серия Б. Комитета при Президенте Российской Федерации по политике информатизации).

**ППП «ОЛИМП»**

Пакет «Олимп» предназначен для автоматизации обработки данных на основе широкого набора современных методов прикладной статистики. Он реализован в расчете на самых разнообразных пользователей – от новичков до экспертов в области статистики.

В состав пакета, кроме основных программ, входят также электронная таблица MNCALC и программное средство «Прикладные социологические исследования (ПСИ)».

*Пакет «ОЛИМП» позволяет* организовать полный цикл исследований по статистическому анализу и прогнозированию данных, начиная с ввода исходных данных, их проверке и визуализации и заканчивая проведением расчетов и анализом результатов.

*С функциональной точки зрения пакет состоит из* следующих программ (процедур): редактора средств графического отображения и утилит преобразования данных, а также программ реализации методов статистического анализа.

Редактор данных обеспечивает возможность ввода, просмотра и редактирования исходных данных (в том числе пропущенных наблюдений).

Средства графического отображения данных позволяют выводить различные виды графиков на экран, а также сохранять их на диске для дальнейшего использования.

Утилиты преобразования данных выполняют арифметические преобразования данных (унарные и бинарные), различные виды сортировки (в том числе по нескольким переменным), агрегирование (объединение по одному признаку) и фильтрование данных (отбор по одному признаку).

*Программы пакета «ОЛИМП»* реализуют следующие методы статистического анализа: корреляционный, регрессионный, дисперсионный, дискриминантный, факторный и компонентный, анализ таблиц сопряженности рядов и др.

Для анализа и прогнозирования динамических данных применяются следующие методы:

* адаптивные методы прогнозирования;
* модели динамической регрессии;
* модели прогнозирования на основе линейной регрессии;
* модели гармонического, спектрального анализа и частотной фильтрации.

Каждая из перечисленных выше моделей может управляться пользователем с помощью параметров, характеризующих эту модель. Такой подход позволяет постепенно осваивать заложенные в программе возможности и облегчает работу с ней.

С помощью **корреляционного** анализа рассчитывается матрица парных корреляций, матрица частных корреляций, а также коэффициенты множественных корреляций.

На основе **регрессионного** анализа решаются следующие задачи: установление форм зависимости (положительная, отрицательная, линейная, нелинейная).

**Компонентный и факторный** анализ – два принципиально различных статистических метода. В программе они объединены в единый блок, поскольку такое объединение оправдано с вычислительной точки зрения.

*Компонентный анализ* служит для определения структурной зависимости между случайными переменными. В результате его использования получается сжатое описание явления, несущее почти всю информацию, содержащуюся в исходных данных.

*Факторный анализ* является более общим методом преобразования исходных переменных по сравнению с компонентным анализом. В задачи факторного анализа входит: определение числа общих факторов, определение оценок общих и специфических факторов.

*Анализ временных рядов* включает в себя расчет статистических характеристик, анализ кривых роста по 16 функциям и некоторые адаптивные параметрические модели для анализа одномерных временных рядов.

Анализ автокорреляции динамического ряда выполняется с помощью графика автокорреляции.

Расчет кривых роста рассматривается как построение парной регрессии, в которой основной переменной является время.

Углубленный анализ предполагает использование адаптивных методов, сезонных методов прогнозирования. Для решения задач частотного анализа могут быть использованы методы частотной фильтрации, гармонического анализа, спектрального анализа.

Электронная таблица MNCALC представляет собой табличный процессор, сходный по своим функциональным возможностям с пакетами LOTUS 1-2-3 или EXCEL.

С точки зрения пользователя пакета «ОЛИМП» база данных MNCALC является таблицей, каждый столбец которой содержит переменную, а строки – значения переменных.

***Преимущество*** MNCALC по сравнению со стандартным редактором пакета ОЛИМП заключается в том, что он позволяет отображать и редактировать сразу весь набор данных. Кроме того, в таблице могут находиться формулы, с помощью которых можно формировать новые переменные на основе существующих. Таблицы могут содержать различную текстовую информацию, позволяющую именовать переменные и комментировать наборы данных.

***Программное средство ПСИ*** предназначено для формирования структуры анкет и ввода данных по этим анкетам. ПСИ содержит пять стандартных типов вопросов, наиболее часто встречающихся в анкетах.

Программное средство ПСИ разработано для подготовки данных с целью их последующей обработки при помощи пакета «ОЛИМП», однако может быть с успехом использовано как средство подготовки данных для других программных продуктов. В ПСИ реализован оригинальный пользовательский интерфейс, позволяющий быстро и легко освоить работу с ним.

**ППП «МЕЗОЗАВР»**

Основное назначение пакета «МЕЗОЗАВР» заключается в проведении разведочного анализа временных рядов. Это касается ситуации, когда необходимо «пощупать» имеющуюся числовую информацию, по усмотрению исследователя применяя различные методы обработки и анализируя получающиеся при этом результаты и их адекватность. Пакет позволяет осуществлять подобные исследования весьма оперативно и эффективно.

*Пакет «МЕЗОЗАВР» используется* для анализа временных рядов умеренной (не более нескольких тысяч наблюдений) длины. Диалог происходит по желанию пользователя на русском или английском языке. Управление осуществляется с помощью меню и клавиш быстрого доступа.

*Под временным рядом* понимается последовательность наблюдений за некоторой числовой характеристикой показателей, сделанных с постоянным шагом во времени (например ежегодно, ежемесячно, каждый час и т.п.). В статистике примерами подобных показателей могут служить на макроэкономическом уровне ежегодные, ежеквартальные, ежемесячные и т.п. объемы производства, поставок, перевозок, потребления; индексы цен и другие макроэкономические показатели; на уровне предприятия – объемы выпуска продукции, затраты, расход ресурсов, эволюция характеристик качества и др.

*Пакет «МЕЗОЗАВР» обладает следующими преимуществами* по вводу и хранению информации:

* имеет свой стандарт файлов, ввод информации в которые осуществляется через встроенный редактор данных типа «электронной таблицы»;
* предоставляет возможность сохранения в стандартных файлах любых данных, полученных в ходе анализа;
* допускает экспорт и импорт информации из текстовых файлов.

Предельная длина одного анализируемого временного ряда равна 16 тыс. значений, однако возможности анализа такого ряда будут весьма ограничены и поэтому наиболее эффективно работать с рядами до 2 – 3 тыс. значений.

Одновременно можно анализировать до 256 рядов, однако их суммарная не должна превышать 60 тыс. чисел. При этом можно пользоваться либо реальными временными шкалами (шаг по времени – начиная от 1 мин. до любого (целого) числа лет), или же условной временной шкалой.

Редактирование данных осуществляется с помощью встроенного табличного редактора или же в графическом режиме, а их преобразование – с помощью дополнительного меню преобразований более специального характера.

Кроме того, есть возможность непосредственно указать предварительное преобразование данных, которое далее учитывается в процессе анализа. При всех арифметических операциях над рядами учитывается совместимость временных шкал.

***Пакет «МЕЗОЗАВР» реализует следующие основные процедуры по анализу временных рядов*:** сглаживание, прогнозирование, фильтрацию, а также построение различных регрессионных зависимостей. Все процедуры снабжены мощной графической поддержкой с большим числом интерактивных возможностей, таких, как установка различных шкал, увеличение любого куска графика и т.д.

### Заключение

В настоящее время рынок математического программного обеспечения очень развит. На нём можно найти свыше тысячи программ, различного уровня, направленности, качества, требованиям к аппаратному обеспечению и, наконец, отличающиеся ценой.

Число статистических пакетов, получивших распространение в России, достаточно велико и спрос на них продолжает возрастать.

В данном реферате я рассказывала про отечественные статистические пакеты и про статистические пакеты в целом. Так же он содержит описание и характеристики некоторых пакетов.

В моей работе приведена таблица общих сведений об СПП, которая наглядно показывает категорию пользователей, цену, продавцов и т.д.

Для наиболее доступного понимания, в реферате описывается легкость использования пакета и его освоение.

В моей работе перечислены виды статистических пакетов.

Так же здесь можно найти описание некоторых отечественных пакетов, таких как «Эвриста», «Stadia», «ОЛИМП» и «Мезозавр».

**Особенности** отечественных пакетов: более простой путь; аккумулируют опыт передовой российской науки; методы анализа сгруппированы по содержательному принципу, а не по фамилии авторов.

**Недостаток** отечественных пакетов – это их малочисленность.

Документация отечественных пакетов не уступает зарубежным пакетам(Fore Cast Expert, Олимп Старт Эксперт).

При впечатляющем разнообразии статистических программных продуктов, которым характеризуется современный мировой и отечественный рынок (по официальным данным Международного статистического института, число наименований СПП приближается к тысяче), крайне важно - как для производителя, так и для потребителя этой продукции - правильно ориентироваться в этом многообразии.

**СПИСОК ЛИТЕРАТУРЫ**

***1.*** *Прикладная статистика: Классификация и снижение размерности: Справ. изд. М., 1989. 607 с.*

***2.*** *Айвазян С.А. Интеллектуализированные инструментальные системы в статистике и их роль в построении проблемно-ориентированных систем поддержки принятия решений// Обозрение проблем прикладной математики. Том 4, #2. М.: Наука; Изд-во "ТВП", 1997.*

***3.*** *Статистические и математические системы//Тысячи программных продуктов: Каталог: Вып. 2. М., 1995. С. 88-92.*

***4.*** *Козырев А.Н. Рынок программного обеспечения в СССР, лицензионные и авторские договоры, цены//Мир ПК. 1989. #3.*

***http://www.exponenta.ru***