

Министерство образования Российской Федерации
Ярославский государственный университет им. П.Г. Демидова

С.В. ТИХОНОВ

**ПРАКТИЧЕСКИЕ ЗАНЯТИЯ
ПО МАТЕМАТИЧЕСКИМ МЕТОДАМ
В БИОЛОГИИ И ЭКОЛОГИИ**

Учебное пособие

Ярославль 2003

ББК

УДК

Тихонов С.В. Практические занятия по математическим методам в биологии и экологии: Учеб. пособие/ Яросл. гос. ун-т. Ярославль, 2003. с.

В пособии излагаются основы анализа и синтеза биологических и экологических данных с использованием вычислительной техники. Последовательно рассматриваются основные этапы работы с разнотипной информацией и математического моделирования биосистем различного уровня организации. Практикум является дополнением к теоретической части соответствующих дисциплин и может использоваться для самостоятельной работы студентов при подготовке к практическим занятиям, контрольным мероприятиям и при обработке данных в ходе выполнения курсовых и дипломных работ.

Пособие предназначено для студентов университета, обучающихся по специальностям "Биология" и "Экология" и направлению "Экология и природопользование".

Печатается по решению редакционно-издательского совета Ярославского государственного университета.

Рецензенты:

ISBN

Ярославский
государственный
университет, 2003
С.В.Тихонов, 2003

Введение

Данное пособие предназначено для студентов, готовящихся стать специалистами в области изучения, охраны и рационального использования разнообразных объектов живой природы.

На практических занятиях по дисциплинам "Математические методы в биологии" и "Математические методы в экологии" студенты знакомятся со специальным программным обеспечением и осваивают основные методы анализа данных и математического моделирования.

Для этого должны использоваться знания, умения и навыки, приобретенные в ходе изучения высшей математики (включая теорию вероятностей) и информатики. Кроме того, необходимо знание основных теоретических положений математической статистики, излагающихся в соответствующих учебниках и лекционной части курса. Предполагается также, что студенты уже знакомы со структурой и функционированием биологических и экологических объектов изучения и имеют представление о методах получения и регистрации биологической и экологической информации.

Пособие написано на основе опыта проведения лабораторных практикумов, проводимых на факультете биологии и экологии ЯрГУ по дисциплинам "Математические методы в биологии" и "Математические методы в экологии", "Системная экология" и "Основы рыбного хозяйства" (раздел "Моделирование популяций и сообществ рыб").

Руководство состоит из четырех основных разделов: в первом дается краткая характеристика программного обеспечения, во втором описывается предварительная обработка данных, в третьем последовательно излагаются основные этапы анализа данных, четвертый посвящен применению метода математического моделирования в биологии и экологии. В начале каждого подраздела приводится краткое описание этапа обработки или используемого метода анализа, затем рассматриваются примеры его применения, а в конце даются упражнения, выполняемые на практических занятиях.

При работе с пособием рекомендуется использование электронных учебников, и решение дополнительных задач в ходе самостоятельной работы.

Содержание практикума естественно не может охватить все многообразие соответствующих программных продуктов, методов и способов анализа и синтеза биологических и экологических данных. Так, вопросы планирования эксперимента и многомерные методы анализа данных (кластерный, дискриминантный и факторный) в посо-

бии не рассматриваются. Но оно может стать той основой, на которой в дальнейшем будет происходить освоение новых методов решения практических задач в области биологии и экологии с учетом научной специализации студентов.

Автор надеется, что пособие поможет разрушить предубеждение студентов по поводу необходимости использования математических методов в биологии и экологии, а также будет полезно при обработке данных в ходе выполнения курсовых и дипломных работ.

Большую помощь в методическом и содержательном аспектах при написании данного пособия оказало совместное проведение практических занятий по применению компьютерных технологий в ихтиологии, гидробиологии и рыбном хозяйстве, проводимых на биологическом факультете ЯрГУ в 1983-1990 гг. с Л.А.Жаковым и В.В.Меншуткиным, за что автор выражает им искреннюю благодарность.

1. Программное обеспечение анализа данных.

Все многообразие программного обеспечения анализа данных можно условно разделить на пакеты общего назначения (MS Excel), специальные интегрированные системы и алгоритмические языки программирования. В свою очередь специальное ПО делится на математическое (Mathematica, Matlab, Maple, Mathcad), статистическое (Statistica, StatGraphics, SPSS, Stadia) и пакеты научной графики.

Наиболее доступными в вузах и широко применяемыми в научно-исследовательских организациях биологического профиля и в контролирующих природоохранных учреждениях являются электронные таблицы MS Excel и пакет Statistica.

1.1. Электронные таблицы MS Excel.

Электронные таблицы MS Excel располагают разнообразными и достаточно эффективными средствами анализа информации, а также ведения баз данных. Сюда относятся формулы, разнотипные встроенные функции, разнообразные типы графиков. Предполагается, что студенты имеют практические навыки использования этого инструментария, полученные при изучении информатики.

Использование функций для вычисления значений

Функции – это заранее определенные формулы, которые выполняют вычисления по заданным величинам, называемым аргументами. Список аргументов может состоять из чисел, текста, логических величин (например, ИСТИНА или ЛОЖЬ), массивов и др. Необходимо следить за соответствием типов аргументов. Кроме того, аргументы могут быть как константами, так и формулами. Эти формулы, в свою очередь, могут содержать другие функции.

Структура функции начинается с указания имени функции, затем вводится открывающая скобка, указываются аргументы, отделяющиеся точками с запятыми, а затем – закрывающая скобка. Перед именем функции вводится знак равенства (=).

Таблица 1

Некоторые математические и статистические функции электронных таблиц.

Категория	Название	Выполняемое действие	Список аргументов
Математические	ABS	Возвращает модуль (абсолютную величину) числа.	(число)
Математические	EXP	Возвращает число «е» возведенное в указанную степень.	(число)
Математические	LN	Возвращает натуральный логарифм числа.	(число)
Математические	LOG	Возвращает логарифм числа по заданному основанию.	(число; основание)
Математические	LOG10	Возвращает десятичный логарифм числа.	(число)
Математические	КОРЕНЬ	Возвращает положительное значение квадратного корня.	(число)
Математические	ОКРУГЛ	Округляет число до указанного количества десятичных разрядов.	(число; количество десятичных разрядов)
Математические	СТЕПЕНЬ	Возвращает результат возведения в степень.	(число; степень)
Математические	СУММ	Суммирует числа.	(диапазон чисел)
Статистические	ДИСП	Оценивает дисперсию по выборке.	(диапазон чисел)
Статистические	ДОВЕРИТ	Возвращает доверительный интервал для среднего генеральной совокупности.	(критический уровень значимости; стандартное отклонение; количество чисел в диапазоне)
Статистические	КОРРЕЛ	Возвращает коэффициент корреляции.	(диапазон чисел)
Статистические	МАКС	Возвращает наибольшее значение	(диапазон чисел)

Статистические	МЕДИАНА	Возвращает медиану	(диапазон чисел)
Статистические	МИН	Возвращает наименьшее значение.	(диапазон чисел)
Статистические	МОДА	Возвращает наиболее часто встречающееся или повторяющееся значение.	(диапазон чисел)
Статистические	СКОС	Возвращает асимметрию распределения.	(диапазон чисел)
Статистические	СРГАРМ	Возвращает среднее гармоническое.	(диапазон чисел)
Статистические	СРГЕОМ	Возвращает среднее геометрическое.	(диапазон чисел)
Статистические	СРЗНАЧ	Возвращает среднее арифметическое.	(диапазон чисел)
Статистические	СТАНДОТКЛОН	Оценивает стандартное отклонение по выборке.	(диапазон чисел)
Статистические	СЧЕТ	Подсчитывает количество чисел (объем выборки).	(диапазон чисел)
Статистические	ЭКСЦЕСС	Возвращает эксцесс распределения.	(диапазон чисел)

В электронных таблицах также имеется пакет "Анализ данных", вызываемый из меню "Сервис" (рис. 1). При отсутствии названия пакета в списке следует использовать пункт "Настройки".

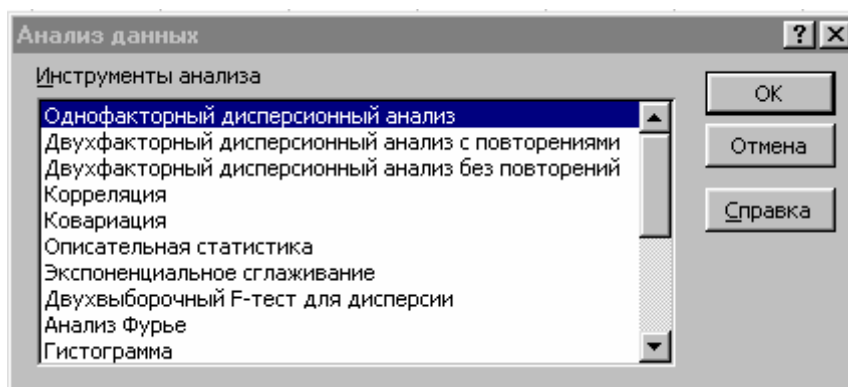


Рис. 1. Общий вид меню пакета "Анализ данных".

Применение отдельных модулей пакета для решения практических задач разбирается в разделе 3.

Таблица 2.

Основные инструменты пакета "Анализ данных".

Дисперсионный анализ	Пакет анализа включает в себя три варианта дисперсионного анализа: однофакторный, двухфакторный с повторениями и без повторений.
Корреляция	В результате анализа формируется корреляционная матрица.
Описательная статистика	Это средство анализа служит для создания статистического отчета, содержащего информацию о среднем значении, показателях вариации, стандартной ошибке и др. параметрах одной или нескольких выборок.
Двухвыборочный F-тест для дисперсий	Двухвыборочный F-тест применяется для сравнения дисперсий двух генеральных совокупностей.
Регрессия	Линейный регрессионный анализ заключается в подборе уравнения и соответствующего графика для набора наблюдений с помощью метода наименьших квадратов. Регрессия используется для анализа воздействия на отдельную зависимую переменную значений одной или более независимых переменных.
Парный двухвыборочный t-тест для средних	Парный двухвыборочный t-тест Стьюдента используется для проверки гипотезы о различии средних для двух выборок данных. В нем не предполагается равенство дисперсий генеральных совокупностей, из которых выбраны данные. Парный тест используется, когда имеется естественная парность наблюдений в выборках, например, когда генеральная совокупность тестируется дважды.
Двухвыборочный t-тест	Двухвыборочный t-тест Стьюдента служит для проверки гипотезы о равенстве средних для двух выборок. Одна форма t-теста предполагает совпадение дисперсий генеральных совокупностей, другая – несовпадение.

1.2. Пакет прикладных программ Statistica.

Общая структура системы

STATISTICA представляет собой интегрированную систему статистического анализа и обработки данных. Система состоит из следующих основных компонентов:

- многофункциональной системы для работы с данными, которая включает в себя электронные таблицы для ввода и задания исходных данных, а также специальных таблиц (*Scroolsheet*) для вывода численных результатов анализа. Для сложной (специализированной) обработки данных в *STATISTICA* имеется модуль *Управления данными*;
- графической системы для визуализации данных и результатов статистического анализа;
- набора статистических модулей, в которых собраны группы логически связанных между собой статистических процедур. В любом конкретном модуле можно выполнить определенный способ статистической обработки, не обращаясь к процедурам из других модулей. Каждый модуль является полноценным *Windows* приложением. Поэтому пользователь имеет возможность одновременной работы как с одним, так и с несколькими модулями. Переключаться между ними можно как между обычными *Windows* приложениями, например, при помощи комбинации клавиш ALT+TAB или, выбирая их на *панели задач*. Все основные операции при работе с данными и графические возможности доступны в любом статистическом модуле и на любом шаге анализа;
- специального инструментария для подготовки отчетов. При помощи текстового редактора, встроенного в систему, можно готовить полноценные отчеты. В *STATISTICA* также имеется возможность автоматического создания отчетов.

Все структурные компоненты *STATISTICA* настолько тесно интегрированы между собой, что разделение на различные компоненты во многом условно и является полезным лишь для изучения системы с методической точки зрения.

STATISTICA работает с четырьмя различными типами документов, которые соответствуют основным структурным компонентам системы. Это:

- электронная таблица *Spreadsheet*, которая предназначена для ввода исходных данных и их преобразования;
- электронная таблица *Scrollsheet* для вывода численных и текстовых результатов анализа;
- график - документ в специальном графическом формате для визуализации и графического представления численной информации;
- отчет - документ в формате *RTF (Расширенный текстовый формат)* для вывода текстовой и графической информации.

В соответствии со стандартами среды *Windows* каждый тип документа выводится в своем собственном окне в рабочей области системы *STATISTICA*. Как только это окно становится активным, изменяется панель инструментов и меню. В них появляются команды и кнопки, доступные для активного документа.

Ввод данных

Данные в *STATISTICA* организованы в виде электронной таблицы - *Spreadsheet*. Они могут содержать как численную, так и текстовую информацию. Данные в электронной таблице могут иметь различные форматы, например, даты, времени, научный формат и др. Электронные таблицы в *STATISTICA* поддерживают различные типы операций с данными, такие как: операции с использованием *буфера обмена Windows*; операции с выделенными блоками значений (аналогично *MS Excel*), в том числе и с использованием метода *Drag-and-Drop* - "*Перетащить и опустить*", автозаполнение блоков и т.д.

Ввести данные в электронную таблицу можно одним из следующих способов.

- Непосредственно ввести их в электронную таблицу с клавиатуры. В *STATISTICA* имеются развитые инструментальные средства для автоматизации ручного ввода данных.
- Вычислить новые данные на основе уже введенных данных при помощи формул, которые можно задать в электронной таблице. При этом имеется возможность быстрого доступа к большому количеству специализиро-

ванных математических и статистических функций, допускается использование логических операторов.

- Воспользоваться данными, подготовленными в другом приложении используя операции копирования данных через *Буфер обмена - Clipboard Windows*.

Vars Cases										
TEXT	VALU	1	2	3	4	5	6	7	8	9
		ПОЛ	ВОЗРАСТ	РОСТ	ВЕС	ОГК1	ОГК2	ОГК3	ДИН1	ДИН2
1		f	6	128,5	24,0	67,0	63,0	59,0	10,0	5,0
2		f	6	135,0	25,5	61,0	57,5	54,0	15,5	10,0
3		f	6	126,0	20,3	63,0	59,0	55,0	9,0	9,0
4		f	6	124,0	20,2	60,0	56,0	53,0	15,0	6,0
5		f	6	128,5	20,4	60,0	56,0	53,0	10,0	10,0
6		f	6	124,6	20,5	68,0	62,0	57,0	6,0	10,0
7		f	6	124,0	19,8	58,0	57,0	56,0	3,0	4,0
8		f	6	124,0	20,0	58,0	57,0	56,0	6,0	4,0
9		f	6	130,0	25,6	62,0	60,0	60,0	3,0	6,0
10		f	6	125,0	27,2	65,0	63,0	62,0	15,0	13,0
11		f	6	124,0	21,1	60,0	58,0	56,0	5,0	9,0
12		f	6	124,0	21,1	59,0	57,5	57,0	10,0	12,0
13		f	6	122,0	23,8	62,0	60,0	59,0	10,0	12,0
14		f	6	112,0	17,0	56,0	54,0	53,0	9,0	5,0
15		f	6	129,0	24,0	59,0	57,0	56,0	14,0	11,0
16		f	6	122,5	24,6	63,0	62,0	61,0	5,0	5,0

Рис. 2. Электронная таблица системы Statistica. Кнопка **Vars** обеспечивает управление столбцами таблицы (переменными - **Variables**), **Cases** – строками (наблюдениями).

Вывод численных и текстовых результатов анализа

Численные результаты статистического анализа в системе *STATISTICA* выводятся в виде специальных электронных таблиц, которые называются таблицами вывода результатов - *Scrollsheets*. Таблицы *Scrollsheet* могут содержать любую информацию (как численную, так и текстовую), от короткой строчки до мегабайтов результатов. Обычно даже в результате простейшего статистического анализа мы получаем на выходе большое количество численной и графической информации. В системе *STATISTICA* эта информация выводится в виде последовательности (очереди) которая состоит из набора таблиц *Scrollsheet* и графиков.

STATISTICA содержит большое количество инструментов для удобного просмотра результатов статистического анализа и их визуализации. Они включают в себя стандартные операции по редактированию таблицы (включая операции над блоками значений, *Drag-and-Drop* - "Перетащить и опустить", автозаполнение блоков и др.), операции удобного просмотра (подвижные границы столбцов, разделение прокрутки в таблице и др.) доступ к основным статистикам и графическим возможностям системы *STATISTICA*. При выводе целого ряда результатов (например, корреляционной матрицы) *STATISTICA* отмечает значимые коэффициенты корреляции цветом. Пользователь так же имеет возможность выделить при помощи цвета необходимые значения в таблице *Scrollsheet*.

Если пользователю необходимо провести детальный статистический анализ промежуточных результатов, то можно сохранить таблицу *Scrollsheet* в формате файла данных *STATISTICA*, и далее работать с ним, как с обычными данными.

Кроме вывода результатов анализа в виде отдельных окон с графиками и таблицами *Scrollsheet* на рабочем пространстве системы *STATISTICA*, в системе имеется возможность создания отчета, в окно которого может быть выведена вся эта информация. Отчет - это документ (в формате *RTF*) который может содержать любую текстовую или графическую информацию. В *STATISTICA* имеется возможность автоматического создания отчета, так называемого автоотчета. При этом любая таблица *Scrollsheet* или график могут автоматически быть направлены в отчет.

Графические возможности системы *STATISTICA*

Система *STATISTICA* обладает широкими графическими возможностями. *STATISTICA* включает в себя большое количество разнообразных категорий и типов графиков (включая научные, деловые, трехмерные и двухмерные графики в различных системах координат, специализированные статистические графики - гистограммы, матричные, категоризованные графики и др.).

В систему *STATISTICA* включено большое количество инструментов настройки всех компонент графиков. Имеется возможность выбора различных типов линий, форматов разметки осей, цветов, легенд, названий и других атрибутов графика. Настроенные атрибуты могут быть сохранены в специальном файле и потом применяться к другим графикам. Доступ ко всем основным командам настройки реализован при помощи

контекстных меню, которые появляются при нажатии на правую кнопку мыши, общего меню и из панели инструментов графика.

Графические средства системы *STATISTICA* доступны в любом статистическом модуле и на любом шаге статистического анализа. Они могут быть использованы в целях:

- визуализации численных и текстовых значений непосредственно из электронной таблицы с исходными данными *STATISTICA* или таблицы *Scrollsheet* с результатами анализа;
- вывода результатов статистического анализа в виде последовательности (очереди) графиков. Для этого в диалоговых окнах всех статистических процедур имеется возможность построения различных, предназначенных именно для этого вида анализа типов графиков.

Визуализация данных является важным этапом статистического анализа. Очень важно подобрать вид графика, который наибольшим образом соответствует исследуемым данным. В систему *STATISTICA* включены удобные инструментальные средства для проведения интерактивного графического анализа данных. При помощи так называемой *Кисти*, включенной в *STATISTICA*, вы можете выделить какие-либо точки на построенном графике (например, выбросы значений переменной) и провести дальнейший статистический анализ без значений, которые соответствуют этим точкам. Эти точки могут быть помечены специальным образом, им может быть присвоена метка. Численные значения, соответствующие выделенным точкам на графике, могут быть отмечены в специальной электронной таблице, которая связана с графиком и просмотрены в специальном *Редакторе данных графика*..

Графический документ в *STATISTICA* может быть сохранен как:

- графический документ в специальном графическом формате системы *STATISTICA* (расширение файла **.stg*), который может быть открыт позже и использован в процессе анализа,
- в графическом растровом формате (расширение файла **.bmp*, **.pcx*).

Основные шаги обработки данных в системе

Статистическая обработка данных в системе *STATISTICA* обычно состоит из следующих основных шагов:

- ввод исходных данных в электронную таблицу системы *STATISTICA*;
- предварительные преобразования данных перед непосредственным применением конкретного статистического метода;
- визуализация данных при помощи того или иного типа графиков;
- статистический анализ при помощи некоторого статистического метода;
- вывод численных, текстовых и графических результатов, как на рабочее пространство системы, так и в файл с отчетом;
- анализ результатов.

Статистические процедуры

Статистические процедуры системы *STATISTICA* сгруппированы в нескольких специализированных статистических модулях (рис. 3). В каждом модуле вы можете выполнить определенный способ обработки, не обращаясь к процедурам из других модулей.

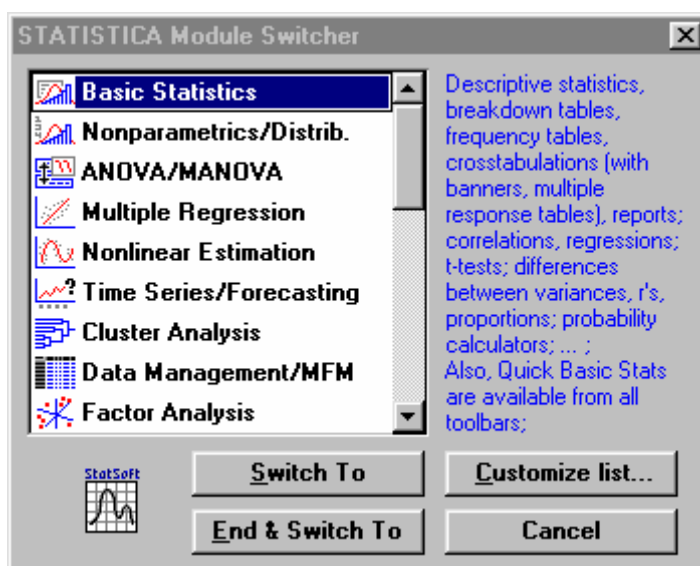


Рис. 3. Переключатель модулей системы Statistica.

Модуль Основные статистики и таблицы

Если вы хотите провести предварительную обработку данных, осуществить разведочный анализ, определить зависимости между ними, разбить их различными спосо-

бами на группы, просмотреть эти группы визуально и определить взаимосвязи между данными, запустите модуль Basic Statistics - Основные статистики (рис. 4). Обычно с этого модуля начинается работа в системе.

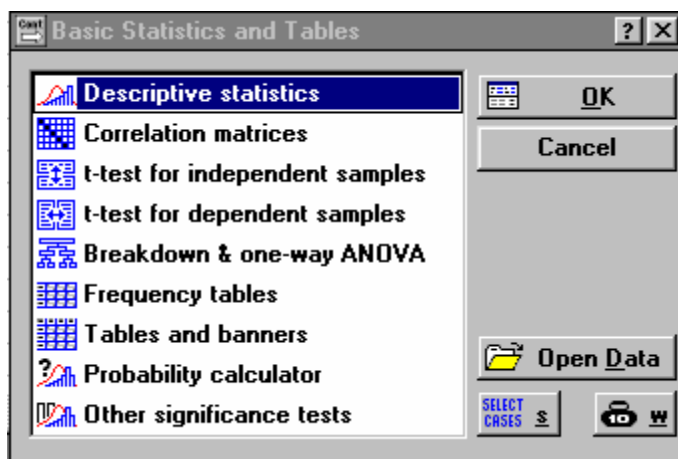


Рис. 4. Стартовая панель модуля **Basic Statistic and Tables** (Основные статистики и таблицы).

Этот статистический модуль включает в себя приведенные ниже группы статистических процедур.

Описательные статистики, группировки, разведочный анализ. *STATISTICA* предлагает широкий выбор методов разведочного статистического анализа. Система может вычислить практически все описательные статистики, включая медиану, моду, квартили, определенные пользователем процентиля, средние и стандартные отклонения, доверительные интервалы для среднего, коэффициенты асимметрии, эксцесса (с их стандартными ошибками), а также многие другие описательные статистики. Практически все описательные статистики могут быть вычислены для данных, разделенных на группы с помощью одной или нескольких группирующих переменных. Например, всего несколько щелчков мыши позволяют разбить данные на группы, например, ПОЛ и ВОЗРАСТ и посмотреть категоризованные гистограммы на графике, а также графики значений, матричные и многие другие типы графиков.

Корреляции. Этот раздел включает средства, позволяющих исследовать зависимости между переменными с использованием коэффициента корреляции Пирсона. Специальные графические возможности позволяют выбрать отдельные точки на диаграмме рассеяния и оценить их вклад в регрессионную кривую или любую другую кри-

вую, подогнанную к данным. Вычисленные корреляционные матрицы выводятся в виде электронных таблиц Scrollsheet.

t - критерии и другие критерии для групповых различий.

Модуль Множественная регрессия

Если вы хотите построить зависимости между многомерными переменными, подобрать простую линейную модель и оценить ее адекватность, воспользуйтесь модулем - Multiple Regression - Множественная регрессия.

Модуль Дисперсионного анализа.

Если вам известны факторы, которые воздействуют на измеряемые вами данные, и вы хотите оценить степень этого воздействия, понять существенно оно или нет, выберите модуль ANOVA. Модуль представляет собой набор процедур общего одномерного и многомерного дисперсионного анализа.

Модуль Непараметрической статистики и подгонки распределений

Если вы хотите проверить различные гипотезы о характере распределения ваших данных, работайте в модуле Nonparametrics/Distribution - Непараметрическая статистика/Распределения. Модуль содержит обширный набор непараметрических критериев согласия, в частности, критерий Колмогорова-Смирнова, ранговые критерии Манна-Уитни, Вилкоксона и многие другие.

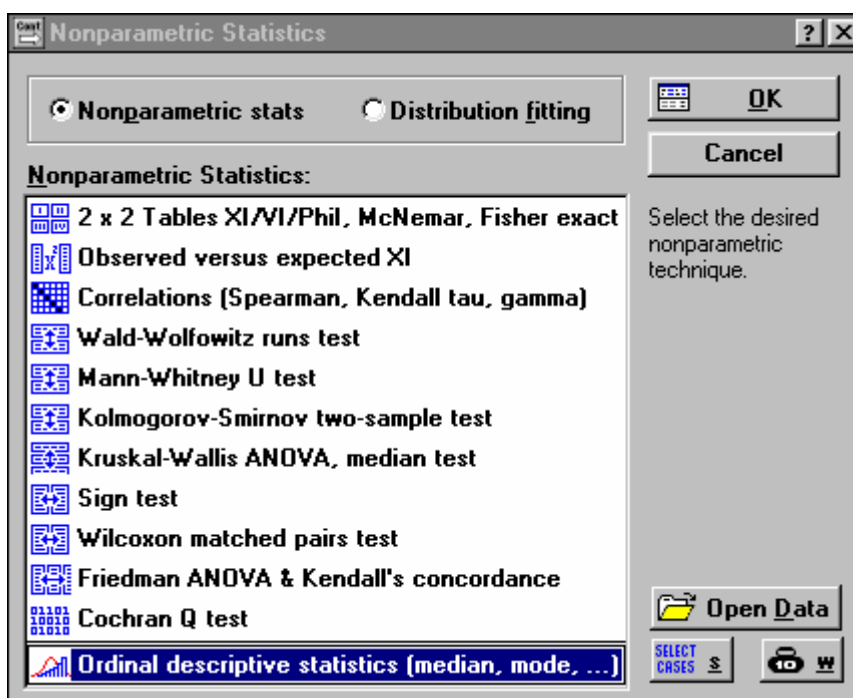


Рис. 5. Стартовая панель модуля **Nonparametric Statistics** (Непараметрическая статистика).

Статистические процедуры модуля позволяют пользователю легко сравнить распределение наблюдаемых величин с большим количеством различных теоретических распределений. Вы можете подогнать к данным нормальное, равномерное, линейное, экспоненциальное, биномиальное, Пуассоновское, геометрическое распределения, распределение Бернулли. Точность подгонки оценивается с помощью критерия хи-квадрат или одновыборочного критерия Колмогорова-Смирнова (параметры подгонки могут контролироваться). Дополнительно подгонка гипотетического распределения к эмпирическому может быть оценена в пользовательских гистограммах (стандартных или кумулятивных) с наложением на них выбранных функций; линейные и столбчатые графики ожидаемых и наблюдаемых частот могут быть получены непосредственно из таблицы с выходными данными. Определенные пользователем функции от одной и двух переменных также легко могут быть построены и наложены на графики.

1.3. Язык программирования Бейсик.

Бейсик (Basic) является простым и удобным языком программирования, который преподается во многих школах в курсе "Информатика", на биологических факультетах российских и зарубежных вузов. На этом языке написано большое количество программ в области биологии и экологии. Настоящий раздел посвящен описанию Бейсика только в том объеме, который необходим для работы с математическими моделями биологических и экологических систем.

Бейсик имеет много различных версий и модификаций. Ниже приводится описание обобщенной версии для интерпретаторов (aBasic, GW Basic) и компиляторов ("Turbo Basic", "Power Basic").

Комментарии

Строка, начинающаяся со служебного слова REM - это комментарий, не оказывающий никакого влияния на работу программы. По правилам Бейсика после слова REM в строке можно писать все, что угодно.

Числа

Числовые константы или просто числа имеют в Бейсике вид весьма близкий к обычному. Знаки десятичной дроби отделяются от целой части не запятой, а точкой. Очень большие или очень маленькие числа записываются в Бейсике в экспоненциальной форме, например:

0.0000025.....2.5E-06
455890000.....4.5589E08

Двузначное число после символа "E" означает, в какую степень надо возвести число 10, чтобы помножив на то, что находится перед символом "E", получить нужную величину.

Целые числа заканчиваются символом "%", что к процентам не имеет никакого отношения. Диапазон изменения целых чисел от -32768% до 32768%. Применять целые

числа следует только тогда, когда величина действительно не может иметь дробной части, например число позвонков у рыб или календарное число данного месяца.

Переменные

Переменной называется величина, которая может принимать различные числовые значения. Например численность рыб в данном водоеме или концентрация растворенного кислорода, выраженная в мг/л. Идентификатор, т.е. название переменной, может представлять собой несколько букв и цифр:

N численность особей в популяции,
WEIGHT средняя масса особи.

Если переменная может принимать значения только в виде целых чисел, то после символов идентификатора можно поставить "%", например:

Y%.....год,
M%.....месяц.

Следует помнить, что Y и Y% - это разные идентификаторы.

Массивы

В биологии и экологии очень часто названию одной величины соответствует не одно число, а несколько чисел, расположенных в определенном порядке. Например, численность популяции, состоящей из трех возрастных групп - это 3 числа, расположенных одно за другим. В информатике такие величины называют массивами. Кроме идентификатора - названия данной величины - массив характеризуется еще и размером. Под размером понимается предельное количество чисел, которое может содержать массив.

В отличие от переменной массивы требуют предварительного описания. Это означает, что до любых действий с массивами вычислительной машине следует сообщить о размерах используемых массивов. Делается это при помощи специального оператора описания, который имеет вид:

DIM N(3)

Здесь DIM - служебное слово (dimension - размерность), Y - идентификатор массива, а 3 - число элементов массива.

Для обращения к элементу массива надо указать его индекс, т.е. порядковый номер. Например, N(3) - численность трехгодовиков.

Индекс может быть не только числом, но и переменной, например N(I). Переменная I в этом случае должна быть целым числом. Если переменная I имеет дробную часть, то она просто отбрасывается.

В том случае, когда значение индекса выходит за пределы описания (например N5(100) или N(-2) в рассматриваемом случае), то выдается сообщение об ошибке.

Возможно описание двумерных массивов, в которых элемент определяется двумя индексами. Описание массива всегда должно предшествовать обращению к конкретному элементу.

Арифметические выражения

Арифметические выражения состояются из переменных и знаков арифметических действий: сложения (+), вычитания (-), умножения (*), деления (/) и возведения в степень (^). Например вычисление биомассы оформляется как:

$$N*W1,$$

где N - численность, а W1 - средняя масса особи.

Скобки имеют в Бейсике свой обычный смысл, т.е. выражения, стоящие внутри скобок выполняются раньше всех остальных. При отсутствии скобок сначала выполняются операции возведения в степень, затем умножение и деление, а сложение и вычитание в последнюю очередь.

Знак минус, стоящий перед одиночной переменной (например -W1), имеет смысл перемены знака числа на обратный, однако написание подряд двух знаков арифметических действий на Бейсике не допускается и приходится пользоваться скобками, например:

$$A*(-W1)$$

Оператор присваивания

Первоначальное значение всех переменных в Бейсике равно нулю. Оператор присваивания служит для придания переменной нового числового значения. Этот оператор состоит из переменной, знака равенства (=) и числа или арифметического выражения. Например:

$$A=752.2 \quad B=N*W1 \quad K=Q1*((T-20)/10)$$

Обратим внимание на последний оператор. Это ни что иное, как вычисление температурной поправки к интенсивности дыхания гидробионтов, определенной при 20 градусах. Т - действительная температура среды, Q1 - коэффициент Q₁₀.

Работа в непосредственном режиме

Начните с самого простого, например, наберите

$$A=2+3 : \text{PRINT } A$$

после чего нажмите клавишу ENTER. На экране должно появиться число 5. Двоеточие означает разделение операторов, а оператор PRINT обеспечивает вывод на дисплей текущего значения той переменной, которая стоит за словом PRINT.

Вычислим интенсивность дыхания рыбы, выраженного в мл кислорода в час при стандартных условиях. Масса рыбы 1000 г.

$$W=1000 : Q=0.498*W^{0.79} : \text{PRINT } Q$$

После нажатия клавиши ENTER на экране должно появиться 116.7426

Работа в непосредственном режиме, по существу, превращает ЭВМ в микрокалькулятор, при этом может использоваться только одна строка дисплея. Собственно поэтому и приходится разделять операторы двоеточием.

Программный режим

В отличие от непосредственного режима в программном режиме строки могут нумероваться целыми числами в порядке возрастания (хотя можно обойтись и без номеров). Практически нумерация ведется числами кратными 10 для того, чтобы облегчить дальнейшее исправление программы.

Запуск программы осуществляется нажатием клавиши F9. При этом происходит чтение и компиляция текста программы построчно. Если строки пронумерованы, то выполняются операторы той строки, которая имеет минимальный номер. Следующей выполняется строка с ближайшим номером в порядке возрастания и т.д.

Порядок выполнения строк может изменяться специальными операторами (GO TO, FOR...NEXT, RETURN и другими), о которых будет сказано ниже.

В конце программы обычно помещают оператор END, которому присваивается наибольший номер строки.

Условный оператор

Условный оператор представляет собой конструкцию, состоящую из слова IF, условного выражения, слова THEN, за которым следует или оператор присваивания или номер строки, к которой следует перейти в случае выполнения условного выражения.

Под условным выражением понимается такое выражение, которое при одних значениях входящих в него переменных может выполняться (быть истинным), а при других нет (быть ложным). Например выражение

$$N1 > N2$$

выполняется при $N1=120$ и $N2=80$, но не выполняется при $N1=120$ и $N2=220$.

В условных выражениях применяются отношения тождества (=), символ которого лишь по внешнему виду совпадает с таким же символом в операторе присваивания, а так же отношения больше (>), меньше (<), равно или больше (>=), равно или меньше (<=) и не равно (<>). В правой и левой части условного выражения могут быть любые арифметические выражения. Например:

$$I+1=J$$

Такая форма совершенно недопустима для оператора присваивания, но вполне приемлема в качестве условного выражения.

Приведем пример условного оператора:

```
10 IF L > 60 THEN Y = N*F
```

Если под L понимается размер рыбы в мм, N - численность популяции, а F - интенсивность промысла, то приведенный оператор будет вычислять вылов (Y) только в том случае, когда размер рыбы больше 60 мм. Это в какой-то мере имитирует действие отцеживающих орудий лова.

Запишем программу, которая определяет минимальную величину из двух предложенных. В гидробиологии такая ситуация возникает, например, при лимитировании продукции фитопланктона (P1) азотом (N) и фосфором (P), растворенными в воде.

```
10 INPUT N : INPUT P
20 P1=N
30 IF N > P THEN P1=P
40PRINT P1
50 END
```

Оператор INPUT обеспечивает ввод числового значения соответствующей переменной с клавиатуры компьютера. Готовность к вводу символизируется появлением на экране дисплея знака "?". Следует набрать желаемое значение и отправить его в оперативную память машины нажатием клавиши ENTER.

Оператор цикла

Оператор цикла обеспечивает повторное выполнение некоторой части программы (называемой телом цикла) заданное число раз. Разберем действие оператора цикла на примере вычисления биомассы популяции (B), состоящей из 10 возрастных групп. Численности особей в возрастных группах располагаются в массиве N, а средние веса особей в массиве W1.

```
10 DIM N(10), W1(10)
20 FOR I=1 TO 10
```

```

30     INPUT N(I)
40     INPUT W1(I)
50     B=B+W1(I)*N(I)
60 NEXT I
70 PRINT B
80 END

```

Оператор, размещенный в строке 20, называется заголовком цикла и состоит из слова FOR, присвоения начального значения циклической переменной ($I=1$), слова TO и конечного значения циклической переменной (10). Конец цикла (строка 60) состоит из слова NEXT и идентификатора циклической переменной (I). Операторы, заключенные между заголовком цикла и его концом, составляют тело цикла (строки 30-50). Выполнение операторов тела цикла повторяется столько раз, сколько единиц заключено между начальным (1) и конечным (10) значением циклической переменной, включая и эти крайние значения. При каждом прохождении тела цикла переменная I увеличивает свое значение на единицу.

В случае необходимости изменения циклической переменной не с шагом равным единице, а с каким-либо другим, следует добавить в заголовок цикла слово STEP с последующим указанием величины шага. Например нахождение произведения всех нечетных чисел от 1 до 15 можно сделать так:

```
J=1 : FOR I=1 TO 15 STEP 2 : J=J*I : PRINT J : NEXT I
```

Шаг циклической переменной может быть отрицательным, но тогда начальное значение должно быть больше конечного.

Начальные и конечные значения циклической переменной сами могут быть переменными величинами. Это открывает возможность обработки последовательности чисел, длина которой заранее не известна.

Операторы цикла можно вкладывать один в другой, например для вычисления элементов двумерных массивов. Необходимо только следить, чтобы границы циклов не пересекались.

Оператор безусловного перехода

Передача управления на любую (естественно реально существующую) строку программы осуществляется оператором безусловного перехода, который состоит из слов GOTO и следующего за ними номера строки или названия метки, куда передается управление. Например выбор коэффициента усвоения пищи (U) для хищников (J=1) или травоядных (J=2) можно осуществить так:

```

10 IF J=1 THEN 40
20 U=0.65
30 GOTO 50
40 U=0.8
50 A=U*C

```

В этом фрагменте программы при J=1 получается значение коэффициента усвоения 0.8, а при J=2 (или больше) U=0.65, т.к. условное выражение в операторе 10 не выполняется и срабатывает оператор GOTO 50. Оператор на строке 40 при этом окажется обойденным.

Функции

В Бейсике имеется целый ряд встроенных математических функций, обращение к которым не требует каких-либо предварительных описаний. Рассмотрим только те функции, которые будут использованы в дальнейшем изложении.

ABS(X) - абсолютная величина переменной или выражения X, заключенного в скобки. Если $X < 0$, то знак минус перед X просто отбрасывается.

INT(X) - целая часть числа или результата вычисления арифметического выражения X, иными словами от X отбрасывается дробная часть.

SQR(X) - квадратный корень из X. EXP(X) - экспоненциальная функция. LOG(X) и LOG10(X) - натуральный и десятичный логарифмы.

Экспоненциальная функция исключительно часто используется в биологии и экологии. Например численность поколения рыб в зависимости от возраста (T) и коэффициента общей смертности (M) можно вычислить при помощи оператора:

$$N = N_0 * \text{EXP}(M * T)$$

RND - функция выработки псевдослучайного числа, равномерно распределенного в диапазоне от 0 до 1. Эта функция используется, главным образом, для имитации случайных событий и величин. Например, гибель особи с вероятностью 0.1 ($W=1$ - особь существует, $W=0$ - особь не существует) можно представить в виде оператора:

```
IF RND < 0.1 THEN W = 0
```

Получить таблицу целых случайных чисел, равномерно распределенных в диапазоне от 1 до 10 можно, например, так:

```
FOR I=1 TO 20 : PRINT 1+INT(RND*10) : NEXT I
```

Для получения случайной величины с распределением похожим на нормальное и с математическим ожиданием равным нулю можно сложить несколько равномерно распределенных и независимых случайных величин, например:

```
FOR I=1 TO 20 : PRINT (RND+RND+RND+RND)-2 : NEXT I
```

Следует помнить, что при каждом проходе программы, функция RND будет выдавать одну и ту же последовательность значений - это удобно при отладке программы, но совершенно недопустимо при использовании метода статистических испытаний (метода Монте-Карло). Чтобы функция RND давала каждый раз неповторяющиеся значения, надо в начале программы поставить оператор RANDOMIZE.

При необходимости введения в программе новых функций, их следует предварительно описать. Например температурную поправку с коэффициентом $Q_{10}=2.3$ можно оформить в виде функции FNK(X). Описание этой функции будет иметь вид:

```
DEF FNK(X)=2.3^((X-20)/10)
```

Это описание должно стоять в начале программы, а в самой программе определение трат на обмен (R) можно вычислить при помощи оператора:

```
R = A*W^0.75*FNK(T)
```

где A - коэффициент, W - масса тела животного и T - температура.

Ввод и вывод

Простейшее использование операторов INPUT и PRINT уже было рассмотрено выше. Для ввода большого количества числовых констант рационально воспользоваться операторами READ и DATA.

Действие оператора READ аналогично действию оператора INPUT с той лишь разницей, что вводимые числа должны быть заранее размещены в отдельной строке, начинающейся словом DATA. Например ввод последовательности масс особей рыб от сеголеток до возраста 3+ может выглядеть так:

```
10 DATA 2,8,12,16
20 READ W1,W2,W3,W4
```

При заполнении массива целесообразно использовать оператор цикла:

```
10 DIM W(4)
20 DATA 2,8,12,16
30 FOR I=1 TO 4 : READ W(I) : NEXT I
```

Оператор вывода информации на экран дисплея (PRINT) может содержать не одну переменную, а несколько. Если переменные разделены запятой, то значения переменных располагаются в зонах по 14 символов. Если переменные разделены точкой с запятой (;), то их значения выводятся друг за другом с промежутком в один пробел.

Строковые переменные

Кроме чисел Бейсик может оперировать с последовательностями латинских и русских букв, а также специальных символов ($\# \$ \% = + ? ! @ \backslash * / < >$). Такие последовательности называются строковыми переменными или просто строками.

Строковые переменные широко используются при обработке текстовой информации, создании баз данных и оформлении результатов расчетов. Например при работе со списками видов целесообразно на вводе и на выводе оперировать с привычными

биологам латинскими названиями животных и растений, чтобы избежать трудоемкой и чреватой многими ошибками ручной кодировки.

Строковые константы уже использовались при рассмотрении оператора PRINT. Напомним, что строковая константа обязательно заключается в кавычки, например:

```
PRINT "Perca fluviatilis L."
```

Идентификаторы строковых переменных должны заканчиваться символом \$, который в некоторых инструкциях именуется "знаком денежной единицы". Приведем примеры операторов присвоения строковых переменных:

```
A$="МЕСЯЦ" A1$="АПРЕЛЬ"
```

Строковые переменные, как и числовые, могут образовывать одномерные и двумерные массивы.

Для строк определена операция объединения или склеивания. Знак этой операции "+". Существо операции склеивания ясно из примера:

```
A2$ = A$ + " " + &A1$
```

в результате A2\$ принимает значение "МЕСЯЦ АПРЕЛЬ". При операциях со строками пробел имеет такие же права и свойства, как и все остальные символы.

Строки можно использовать в условных выражениях. Для них вполне естественны отношения эквивалентности (=) и неравенства (<>). Например:

```
IF A1$="АПРЕЛЬ" THEN 300
IF S$<>"ESOX" THEN 400"
```

Файлы

Файлом называется совокупность последовательно расположенных данных (числовых или строковых) размещенных на внешних носителях и имеющих собственное имя. Перед обращением к файлу его необходимо "открыть", т.е. связать имя файла с определенным каналом ввода-вывода. Для этого используется оператор OPEN. После

завершения работы с файлом его необходимо "закрыть", т.е. отключить от канала ввода-вывода. Это производится оператором CLOSE.

Если файл выводится (например для печати или долговременного хранения), то оператор открытия файла может иметь вид:

OPEN "STATE" FOR OUTPUT AS FILE #1

Здесь "STATE" - название файла, "#1" - это номер канала ввода-вывода.

Для закрытия файла используется оператор CLOSE #1. Оператор CLOSE без всяких указаний на номер канала закрывает все открытые к этому времени файлы.

Подпрограммы

Если программа достаточно длинна и допускает выделение логически обособленных частей или же некоторые группы операторов встречаются в программе несколько раз, то целесообразно пользоваться подпрограммами. Обращение к подпрограмме осуществляется оператором GOSUB, за которым следует начальный номер строки или метка подпрограммы, в которую передается управление. Подпрограмма выполняется до тех пор, пока не встретится оператор RETURN, который возвращает управление к строке, следующей непосредственно за GOSUB.

2. Подготовка первичных данных для статистической обработки.

2.1. Характер биологических и экологических данных.

Для биологических и экологических объектов характерно варьирование значений отдельных признаков. Подобная естественная вариабельность обусловлена влиянием на изучаемые объекты многочисленных случайных факторов. Кроме того, на величине признаков сказывается погрешность измерений.

Поскольку признаки могут принимать различные значения, они называются переменными величинами или просто переменными (variables).

В математической статистике различают три типа значений переменных - количественные, номинальные (качественные, атрибутивные, категориальные) и порядковые (ранговые, ординальные):

1. Значения количественных переменных являются числовыми, могут быть упорядочены и для них имеют смысл различные вычисления, например, средних величин и показателей вариации. Количественные признаки делятся на счетные или меристические (число ветвистых лучей в спинном плавнике рыбы) и мерные или метрические (размер, вес особи). Первые варьируют дискретно, вторые - непрерывно.

2. Значения номинальных переменных (например: пол животного, вид растений, цвет глаз) являются нечисловыми, они означают принадлежность к некоторым классам и не могут быть упорядочены или непосредственно использованы в вычислениях.

3. Порядковые переменные занимают промежуточное положение: их значения упорядочены (например: стадия развития животного или растения, уровень эвтрофикации водоема, степень загрязнения окружающей среды), но не могут быть с уверенностью измерены и сопоставлены количественно. Они в большей или меньшей степени обладают качеством, выраженным данной переменной. Однако они не позволяют сказать "на сколько больше" или "на сколько меньше".

2.2. Получение исходных данных для обработки и анализа.

Источниками первичных данных являются: наблюдения (включая мониторинг), эксперименты в полевых и лабораторных условиях. Для получения исходных данных

как правило производится ограниченная выборка из некоторой, часто неограниченной, генеральной совокупности или популяции объектов или явлений. Выборочный метод широко применяется в практике биологических и экологических исследований.

Когда для каждого объекта в выборке измерено значение одной переменной, и выборка называется одномерной. Если же для каждого объекта регистрируются значения двух или нескольких переменных, то такие данные называются многомерными.

Основной задачей статистики является получение на основе анализа выборки достоверных сведений об интересующих исследователя характеристиках генеральной совокупности. Поэтому основным требованием в выборке является ее репрезентативность, т.е. правильная представимость в ней пропорций генеральной совокупности. Одним из способов достижения репрезентативности является такая организация наблюдения или эксперимента, при которой элементы выборки извлекаются из генеральной совокупности случайным образом.

Выборочные данные регистрируются в специальных журналах, дневниках, бланках, ведомостях и т.д. Например:

1. Учетная (маршрутная) ведомость встречаемости птиц.
2. Бланки обработки гидробиологических проб.
3. Ихтиологический журнал траловой съемки озера.
4. Лабораторный журнал обработки проб.

В качестве материала для обработки и анализа на практических занятиях могут быть использованы данные, полученные:

- при выполнении курсовых и дипломных работ;
- при выполнении исследовательских работ во время прохождения студентами учебно-полевой практики по ботанике и зоологии;
- на лабораторных занятиях по физиологии человека и животных;
- при выполнении научно-исследовательских работ аспирантами и преподавателями кафедр;
- из учебников, учебных пособий и монографий по биологической, экологической и природоохранной тематике;
- из разнотипных баз данных с первичной информацией.

При необходимости морфометрическую и физиологическую информацию можно получить для студентов, используя соответствующую приборную базу факультета. По крайней мере, частоту сердечного пульса перед контрольной работой и через некоторое время после ее окончания можно легко измерить прямо в компьютерном классе.

2.3. Ввод данных.

Прежде чем приступить к непосредственному вводу первичных данных в таблицы пакетов MS Excel или Statistica, необходимо выполнить несколько предварительных процедур:

1. Сохранение файла (пока еще без данных) на жестком диске, причем для англоязычных версий пакета Statistica желательно, чтобы название файла и директории, в которых он будет размещен, имели названия из латинских символов.
2. Создание структуры базы данных.
3. Организация проверки данных при вводе.

В электронных таблицах рекомендуется разнести первичные данные и результаты обработки по разным листам книги, предварительно подписав закладки листов.

Для того чтобы задать структуру базы данных, необходимо дать названия столбцам таблицы и указать характеристики вводимых данных. При этом отдельные переменные (признаки объекта исследования) следует располагать в отдельных столбцах таблицы. В системе Statistica для каждого столбца можно задать параметры и провести экспресс-анализ данных (рис. 6, кнопки Values/Stats и Graphs).

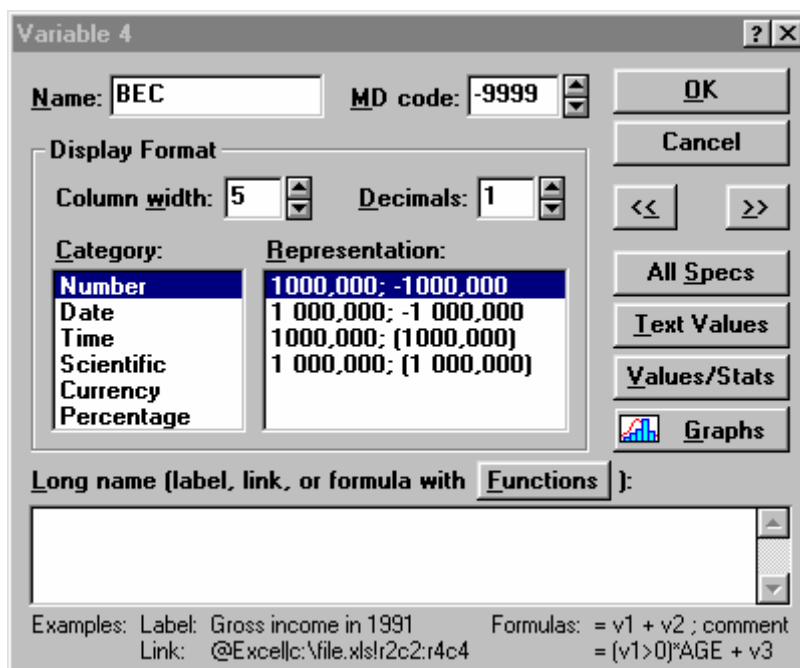


Рис. 6. Задание параметров переменной в пакете Statistica. **Name** - название переменной, **MD code** - коды пропущенных значений, **Column width** - ширина столбца, **Text Values** - работа с кодами символьных переменных, **Values/Stats** - представление всех вариантов значений переменных в виде ранжированного ряда и основных статистических показателей, **Graphs** - вывод гистограммы с кривой нормального распределения.

Для организации проверки вводимых значений в MS Excel можно воспользоваться процедурой условного форматирования путем задания диапазона изменения значений переменной. Например, при вводе данных по температуре человека, можно задать диапазон возможных значений в соответствии с нормой 36,0-37,0°C. Выход за границы данного диапазона свидетельствует или о заболевании, или об ошибке в источнике информации, или ошибке при вводе данных.

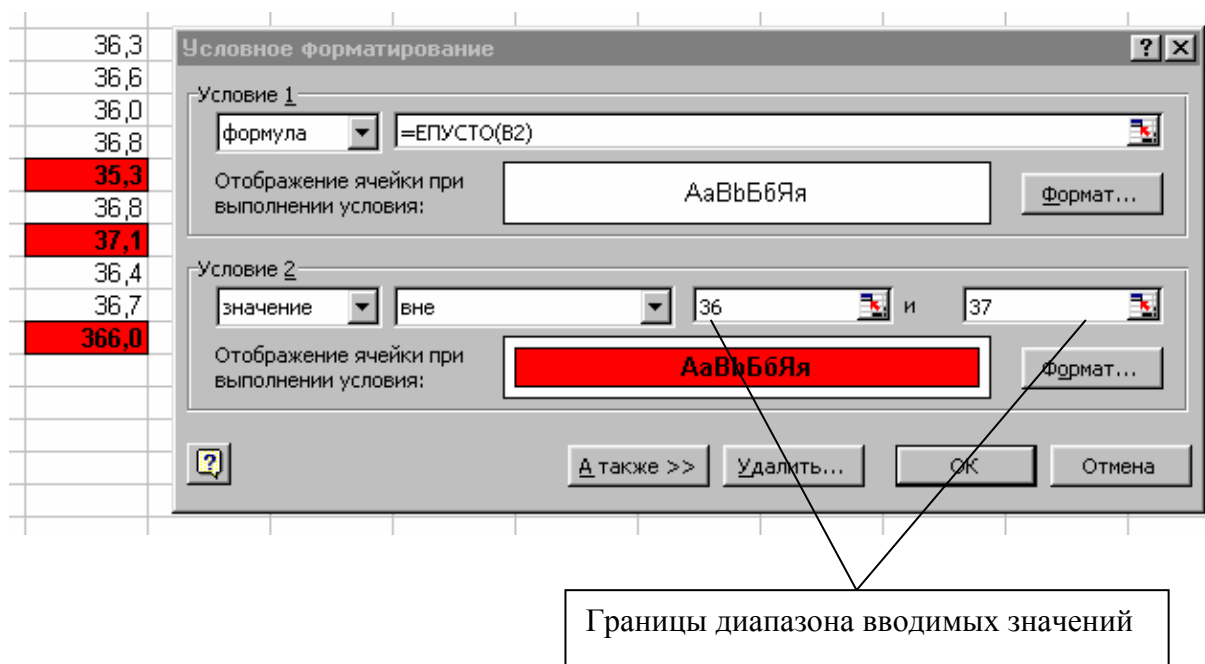


Рис. 7. Проверка вводимых данных в MS Excel с использованием процедуры условного форматирования.

УПРАЖНЕНИЕ 2.1. Ввод данных в таблицу MS Excel.

Задание. Произвести ввод данных в таблицу MS Excel.

Исходные данные взять из таблицы 2.

Таблица 3

Количество лучей в хвостовом плавнике камбалы (n=100).

53	51	52	55	56	49	51	52	54	56
54	53	52	53	51	55	53	55	53	54
51	51	56	54	54	53	54	53	55	53
52	55	53	53	56	53	52	56	52	52
56	55	50	54	49	54	54	55	54	55
52	51	55	52	55	54	51	54	53	54
54	56	54	55	53	53	56	55	54	53
55	52	53	52	51	55	53	54	51	50
53	54	55	52	55	52	53	50	53	52
58	57	57	58	56	57	56	58	57	57

Порядок работы:

1. Организовать проверку данных при вводе, указав границы диапазона вводимых чисел в рабочем окне процедуры условного форматирования.
2. Ввести в столбец "А" электронной таблицы исходные данные. При этом умышленно ввести несколько чисел вне диапазона (например, 5 и 499), проверить реакцию на это программы и немедленно исправить ошибочные значения.
3. Сохранить таблицу с данными на жестком диске для дальнейшего использования.

3. Статистическая обработка и анализ данных.

3.1. Предварительная обработка выборочных данных

Эмпирические данные, полученные в результате исследования (наблюдения или эксперимента), как правило, первоначально представляют беспорядочный набор чисел. Для того чтобы на основании их анализа сделать правильные научные выводы, необходимо упорядочить исходные данные путем группировки.

Результаты упорядочения эмпирических данных могут быть представлены в виде статистических рядов и их графиков. Статистические ряды подразделяются на:

- ранжированные;
- вариационные;
- временные (динамические);
- эмпирические ряды регрессии.

Ранжированные и вариационные ряды иллюстрируют варьирование признака.

Ранжированный ряд - одинарный, применяется для выборок малого объема ($n < 30$). Создается путем сортировки количественных эмпирических данных по возрастанию или убыванию значений признака. По ранжированному ряду можно определить размах изменчивости и найти наиболее часто встречающееся значение.

Вариационный ряд (ряд распределения) - двойной, применяется для выборок как малого, так и большого объема. При распределении дискретных величин получается прерывистый вариационный ряд (безинтервальный), а при распределении непрерывно варьирующих признаков - непрерывный вариационный ряд (интервальный). Последние бывают равно- и неравноинтервальные.

Временные ряды - двойные, отражают изменение значения признака во времени.

Эмпирические ряды регрессии - двойные, отражают связь между признаками.

Графики. По данным вариационного ряда можно построить следующие графики и распределения: гистограмму, полигон, кумуляту. Гистограмма и полигон распределения строятся в осях: значения классовых вариантов (абсцисса) - частоты классовых вари-

ант (ордината). Кумулята строится в осях: значения классовых вариантов (абсцисса) - накопленные частоты классовых вариантов (ордината).

По данным рядов динамики строится график в осях: время (абсцисса) - значение признака (ордината).











На основе эмпирических рядов регрессии можно построить графики (диаграммы рассеяния), отражающие связь между сопряженными признаками.

Вышеперечисленные графики служат для наглядной иллюстрации закономерностей, заключенных в статистических рядах.

Часто в экспедиционных условиях вариационные ряды удобнее формировать сразу в полевых дневниках или журналах учетных съемок. Это бывает необходимо, например, при массовых промерах животных и растений. Для этого применяется метод точкования - разнесение отдельных вариантов выборочной совокупности по заранее определенным классам будущего вариационного ряда, используя следующий шифр частот:

Таблица 4

Шифры частот вариационного ряда.

									
1	2	3	4	5	6	7	8	9	10

При обработке на компьютере вариационный ряд, полученный таким способом, может использоваться непосредственно, а при необходимости преобразован в ранжированный ряд.

УПРАЖНЕНИЕ 3.1. Построение ранжированного ряда.

Задание. Построить ранжированные ряды с использованием электронных таблиц MS Excel.

Исходные данные:

1. Размеры тела горала (в см): 115 106 113 109 110 118 117 110.
2. Вес птенцов сорокопута (в г): 12,6 14,6 20,0 24,0 26,0 27,6 30,6 25,0 37,5 38,0 13,8 17,0 21,8 22,0 23,0 22,0 11,0.

3. Количество личинок жуков-щелкунов на 1 м² почвы: 3 7 1 11 7 8 7 2 7 3 6 3 1 3.
4. Количество микроспоридий в тканях капустной белянки:
- а) самки: 84 58 18 135 107 46 129 97 51 132 71 178 89;
- б) самцы: 317 298 312 277 328 319 299 345 285 362 320.
5. Число делящихся клеток в поле зрения: 53 48 56 59 59 58 53 42 52 43 52 53 50 36 42 52 43 52 52 49 59 59 57 42 52 43 50.
6. Диаметр клетки (в мк): 14 19 28 18 22 23 20 21 27 17 24 21 20 21 23 19 23 22 21 22 24 20 24 20.
7. Время выживания моллюсков лигий (в ч):
- а) в 100%-ной морской воде: 9 4 9 9 9 10 9 5 9 6 7 9 9 9 10 9 9 7 7 9;
- б) в 75%-ной морской воде: 7 7 7 10 7 7 5 7 7 7.
8. Длина тела живородящей ящерицы (в мм): 55 54 45 73 63 62 58.
9. Количество зубцов на 1 см края листа у нивяника обыкновенного: 6 7 5 1 2 5 4 8 9 8 9 8 7 10 10 11 9 8 8 9 10.

УПРАЖНЕНИЕ 3.2. Построение вариационного ряда.

Задание. Построить вариационный ряд и соответствующие графики: (гистограмму, вариационную кривую, полигон и кумуляту).

Источник данных. Таблица 3.

Порядок выполнения.

1. В тетради построить вариационный ряд путем точкования, добавить строку накопленных частот.
2. Построить вариационную кривую и кумуляту.
3. Перенести вариационный ряд в электронные таблицы MS Excel.
4. Используя мастер диаграмм, построить график распределения частот в выборке в виде гистограммы, полигона, вариационной кривой, кумуляты.
5. Таким же образом в электронных таблицах обработать дополнительные наборы данных.

Дополнительные наборы данных:

1. Число бедренных пор у живородящих ящериц: 8 9 11 10 11 10 8 9 9 7 9 9 7 11 10 10 11 11 9 10 9 10 10 9 10 10 10 11 8 11 .10 9 11 10 10 12 10 9 8 10 10 9 10 8 10 9 9 8 9 8 9 8 10 8 9 8 9 9 9 8 10 11 10 11 14 11 11 9 11 9 9 9 9 10 10 8 12 9 12 9 13 9 10 11 11 11 11 11 11 10 11 12 11 11 11 11 10 11 10 10 9 10 9 9 9 11 12 11 10 11 12 11 10 11 9 11 12 8 14 12 15 12 11 13 12 12 11 10 12 11 10 11 9 5 12 10 11 11 9 11 10 11 11 12 11 9 11 10 7 7 10 11 9 6 7 8 11.

2. Количество язычковых цветков в соцветиях нивяника обыкновенного (Карпаты):

а) сухой луг: 21 23 23 20 18 21 25 19 29 24 23 21 27 21 20 20 8 16 12 24 21 15 21 22 18 19 17 20 20 22 21 21 20 20 21 13 15 14 21 20 22 18 21 18 13 15 22 16 14 18 21 25 13 12 13 17 14 15 17 20 20 16 17 21 28 21 21 24 19 21 20 14 13 14 16 20 15 14 20 13 20 19 18 13 17 20 21 21 22 21 19 15 13 21 21 10 14 13 13 16 17 13 21 21 19 14 18 12 27 21 13 8 8 16 21 18 13 14 15 13 20 19 21 18 16 22 20 20 16 13 16 20 18 12 12 11 17 13 14 20 16 15 14 17 15 21 15 12 8 21 21 13 16 16 12 19 13 21 14 18 14 13 15 17 19 18 15 17 13 11 15 16 14 13 13 20 13 13 12 13 13 13 10 18 17 17 21 20 13 10 13 14 13 14 16 12;

б) сырой луг: 25 20 25 26 21 24 16 28 23 28 28 32 23 23 26 22 .29 34 27 28 25 27 27 25 25 23 29 22 24 23 23 27 21 32 24 36 27 29 22 23 24 22 24 24 30 34 26 22 22 27 26 24 27 28 22 21 23 24 32 31 21 23 24 28 25 26 28 21 22 24 22 24 23 28 23 23 23 23 23 21 35 34 33 30 28 34 37 32 31 23 28 33 22 32 34 24 32 24 27 23 23 33 27 34 27 32 33 18 28 24 34 30 32 24 ИЗ* 28 32 24 21 23 30 26 21 33 30 23 27 25 29 32 30 30 34 26 22 32 23 34 32 23 33 22 26 26 21 26 24 33 27 27 24 24 33 35 24 33 21 32 33 23 30 25 30 24 27 28 26 33 35 26 31 31 32 24 21 34 23 26 36 40 31 24 23 26 26 31 22 26 31 26 32 30 32 33 27;

в) опушка леса: 21 22 24 26 23 22 24 21 28 30 21 20 21 22 23 18 29 24 20 24 22 16 23 24 23 22 21 21 30 20 25 18 22 21 19 22 23 23 19 21 22 27 21 21 21 23 26 16 21 21 25 24 17 23 34 24 34 17 24 23 33 23 33 25 23 22 24 20 21 24 23 16 31 16 21 19 21 26 19 19 21 22 21 20 20 22 23 21 21 22 22 25 29 20 24 21 25 25 21 21 25 24 26 21 27 19 20 21 19 21 21 24 20 20 24 24 21 29 23 21 30 25 20 19 21 23 21 21 22 21 24 22 20 22 21 24 21 25 22 21 23 21 23 30 21 20 18 28 27 21 17 22 28 15 34 24 22 22 25 22 29 "21 29 24 22 21 24 20 19 30 18 18 17 21 21 21 23 21 21 13 21 17 21 22 29 26 23 23 28 15 24 27 26 22 26 20 19 23.

3. Длина зерен пшеницы (в мм): 5,39 5,42 5,38 5,47 5,51 5,30 5,40 5,40 5,28 5,45,46,5,53 5,55 5,47 5,24 5,44 5,54 5,66 5,43 5,42 5,43 5,52 5,45 5,26 5,33 5,42 5,50 5,49 5,47 5,52 5,48 5,34 5,36 5,59 5,45 5,44 5,34 5,33 5,41 5,54 5,49 5,35 5,40 5,26 5,50 5,46 5,41 5,55 5,31 5,45 5,54 5,32 5,52 5,39 5,62 5,40 5,23 5,45 5,47 5,40 5,42 5,45 5,32 5,44 5,58 5,50 5,36

5,44 5,50 5,37 5,47 5,50 5,44 5,28 5,31 5,64 5,46 5,47 5,57 5,58 5,45 5,37 5,48 5,46
5,51 5,29 5,42 5,69 5,60 5,45 5,38 5,46 5,52 5,43 5,18 5,61 5,36 5,39 5,44 5,41.

УПРАЖНЕНИЕ 3.3 Преобразование вариационного ряда в ранжированный.

Задание. Преобразовать вариационный ряд в ранжированный.

Источник данных. Использовать вариационный ряд по распределению количества лучей в хвостовом плавнике камбалы, построенный методом точкования в тетрадах.

Порядок выполнения.

1. Написать программу на Бэйсике.
2. Представить ранжированный ряд в виде текстового файла.

3.3. Основные характеристики выборочной совокупности.

Средние величины

Средняя арифметическая. Наиболее часто применяется и представляет собой частное от деления суммы всех вариантов совокупности на объем совокупности.

Средняя квадратическая. Для того чтобы вычислить среднее арифметическое значение площади на основании замеров линейного показателя x_i , характеризующего эту площадь, необходимо определить среднюю квадратическую для линейного показателя.

Например, для того, чтобы найти среднее значение площади круга или поверхности шара по результатам замера диаметра этих фигур, необходимо вычислить среднее значение диаметра. Полученное значение можно потом подставить в соответствующие формулы для нахождения площади упомянутых фигур

Средняя кубическая. Для того чтобы вычислить среднее арифметическое значение объема на основании замеров линейного показателя x_i , характеризующего этот объем, необходимо определить среднюю кубическую для линейного показателя.

Среднюю квадратическую и среднюю кубическую вычисляют в тех случаях, когда определяющим признаком является соответственно площадь тела или его объем. Например, средний диаметр корзины подсолнечника находят по формуле средней квадратической, а средний диаметр горошины - по формуле средней кубической.

Средняя гармоническая. Применяется для нахождения среднего значения скорости передвижения человека или животного, а также скорости протекания какого-либо процесса

Средняя геометрическая. Используется при исследовании скорости прироста какой-то величины с течением времени. Средняя геометрическая характеризует процесс.

Средняя геометрическая обычно используется при анализе признаков, величина которых во времени изменяется по закону геометрической прогрессии. Сюда относятся: изменение веса тела в начальном периоде роста организма; рост численности популяции в естественных условиях; рост поголовья стада в животноводстве и т.д.

Скорость прироста часто выражают в относительных величинах. Относительную скорость роста можно вычислить по формуле Ч.Майнота:

$$V = \frac{t_2 - t_1}{t_1} \cdot 100\%,$$

где t_1 и t_2 – значения признака в начале и конце исследуемого отрезка времени.

Если вычислены величины относительной скорости роста V_1, V_2, V_3 и т.д. для последовательных равных промежутков времени, то средняя относительная скорость роста (средняя геометрическая) для всего периода исследования вычисляется по следующей формуле:

$$\bar{x}_g = \sqrt[n]{V_1 \cdot V_2 \cdot V_3 \dots V_n}$$

Мода - это значение, наиболее часто встречающееся в выборочной совокупности. На гистограмме мода соответствует самому высокому столбцу, а на полигоне или вариационной кривой – самой высокой точке графика.

Медиана - является серединой вариационного ряда. По обе стороны от медианы находится одинаковое количество вариантов. Медиану достаточно просто определить графически с помощью кумуляты. Для этого на оси ординат откладываем величину $n/2$, а на оси абсцисс находят M_e

Квантили отсекают в пределах вариационного ряда определенную часть его членов.

Квартили - это три значения признака (Q_1, Q_2, Q_3), делящие ранжированный ряд на четыре равные части. Аналогично, девять децилей делят ряд на 10 равных частей, а 99 перцентилей - на 100 равных частей.

Показатели вариации

Средние величины ничего не говорят о степени вариабельности признаков. Одинаковые значения средней могут иметь выборочные совокупности, варьирующие в разных пределах.

Для характеристики вариабельности биологических и экологических признаков применяются следующие показатели: пределы вариации, размах вариации, дисперсия, среднее квадратическое отклонение, коэффициент вариации.

Пределы вариации (лимиты) - это значения минимальной и максимальной вариант выборочной совокупности. Лимиты указывают границы варьирования признака.

Пределы вариации (лимиты) - это значения минимальной и максимальной вариант выборочной совокупности. Лимиты указывают границы варьирования признака.

Размах вариации - представляет собой разность между максимальной и минимальной вариантами

Дисперсия и стандартное отклонение (среднее квадратическое отклонение). Характеризуют разброс значений вокруг среднего значения.

Коэффициент вариации. Все рассмотренные выше показатели вариации являются величинами именованными, имеющими размерность вариант. Для того, чтобы можно было сравнить между собой степень вариабельности признаков, имеющих как одинаковые, так и разные размерности, введен безразмерный показатель - коэффициент вариации.

$$C = \frac{s_x}{x} \cdot 100\%$$

Распределения

Асимметрия и эксцесс. При изучении распределений, отличных от нормального, возникает необходимость количественно оценить это различие. С этой целью вводят специальные характеристики, в частности асимметрию и эксцесс. Для нормального распределения эти характеристики равны нулю. Поэтому если для изучаемого распре-

деления асимметрия и эксцесс имеют небольшие значения, то можно предположить близость этого распределения к нормальному. Наоборот, большие значения асимметрии и эксцесса указывают на значительное отклонение от нормального.

Наряду с простыми, одновершинными эмпирическими распределениями могут быть и полимодальные (многовершинные) распределения, частным случаем которых являются бимодальные. Это означает гетерогенность (неоднородность) исходного материала или влияние специфического фактора.

Основные выборочные характеристики могут быть вычислены в электронных таблицах MS Excel в модуле "Описательные статистики" пакета "Анализ данных" (табл. 5).

Таблица 5

Результаты работы модуля "Описательная статистика" пакета "Анализ данных" MS Excel.

<i>РОСТ</i>		<i>ВЕС</i>	
Среднее	119,8816901	Среднее	21,61408451
Стандартная ошибка	0,729676508	Стандартная ошибка	0,374376992
Медиана	120	Медиана	21,1
Мода	117	Мода	20
Стандартное отклонение	6,148363542	Стандартное отклонение	3,154556603
Дисперсия выборки	37,80237425	Дисперсия выборки	9,951227364
Эксцесс	-0,293650547	Эксцесс	-0,227346052
Асимметричность	-0,039091917	Асимметричность	0,407386664
Интервал	31	Интервал	14,2
Минимум	104	Минимум	15,5
Максимум	135	Максимум	29,7
Сумма	8511,6	Сумма	1534,6
Счет	71	Счет	71

В пакете Statistica для этих целей служит раздел "Описательная статистика" модуля "Основные статистики и таблицы" (рис. 8).

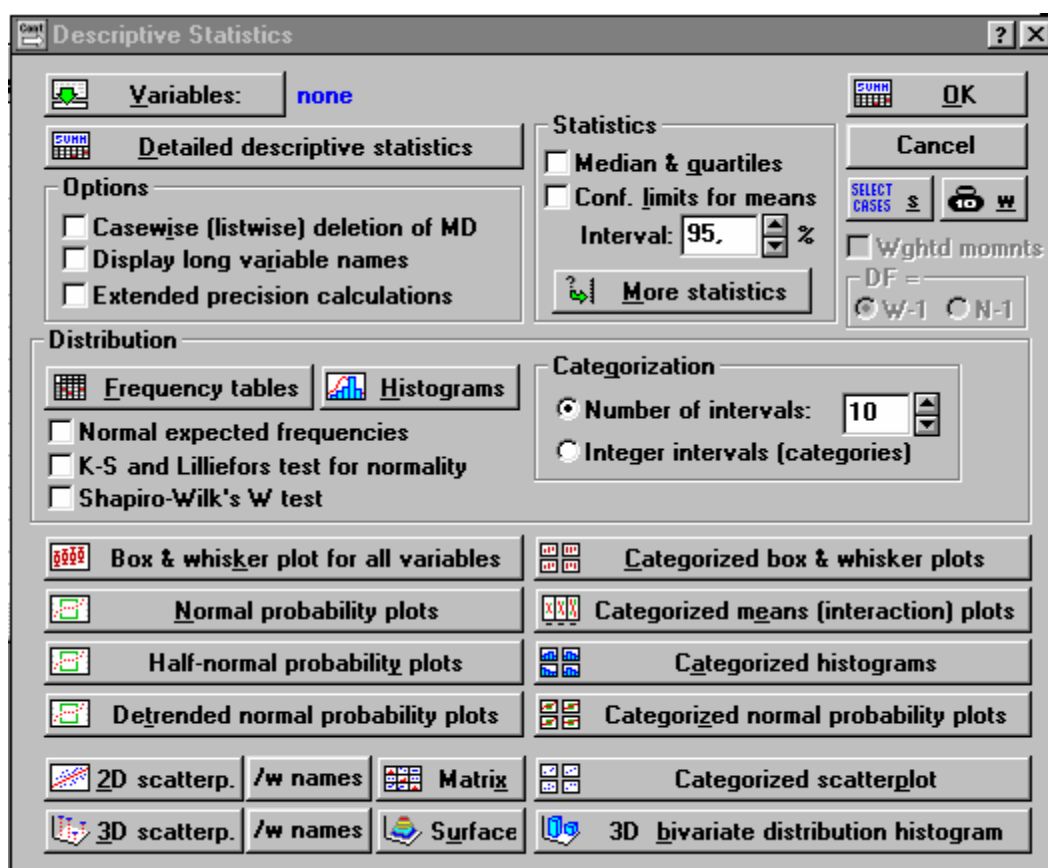


Рис. 8. Диалоговое окно "Описательная статистика".

Названия основных кнопок:

Variables – выбор переменных для анализа,

Frequency tables – построение вариационного ряда,

Histograms – построение гистограммы и кривой нормального распределения,

More statistics – выбор статистических показателей из списка.

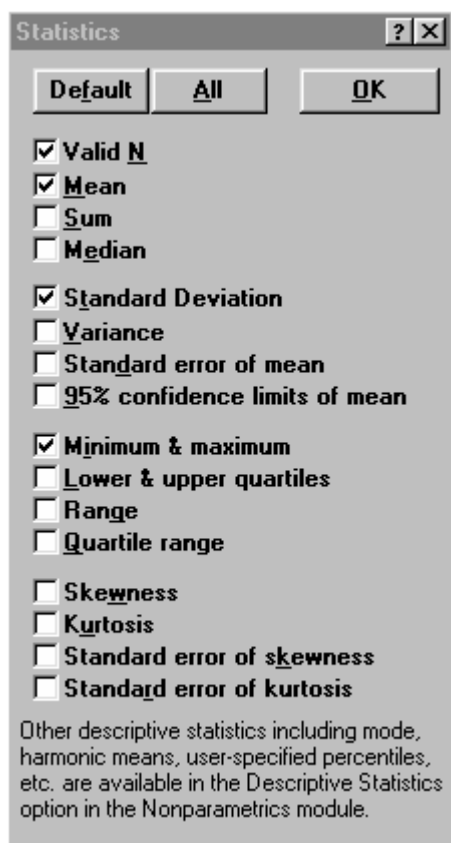


Рис. 9. Основные описательные статистики.

Valid N - объем выборки, **Mean** - среднее арифметическое, **Median** - медиана;

Standard Deviation - стандартное отклонение, **Variance** - дисперсия,

Standard error of mean - стандартная ошибка,

95% confidence limits of mean - доверительный интервала для среднего;

Minimum & maximum - минимальное и максимальное значения (лимиты),

Lower & upper quartiles - нижняя и верхняя квартили,

Range - размах вариации, **Quartile range** - квартильный размах;

Skewness - асимметрия, **Kurtosis** - эксцесс,

Standard errors of skewness & kurtosis - ошибки асимметрии и эксцесса.

УПРАЖНЕНИЕ 3.4. Вычисление средних величин.

Задание 1. По исходным данным вычислить среднюю арифметическую, среднюю квадратическую и среднюю кубическую. Сравнить эти средние между собой.

Исходные данные. Использовать вариационный ряд, отражающий распределение горошин по диаметру x (мм):

x	4.8	5.2	5.6	6.0	6.4	6.8
p	4	10	14	13	8	1

Порядок работы.

1. Вычисления произвести в электронных таблицах Ms Excel по соответствующим формулам:

$$\bar{x} = \frac{\sum x_i p_i}{n} \text{ – среднее арифметическое,}$$

$$\bar{x}_{\text{кв.}} = \sqrt{\frac{\sum x_i^2 p_i}{n}} \text{ – среднее квадратическое,}$$

$$\bar{x}_{\text{кб.}} = \sqrt{\frac{\sum x_i^3 p_i}{n}} \text{ – среднее кубическое}$$

2. Сравнить эти средние между собой.

Задание 2. Рассчитать средний геометрический прирост длины и веса ребенка н первом году жизни. Приросты рассчитать по формуле Ч.Майнота.

Исходные данные. Табл. 6.

Таблица 6

Рост ребенка на первом году жизни.

Возраст, месяцы	Длина тела, см	Масса тела, г
0	50	3250
1	54	4000
2	58	4700
3	60	5350
4	62	5950
5	64	6500
6	65	7000
7	66	7450
8	67	7850
9	68	8250
10	69	8500
11	71	8750
12	72	9000

УПРАЖНЕНИЕ 3.5. Расчет основных выборочных характеристик.

Задание 1. Рассчитать основные характеристики выборочной совокупности в MS Excel.

Исходные данные. Использовать таблицу 2.

Порядок работы. Используя формулы, функции и модуль "Описательная статистика" из пакета "Анализ данных" рассчитать:

1. Объем выборки.
2. Среднее арифметическое.
3. Пределы вариации (лимиты).
4. Размах вариации.
5. Дисперсию.
6. Стандартное отклонение.
7. Коэффициент вариации.
8. Стандартную ошибку.
9. Границы доверительного интервала.

Задание 2. Вычислить те же показатели, используя пакет Statistica.

3.3. Критерии достоверности статистических оценок.

Поскольку математическая статистика имеет дело с данными, достаточно ограниченными по объему и в которых интересующие исследователя закономерности искажены различными случайными факторами, то большинство статистических вычислений сопровождается проверкой некоторых предположений или гипотез о источнике этих данных.

Основное проверяемое предположение называется нулевой гипотезой и формулируется как отсутствие различий, влияния фактора, связи между признаками и т.п. Другое проверяемое предположение называется конкурирующей или альтернативной гипотезой.

Для нулевой гипотезы задается критерий ее проверки, согласно которому вычисляется (по заданному правилу или формуле) значение соответствующей статистики

и уровень значимости P , который представляет собой вероятность отвергнуть нулевую гипотезу тогда, когда она на самом деле верна (так называемая ошибка 1-го рода).

Если вычисленное значение P превосходит выбранный уровень $P_{кр}$, который называется критическим значением (обычно $P_{кр}=0.05$), то нулевая гипотеза может быть принята (данные не противоречат нулевой гипотезе), в противном случае (при $P<0.05$, иногда для уверенности при $P<0.01$ или даже $P<0.001$) нулевая гипотеза отвергается и принимается альтернативная гипотеза.

Анализ эмпирических распределений

При анализе формы вариационной кривой производится оценка степени соответствия исходного эмпирического распределения предполагаемому теоретическому. Сам процесс выявления соответствия осуществляется путем сравнения эмпирического распределения с уже известными теоретическими (нормальным, экспоненциальным, биномиальным, пуассоновским и др.) и носит название подгонки. Особенно часто возникает необходимость определения нормальности эмпирического распределения. От этого, как правило, зависит выбор метода дальнейшего анализа данных.

Рассмотрим проверку нормальности эмпирического распределения на уже известном наборе данных (табл. 3) в пакете Statistica с использованием критерия согласия χ^2 Пирсона. Поскольку этот критерий относится к группе непараметрических, необходимо задействовать модуль Nonparametric Statistics (рис. 5) и его раздел Distribution fitting (подгона распределений) (рис. 10). В появившемся окне надо выбрать исследуемую переменную из списка (Variable), и нажать на кнопку **Graph**.

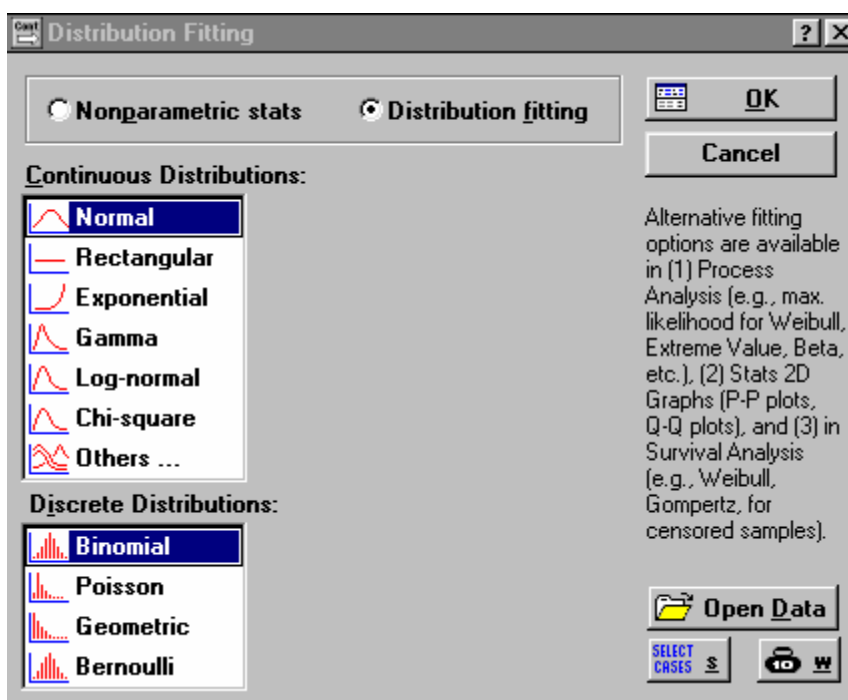


Рис. 10. Окно выбора теоретического распределения. **Continuous Distribution** – непрерывные распределения, **Discrete Distribution** – дискретные.

На графике (рис.) изображена гистограмма частотного распределения признака и соответствующая кривая нормального распределения. В заголовке графика указаны параметры тестов Колмогорова-Смирнова и χ^2 Пирсона. Для того, чтобы сделать заключение о нормальности эмпирического распределения, предварительно необходимо сформулировать нулевую гипотезу: распределение количества лучей в хвостовом пла- вике камбалы не отличается от нормального. Достигнутый уровень значимости ($p=0,166$) больше критического ($p_{кр}=0.05$) свидетельствует о том, что эмпирическое распределение не отличается от нормального.

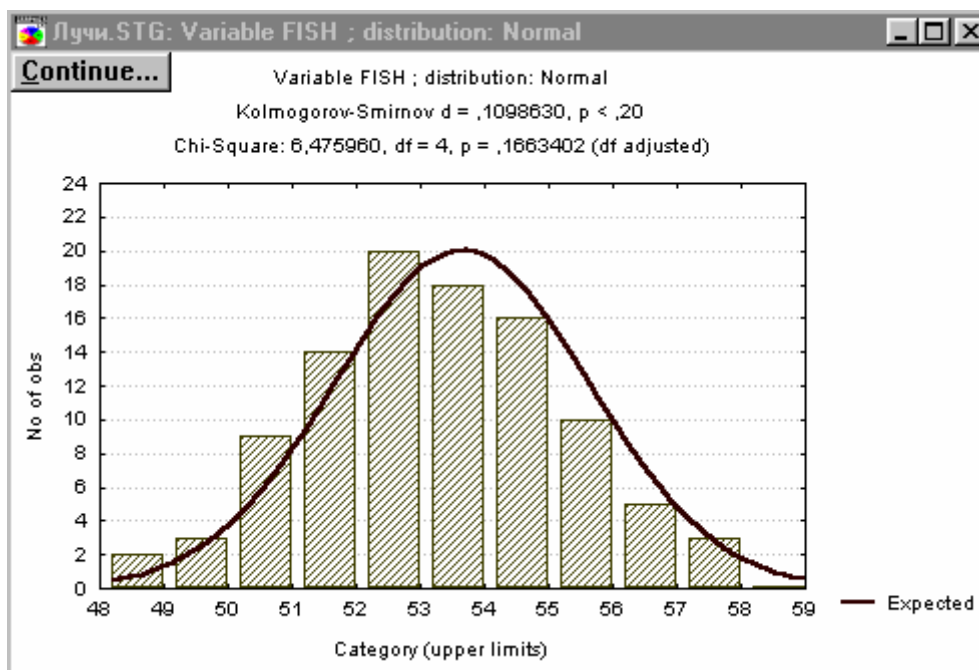


Рис. . Сравнение эмпирического распределения количества лучей в хвостовом плавнике камбалы с теоретическим нормальным.

Выявление достоверности различий между выборками

Для выявления достоверности различий между двумя выборками применяются параметрические или непараметрические критерии. К параметрическим относятся t-критерий Стьюдента и F-критерий Фишера.

t-критерий является наиболее часто используемым методом обнаружения различия между средними двух выборок. Теоретически, t-критерий может применяться, даже если размеры выборок очень небольшие (например, 10; некоторые исследователи утверждают, что можно исследовать выборки меньшего размера), и если переменные нормально распределены, а дисперсии наблюдений в группах не слишком различны. Предположение о нормальности можно проверить, исследуя распределение (например, визуально с помощью гистограммы) или применяя какой-либо критерий нормальности (например χ^2). Равенство дисперсий в двух группах можно проверить с помощью F-критерия Фишера. Если условия применимости t-критерия не выполнены, следует использовать непараметрические альтернативы t-критерия (например, U-тест Манна-Уитни).

В начале сравнения двух выборок рекомендуется провести графический анализ с использованием диаграмм размаха, выводя на экран одновременно значения средних, стандартных ошибок и границ доверительного интервала (рис. 11).

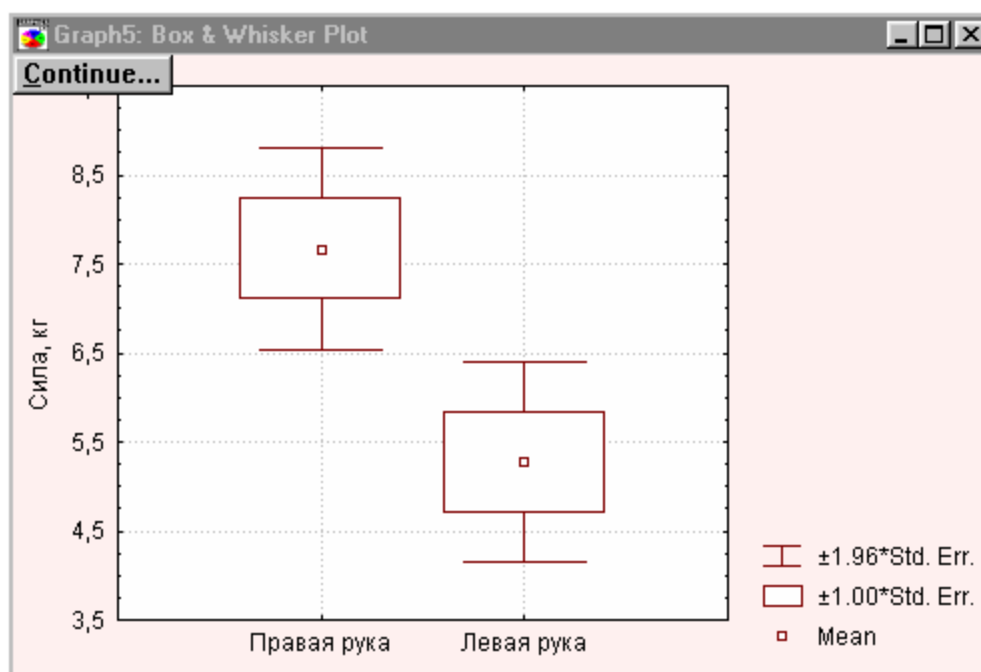


Рис. 11. Диаграммы размаха значений признака в двух попарно связанных выборках. **Mean** - среднее значение, $\pm 1.00 \text{ Std. Err.}$ - стандартная ошибка, $\pm 1.96 \text{ Std. Err.}$ - границы доверительного интервала.

Непересекающиеся границы доверительных интервалов для средних свидетельствуют о наличии достоверных различий в силе правой и левой руки у детей младшего школьного возраста.

В качестве примера рассмотрим применение t-критерия Стьюдента для выяснения достоверности различий средних величин двух независимых выборок в электронных таблицах. Воспользуемся базой данных по детям младшего школьного возраста (Приложение 1). Сравним вес детей двух групп - 6 и 7 лет. Нормальность распределения значений признака в выборке можно определить по гистограмме, сравнивая ее с соответствующей теоретической кривой нормального распределения (рис. 12). Эту процедуру надо применить для каждой выборки. Более надежную оценку дает использование критерия согласия Пирсона χ^2 или Колмогорова-Смирнова, рассматриваемые далее.

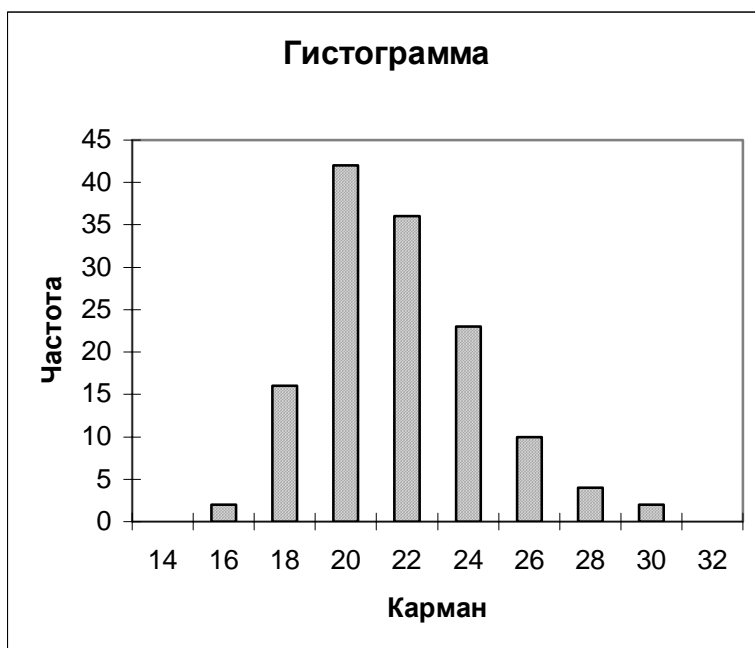


Рис. 12. Гистограмма распределения детей по весу в выборке 1.

Равенство дисперсий проверим с помощью F-критерия Фишера. Предварительно следует сформулировать нулевую гипотезу: различия между дисперсиями веса двух возрастных групп детей не достоверны, а если и наблюдаются между выборочными дисперсиями, то это случайное явление, в генеральной совокупности эта разница равна нулю.

Таблица 7

Результаты проверки гипотезы о равенстве генеральных дисперсий с помощью F-критерия Фишера.

Двухвыборочный F-тест для дисперсии		
	<i>Переменная 1</i>	<i>Переменная 2</i>
Среднее	20,94740741	23,25068493
Дисперсия	7,353258154	6,132534247
Наблюдения	135	73
Df	134	72
F	1,199057006	
P(F<=f) одностороннее	0,198478853	
F критическое одностороннее	1,421813778	

Достигнутый уровень значимости (0,198478853) меньше критического (0.05) и нулевую гипотезу о равенстве дисперсий следует принять.

Поскольку выборочные распределения близки к нормальному, а по результатам предыдущего теста можно говорить о равенстве генеральных дисперсий, мы имеем полное право использовать t-критерий Стьюдента, так как выполнены условия его применения. Кроме того, немаловажное значение имеют и объемы сравниваемых выборок. В данном случае, они равны 135 и 73, что вполне приемлемо для применения данного критерия. Из списка модулей выбираем "Двухвыборочный t-тест с одинаковыми дисперсиями", имея виду, что выборки независимы, иначе пришлось бы использовать "Парный двухвыборочный t-тест для средних". В случае выявления достоверных различий между дисперсиями, можно воспользоваться модулем "Двухвыборочный t-тест с различными дисперсиями", в котором используется специальный метод расчета t-критерия.

Предварительно следует сформулировать нулевую гипотезу: различия между средним весом двух возрастных групп детей не достоверны, а если и наблюдаются между выборочными средними, то это случайное явление, в генеральной совокупности эта разница равна нулю. Результаты анализа представлены в табл. 8.

Таблица 8

Результаты проверки гипотезы о равенстве генеральных средних с помощью t-критерия Стьюдента.

Двухвыборочный t-тест с одинаковыми дисперсиями		
	<i>Переменная 1</i>	<i>Переменная 2</i>
Среднее	20,94740741	23,25068493
Дисперсия	7,353258154	6,132534247
Наблюдения	135	73
Объединенная дисперсия	6,926597371	
Гипотетическая разность средних	0	
Df	206	
t-статистика	-6,023972001	
P(T<=t) одностороннее	3,86085E-09	
t критическое одностороннее	1,652283572	
P(T<=t) двухстороннее	7,72171E-09	
t критическое двухстороннее	1,971548045	

Пусть вас не смущает форма записи достигнутого уровня значимости $7,72171\text{E-}09$, поскольку это просто компактный способ представления числа $0,00000000772171$ или, по-другому, $7,72171 \cdot 10^{-9}$.

Поскольку величина достигнутого уровня значимости меньше критического уровня 0.05 (и даже 0.001), нулевая гипотеза отклоняется на высоком уровне значимости ($p < 0.001$) и можно сделать заключение о достоверном различии веса детей двух возрастных групп.

Параметрический F-критерий Фишера и непараметрический критерий согласия Пирсона можно использовать и самостоятельно, вне связи с процедурой исследования применимости t-критерия Стьюдента.

Упражнение 3.6. Оценка достоверности различия между средними арифметическими независимых выборочных совокупностей.

Задание. Для исследуемых выборочных совокупностей определить достоверность различия между средними арифметическими.

Исходные данные. База данных по детям младшего школьного возраста (Приложение 1). Использовать следующие параметры: рост, вес, окружность грудной клетки и силу правой и левой руки. В качестве сравниваемых выборок использовать половые и возрастные группы детей.

Порядок выполнения задания. Оценку провести с использованием пакета "Анализ данных" электронных таблиц MS Excel и модуля "Basic statistics" пакета Statistica. В последнем случае для проверки нормальности выборочных распределений следует воспользоваться критерием согласия Пирсона в модуле "Непараметрическая статистика".

Упражнение 3.7. Оценка достоверности различия между средними арифметическими независимых выборочных совокупностей.

Задание. Проверить гипотезу о достоверности различий между попарно связанными выборками.

Источник данных. Измерить пульс студентов до начала и через полчаса после проведения контрольной работы и внести данные в сводную таблицу.

3.4. Дисперсионный анализ.

Понять суть и эффективность дисперсионного анализа проще на разборе конкретного примера. Обратимся к данным С.П. Китаева (1984) о зависимости биомассы бентоса озер разных природных зон от площади озер.

Таблица 9

Биомасса бентоса (г/м²) в зависимости от площади озера и природной зоны.

Площадь озер	Природная зона			
	Тундра	Северная тайга	Средняя тайга	Смешанный лес
<10 га	5.00	4.04	5.17	3.73
10-50	3.21	3.85	3.92	7.56
50-100 га	3.75	4.02	5.10	8.74
100-500 га	3.85	3.32	3.80	8.75
500-1000 га	5.39	2.23	3.77	15.39
1000-5000 га	3.59	2.11	2.56	11.85
>5000 га	8.44	1.87	1.94	9.84

При анализе таблицы трудно определить, как зависит биомасса бентоса от площади озера и природной зоны, в которой оно расположено. Единственно, что действительно бросается в глаза - это высокие биомассы бентоса в озерах площадью 500-5000 га в зоне смешанных лесов.

Попробуем выяснить, существует ли связь между биомассой бентоса, площадью озера и природной зоной, в которой расположено озеро. В данном случае можно использовать метод дисперсионного анализа, который широко применяется в биологии и экологии. Для этого воспользуемся модулем "Двухфакторный дисперсионный анализ без повторений" пакета "Анализ данных" из MS Excel. Предварительно необходимо сформировать дисперсионный комплекс, в котором столбцы представляют градации фактора "Природная зона", а строки – градации фактора "Площадь озера". Затем следует сформулировать нулевую гипотезу: рассматриваемые факторы не влияют на результирующий признак (площадь озера). В результате анализа на отдельном листе будет выведена следующая таблица:

Таблица 10

Результаты двухфакторного дисперсионного анализа.

Двухфакторный дисперсионный анализ без повторений						
<i>ИТОГИ</i>	<i>Счет</i>	<i>Сумма</i>	<i>Среднее</i>	<i>Дисперсия</i>		
<10 га	4	17,94	4,485	0,500833333		
10-50	4	18,54	4,635	3,904566667		
50-100 га	4	21,61	5,4025	5,290825		
100-500 га	4	19,72	4,93	6,5426		
500-1000 га	4	26,78	6,695	35,26596667		
1000-5000 га	4	20,11	5,0275	21,07109167		
>5000 га	4	22,09	5,5225	17,77589167		
Тундра	7	33,23	4,747142857	3,26922381		
Северная тайга	7	21,44	3,062857143	0,930190476		
Средняя тайга	7	26,26	3,751428571	1,427147619		
Смешанный лес	7	65,86	9,408571429	13,07804762		
Дисперсионный анализ						
<i>Источник вариации</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-Значение</i>	<i>F критическое</i>
Строки	13,202	6	2,200333333	0,399956952	0,869358583	2,66130229
Столбцы	172,0297	3	57,34322262	10,42333913	0,000334607	3,159911444
Погрешность	99,02566	18	5,501425397			
Итого	284,2573	27				

Рассмотрим результаты дисперсионного анализа. Достигнутый уровень значимости $p=0,869$ для площади озер больше критического (0,05). Следовательно, нулевую гипотезу необходимо принять. Для природной зоны $p=0,00033$, и нулевую гипотезу следует отвергнуть на уровне значимости 0,001. Альтернативная же гласит, что расположение озера в той или иной природной зоне достоверно влияет на биомассу бентоса.

Упражнение 3.7. Однофакторный дисперсионный анализ.

Задание. Определить влияние температуры на активность амилазы слюны человека при постоянном значении pH среды.

Для исходных данных с помощью дисперсионного анализа установить, влияют ли регулируемый фактор на величину признака и определить уровень значимости сделанного заключения.

Исходные данные. Получены по результатам биохимического анализа проб слюны, взятых у трех обследуемых одного пола и возраста, физически здоровых (то есть имеется три повторности). Исходные данные представить в виде однофакторного комплекса.

Таблица 11

Результаты эксперимента по исследованию влияния температуры на активность амилазы.

Температура, С°	Активность (усл. ед.) по повторностям		
	1	2	3
20	5.40	4.50	9.80
30	14.30	10.36	20.35
40	29.50	25.00	37.10
50	42.00	40.20	44.60

Упражнение 3.8. Двухфакторный дисперсионный анализ.

Задание. Определить влияние температуры (фактор В) и pH среды (фактор А) на активность амилазы слюны человека при постоянном значении pH среды.

Для исходных данных с помощью дисперсионного анализа установить, влияет ли регулируемый фактор на величину признака и определить уровень значимости сделанного заключения.

Исходные данные. Получены по результатам биохимического анализа проб слюны, взятых у тех же трех обследуемых. Исходные данные представить в виде двухфакторного комплекса с повторениями.

Таблица 12

Результаты эксперимента по исследованию влияния температуры и pH на активность амилазы.

Температура, С°	Активность (усл. ед.) по повторностям								
	pH 6.8			pH 7.2			pH 7.6		
20	12.3	16.1	16.9	5.4	4.5	9.8	3.6	3.3	5.7
30	28.2	35.1	38.5	14.3	10.4	20.3	7.8	7.6	11.2
40	57.8	71.6	77.8	29.5	25.0	37.1	15.4	16.0	21.4
50	79.8	99.0	107.5	42.0	40.2	44.6	21.3	22.1	29.6

3.5. Корреляционный анализ.

Наличие корреляционных связей между варьирующими признаками обнаруживается на всех уровнях организации живой материи. Ключевым понятием, описывающим связи между переменными величинами, является корреляция.

Корреляция представляет собой меру зависимости переменных и может быть выражена количественно коэффициентом корреляции. Знак коэффициента корреляции определяет направление связи ("+" – положительная, "-" – отрицательная). Так, на диаграмме рассеяния (рис.) "облако точек" и линия регрессии ориентированы в направлении из нижнего левого угла координатной плоскости в правый верхний, что указывает на положительную корреляцию между ростом и весом.

Коэффициенты корреляции изменяются в пределах от -1 до +1. Значение -1 означает, что переменные имеют строгую отрицательную корреляцию, при +1 - строгую положительную. Значение 0 означает отсутствие корреляции.

При $r > 0$ (по абсолютному значению) – есть корреляция. В этом случае можно говорить о степени сопряженности между варьирующими признаками. Если $r < 0.3$, то это слабая связь. Если $0.3 < r < 0.5$, то эта связь умеренная; $0.5 < r < 0.7$ – значительная корреляция; $0.7 < r < 0.9$ – сильная корреляция; $r > 0.9$ – очень сильная связь, приближающаяся к функциональной.

Корреляция между парой переменных называется *парной корреляцией*, а ее величина определяется *коэффициентом парной корреляции*. Если исследуется зависимость между двумя переменными, измеренными в *интервальной шкале*, наиболее подходящим коэффициентом будет коэффициент корреляции Пирсона r , называемый так-

же коэффициентом линейной корреляции. Эта корреляция наиболее популярна, поэтому часто, когда говорят о корреляции, имеют в виду именно корреляцию Пирсона.

Корреляционный анализ рекомендуется проводить по определенной схеме:

1. Выявление наличия связи между признаками.
2. Определение характера распределения переменных (нормальное или нет).
3. Определение формы связи (линейная или нелинейная).
4. Определение направления связи (положительная или отрицательная).
5. Выбор метода расчета показателя связи (параметрический или непараметрический).
6. Расчет коэффициента корреляции (степени сопряженности признаков).
7. Расчет ошибки и оценка достоверности коэффициента корреляции.

В электронных таблицах MS Excel имеется возможность вычисления коэффициента линейной корреляции Пирсона с помощью функции КОРРЕЛ или модуля "Корреляция" пакета "Анализ данных". Результаты в первом случае выводятся в виде коэффициента корреляции, во втором – в виде корреляционной матрицы без указания достоверности коэффициентов корреляции. Кроме того, с помощью мастера диаграмм можно построить диаграмму рассеяния.

В пакете Statistica (рис. 13) одновременно с коэффициентом корреляции Пирсона определяется его достоверность, о которой можно судить по достигнутому уровню значимости – при $p < 0,05$ коэффициент корреляции достоверно отличается от нуля.

Для выявления наличия связи, ее формы и направления полезно воспользоваться диаграммами рассеяния. На рис. 14 и 15 представлены две зависимости между переменными и соответствующие коэффициенты линейной корреляции. Во втором случае, поскольку зависимость между двумя использованием линейного коэффициента корреляции Пирсона некорректно, необходимо использовать непараметрический аналог – коэффициент корреляции Спирмена.

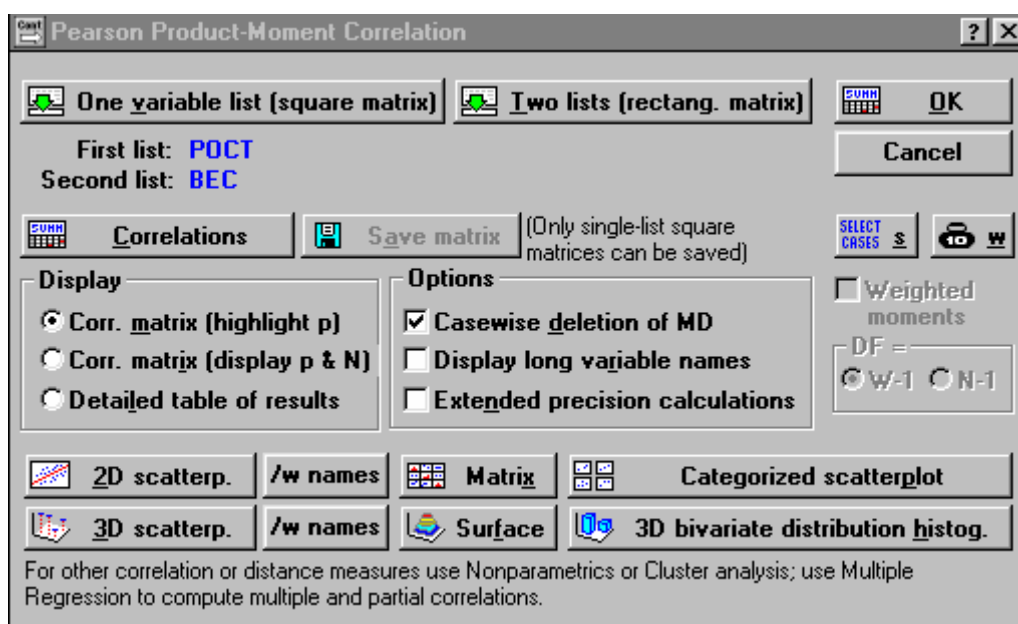


Рис. 13 . Диалоговое окно "Корреляция Пирсона".

Названия основных кнопок:

One variable list (square matrix) – один список переменных (квадратн. матрица),

Two lists (rectang. matrix) – два списка (прямоугольная матрица),

Correlations – построение корреляционной матрицы,

2D scatterp. – двумерные диаграммы рассеяния.

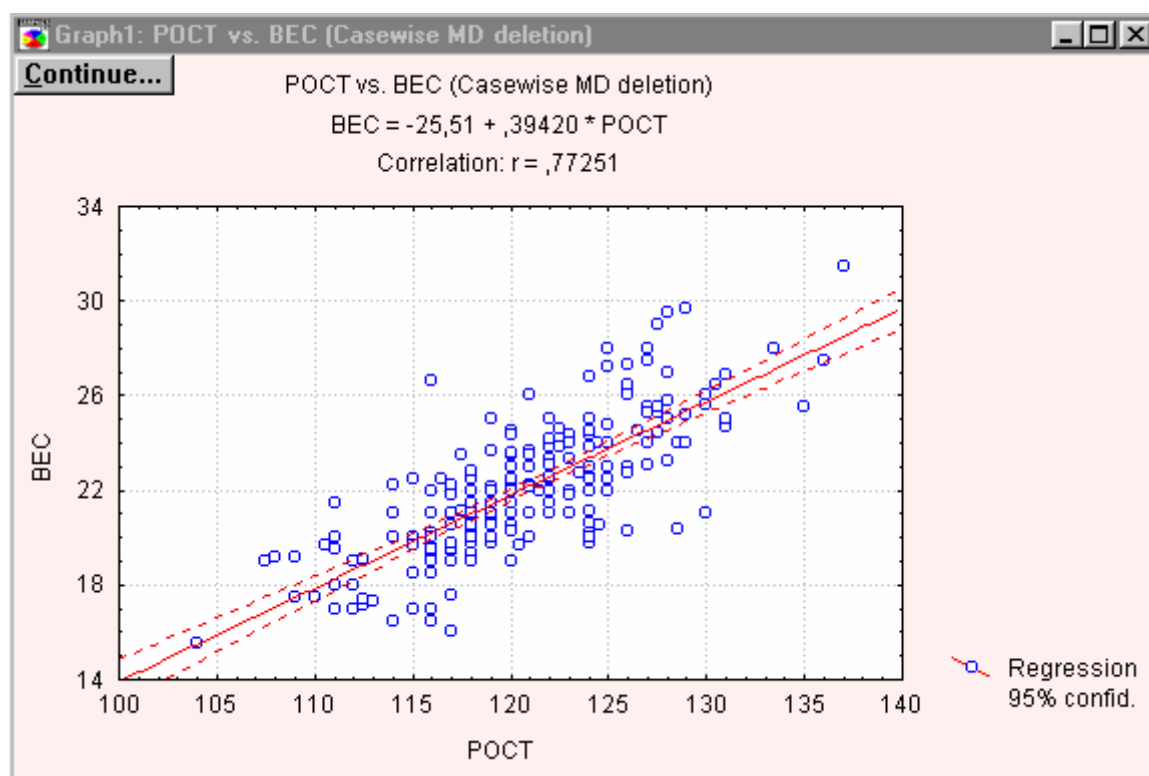


Рис. 14 . Диаграмма рассеяния, характеризующая связь между ростом и весом детей 6-7 лет. **Regression 95% confid.** – 95%-ный доверительный интервал для линии регрессии.

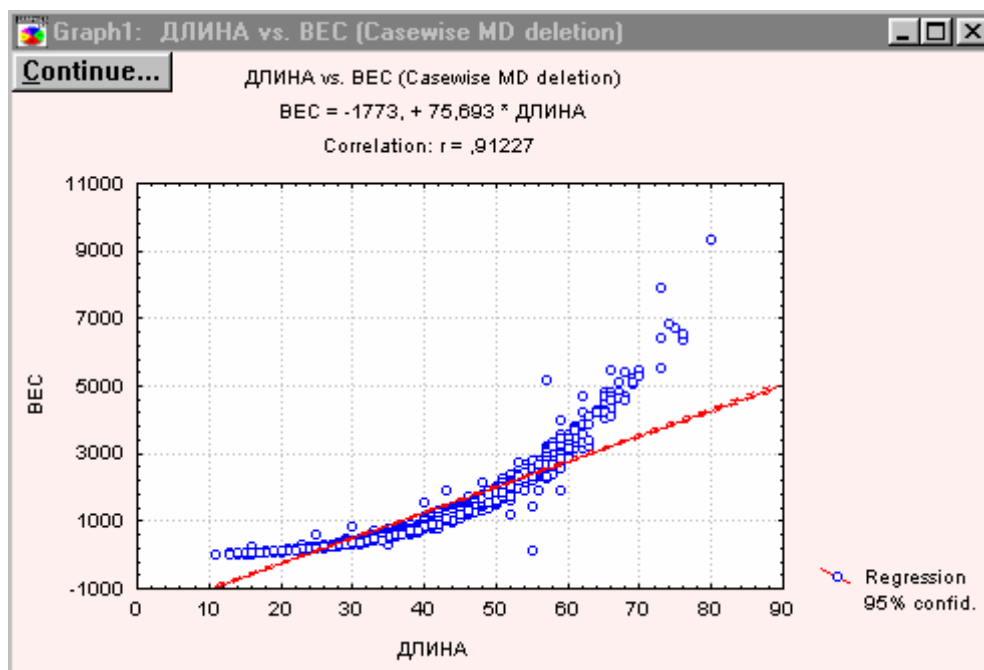


Рис. 15. Диаграмма рассеяния, характеризующая связь между длиной и весом судака.

УПРАЖНЕНИЕ 3.9. Вычисление коэффициента корреляции.

Задание. Для заданной совокупности попарно связанных значений нескольких признаков вычислить коэффициент линейной корреляции Пирсона, оценить его достоверность, установить степень сопряженности между признаками и направление связи.

Исходные данные.

1. Базы данных по детям младшего школьного возраста (Приложение 1)
2. База данных по лимнологическим параметрам озер (Приложение 2).
3. Вес теляпии в г (x), коэффициент зрелости (k), количество икринок (y) и вес одного ооцита в мг (o):

x: 13,1 14,5 14,6 17,4 27,5 17,5 12,6 11,9 11,3 22,3 22,9 10,1 10,2

k: 5,1 6,3 5,8 6,1 7,6 5,2 6,4 6,3 6,2 5,0 5,9 5,0 5,7

y: 319 307 239 290 429 270 252 243 232 201 383 280 254

o: 1,6 2,2 2,8 3,1 3,9 2,6 2,6 2,1 2,2 3,3 2,7 1,5 1,1

4. Длина крыла полевого жаворонка в мм (x) и географическая широта места в градусах (y):

x: 104 101 98 111 103 96 95 94 105 110 101 100 92 100 91 84 95 102 115 115 112

y: 67 56 38 47 60 44 45 52 52 50 48 60 39 52 37 42 43 62 44 57 61

Порядок работы.

1. В электронных таблицах MS Excel и пакете Statistica построить диаграммы рассеяния значений по парам признаков. Выявить наличие корреляции и направление связи.

2. Построить корреляционные матрицы. Определить степень и направление связи между признаками, оценить достоверность коэффициентов корреляции.

Не следует обольщаться большими значениями коэффициента корреляции при малом объеме выборки. Проведите следующий эксперимент: из большой выборки используйте ее части разного объема (по возрастанию) и посмотрите на закономерность изменения коэффициента корреляции.

3.6. Регрессионный анализ.

Одним из основных методов исследования зависимости между двумя переменными является регрессионный анализ, предназначенный для подбора аналитической функции (модели), наиболее точно описывающей эмпирическую зависимость. В качестве такой функции используются:

- простые зависимости: линейная, экспоненциальная, степенная, логистическая, полиномиальная и др.;
- произвольные функции, задаваемые посредством алгебраических операций и функций.

Регрессионный анализ позволяет:

- проверить статистическую нулевую гипотезу об отсутствии различий между эмпирической зависимостью и теоретической функцией;

- предсказать значения зависимой переменной при промежуточных или лежащих вне диапазона измерения значениях независимой переменной.

При отсутствии априорных предположений о том, какому закону соответствует эмпирическая зависимость между переменными, целесообразно проверить несколько различных регрессионных моделей и среди тех, для которых нулевая гипотеза принимается, остановиться на модели с большим коэффициентом аппроксимации (r^2).

Для получения представления о характере связи между переменными (линейном или нелинейном) полезно построить функциональный график зависимости переменной-отклика от независимой переменной (рис. 14-15).

В электронных таблицах регрессионный анализ реализован в модуле "Регрессия" пакета "Анализ данных". Рабочее окно модуля представлено на рис. 16.

Выбор приемлемой функции осуществляется в окне "Линия тренда" (рис. 17), исходя из величины коэффициента аппроксимации (r^2). Для вывода на экран коэффициента аппроксимации и уравнения регрессии необходимо перейти на закладку "Параметры".

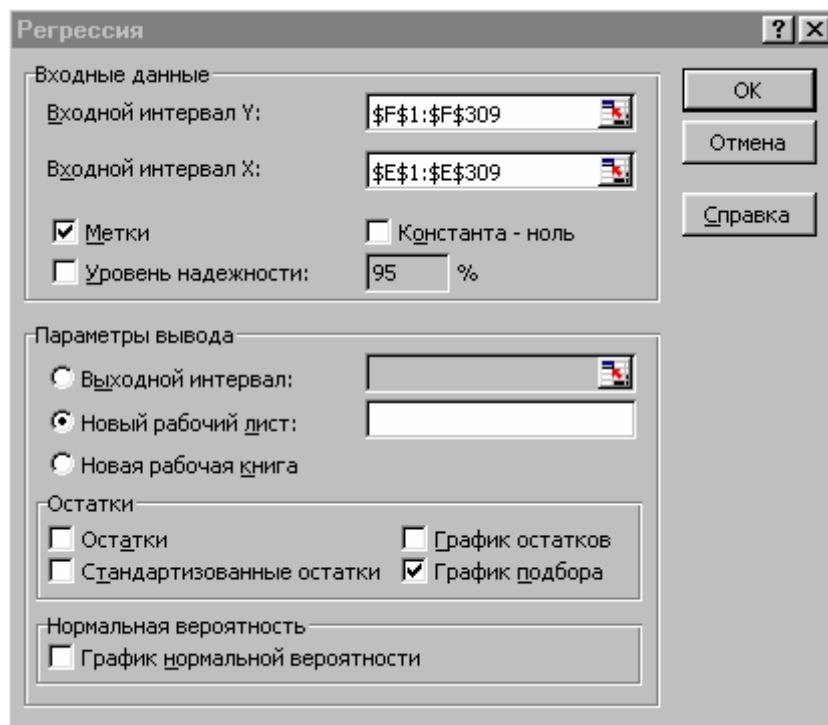


Рис. 16. Рабочее окно модуля "Регрессия".

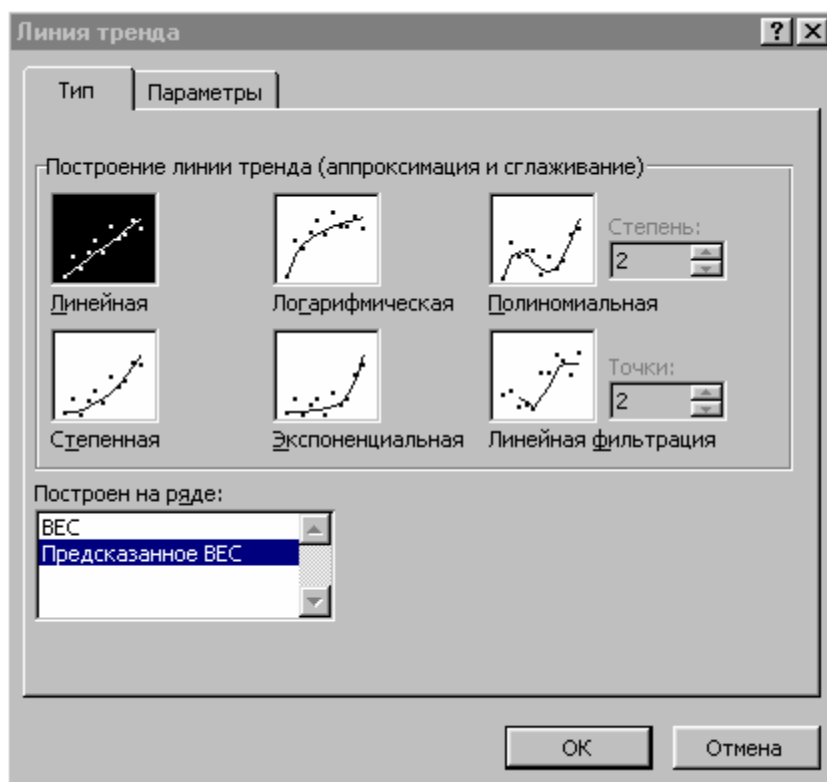


Рис. 17. Выбор типа зависимости между двумя переменными в модуле "Регрессия".

В пакете Statistica подбор теоретической линии регрессии к эмпирической зависимости можно провести в модуле "Нелинейное оценивание". В качестве примера рассмотрим анализ соотношения размера (см) и веса (г) судака на достаточно большой представительной выборке ($n=1813$). На рис. 18. представлен выбор функциональной зависимости веса от размера в виде уравнения $v2=a*v1^b$, где $v1$ - размер рыбы, $v2$ - вес рыбы, a и b – коэффициенты регрессионного уравнения, подбираемые компьютером.

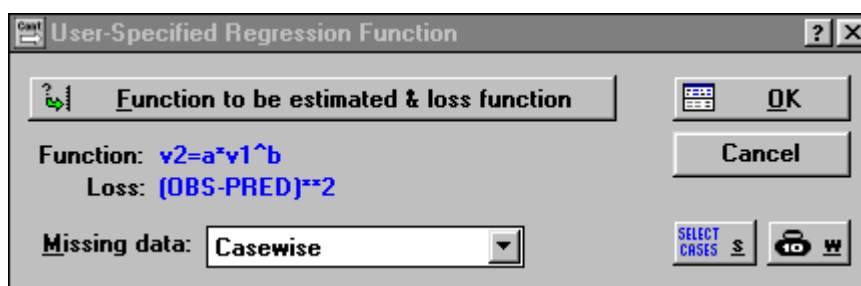


Рис. 18. Окно определения регрессионной функции пользователем.

Результаты подгонки теоретической линии регрессии к эмпирическим данным представлены на рис. Коэффициент корреляции в этом случае равен 0,989, а при использовании линейной зависимости – 0,912 (рис. 15, 19).

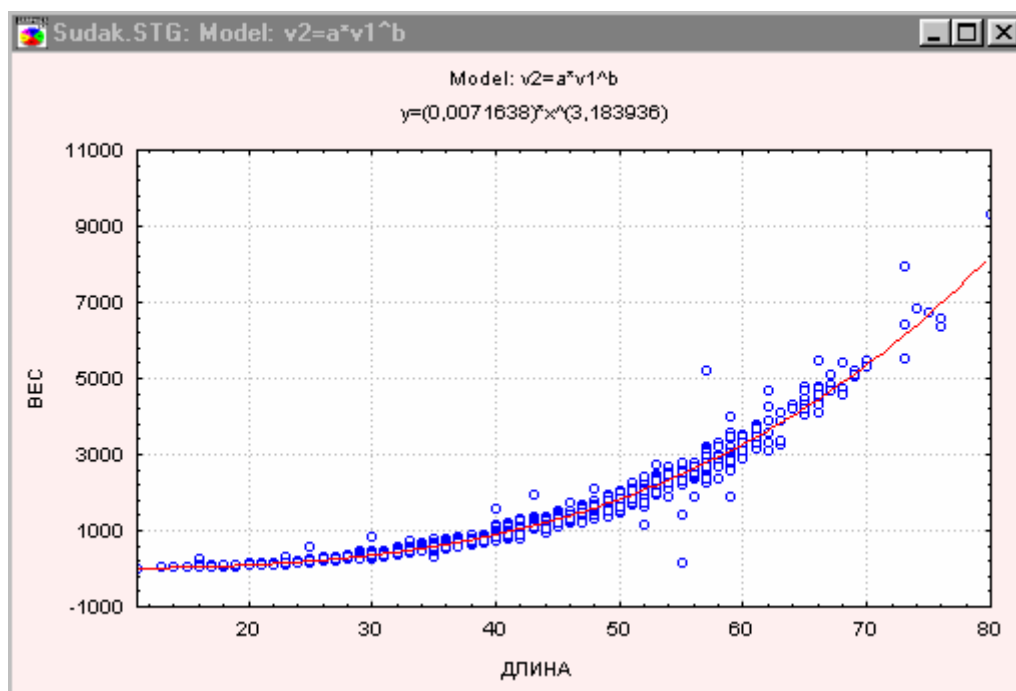


Рис. 19. Результаты подгонки теоретической степенной линии регрессии к эмпирическим данным.

УПРАЖНЕНИЕ 3.10. Регрессионный анализ.

Задание. Для заданной совокупности попарно связанных значений нескольких признаков провести регрессионный анализ

Исходные данные.

1. Базы данных по детям младшего школьного возраста (Приложение 1)
2. База данных по лимнологическим параметрам озер (Приложение 2).
3. Длина озерной лягушки в см (x) и индекс длинноноготности (y):

x :	2	4	6	8	10	12
y :	2,1	2,6	2,6	2,4	2,2	2,0
f :	118	187	134	112	65	5

4. Возраст в годах (x) и средний вес в г (y) у осетра:

x :	1	2	3	4	5	6	7	8
y :	49	82	119	150	206	275	357	375

4. Высота растения в см (x) и длина чашечки цветка в мм (y) у зубчатки поздней:

x :	49,5	89,5	129,5	169,5	209,5	249,5	289,5
y :	4,57	4,66	4,80	4,96	5,10	4,99	5,25

Порядок работы.

1. В электронных таблицах MS Excel и пакете Statistica построить диаграммы рассеяния значений по различным парам признаков.
2. Подобрать к эмпирическим зависимостям теоретические линии регрессии.
3. Осуществить вычисление функции при различных значениях аргумента.

Множественная регрессия

Общее назначение множественной регрессии (этот термин был впервые использован в работе Пирсона - Pearson, 1908) состоит в анализе связи между несколькими независимыми переменными (называемыми также регрессорами или предикторами) и зависимой переменной. Например, если мы рассматриваем зависимость между биомассой ерша (бентофаг) и факторами внешней среды, то в данном случае биомасса ерша - это зависимая переменная, а факторы внешней среды - независимые переменные или предикторы (абиотические – содержание растворного кислорода у дна; биотические – биомасса бентоса, биомасса хищников, биомасса бентофагов-конкурентов; антропогенные - интенсивность вылова).

Эту информацию можно использовать при анализе с помощью множественной регрессии для построения регрессионного уравнения в следующем виде:

$$\text{Биомасса ерша} = 0.3 \cdot \text{Кислород} + 0.8 \cdot \text{Бентос} - 0.4 \cdot \text{Хищники} - 0.1 \cdot \text{Конкуренты} - 0.2 \cdot \text{Вылов}$$

В биологии и экологии процедуры множественной регрессии чрезвычайно широко используются в исследованиях. В общем, множественная регрессия позволяет исследователю задать вопрос (и, вероятно, получить ответ) о том, "что является лучшим предиктором для...". Термин "множественная" указывает на наличие нескольких предикторов или регрессоров, которые используются в модели.

Анализ временных рядов

Существуют две основные цели анализа временных рядов: (1) определение природы ряда и (2) прогнозирование (предсказание будущих значений временного ряда по настоящим и прошлым значениям). Обе эти цели требуют, чтобы модель ряда была идентифицирована и, более или менее, формально описана. Как только модель определена, вы можете с ее помощью интерпретировать рассматриваемые данные.

Как и большинство других видов анализа, анализ временных рядов предполагает, что данные содержат систематическую составляющую (обычно включающую несколько компонент) и случайный шум (ошибку), который затрудняет обнаружение регулярных компонент. Большинство методов исследования временных рядов включает различные способы фильтрации шума, позволяющие увидеть регулярную составляющую более отчетливо.

Большинство регулярных составляющих временных рядов принадлежит к двум классам: они являются либо трендом, либо сезонной составляющей. Тренд представляет собой общую систематическую линейную или нелинейную компоненту, которая может изменяться во времени. Сезонная составляющая - это периодически повторяющаяся компонента. Оба эти вида регулярных компонент часто присутствуют в ряде одновременно.

Анализ временных рядов реализован в модуле Time series системы Statistica.

4. Математическое моделирование биологических и экологических систем.

4.1. Моделирование как метод исследования сложных систем.

С точки зрения системного подхода, все биологические и экологические объекты на любом уровне организации являются сложными системами. Их исследование осуществляется в рамках системного анализа, основным методом которого служит моделирование.

С появлением вычислительной техники перед биологами и экологами открылись новые возможности экспериментирования, так трудно реализуемые в природных условиях. Характер этого экспериментирования особый. Он позволяет производить опыты не с природными объектами, а с их подобиями, моделями, отражающими главные свойства, способы функционирования и взаимодействия, как живых организмов между собой, так и со средой их обитания.

Моделированием называется изучение свойств одного объекта путем исследования свойств другого, имеющего определенное сходство с первым. Тот объект, который непосредственно исследуется, принято называть моделью, а тот, к которому относится вывод - оригиналом.

На модели можно изучать только те свойства, которые на нее перенесены с оригинала. Моделирование позволяет масштабировать время функционирования модельных систем (сокращать или увеличивать) и выявлять "белые пятна" в наших представлениях об объектах исследования. Модель приобретает особую ценность, когда она позволяет прогнозировать их поведение.

4.2. Классификация моделей.

Существует большое количество классификаций моделей. Ниже приводится один из вариантов классификации моделей по способу их реализации:

Материальные

1. Геометрически подобные
2. Субстратно подобные

Идеальные

1. Неформализованные
2. Частично формализованные
3. Вполне формализованные

Материальные

Геометрически подобные - масштабные, воспроизводящие пространственно-геометрические характеристики оригинала безотносительно к его субстрату (макеты, муляжи органов животных и растений и др.).

Субстратно-подобные - воспроизводящие с масштабированием в пространстве и времени свойств и характеристик оригинала той же природы, что и модель (аквариумы, чучела животных, гербарии и др.).

Неформализованные

Концептуальные - системы представлений об объекте оригинале, сложившиеся в человеческом мозгу.

Частично формализованные

Вербальные - описание свойств и характеристик оригинала на некотором естественном языке.

Графические иконические - средствами графики представляют черты, свойства и характеристики оригинала, реально или хотя бы теоретически доступные непосредственному зрительному восприятию (изображения животных и растений, географические карты и др.).

Графические условные - воспроизводят средствами графики свойства и характеристики объекта-оригинала, которые даже в принципе не могут наблюдаться визуально. К числу подобных моделей относятся всевозможные графики, диаграммы и схемы, гистограммы, фазовые портреты и номограммы полученные в результате наблюдений и экспериментальных исследований.

Вполне формализованные

Графоаналитические - блок-схемы, в которых структура системы задается графически (блок-схема экосистемы с указанием структурообразующих связей).

Аналитические. При аналитическом моделировании тот или иной процесс выражается дифференциальным уравнением или системой таких уравнений (регрессионные модели, модели роста популяций, модель "хищник-жертва" и др.).

Аналитическое моделирование имеет два недостатка. Первый из них заключается в том, что если и удастся описать биологические процессы системами дифференциальных уравнений, то решить их часто бывает слишком затруднительно. Но главный недостаток не в этом. Многие жизненные процессы дискретны по существу и состоят из мало похожих между собой фаз. Между тем в основу аналитического моделирования положен подход непрерывности.

Алгоритмические (имитационные) математические модели. В них элементы и процессы оригинала описываются абстрактными символами и операциями над ними. При имитационном моделировании логическая структура исследуемого явления, последовательность протекания в нем процессов и состав информации сохраняются. При этом поведение системы в целом изучается на основании поведения ее частей и взаимодействия между этими частями.

Аналитические и имитационные модели объединяются в группу математических моделей.

Кроме вышеприведенной классификации математические модели можно разделить на несколько альтернативных групп:

- непрерывные, дискретные;
- детерминированные, вероятностные (стохастические);
- точечные, пространственные.

Упражнение 4.1. Классификация моделей.

Задание. Классифицировать модели, применяемые в процессе изучения биологических или экологических дисциплин на 1-3 курсах факультета.

Порядок выполнения.

1. Заполнить соответствующие анкеты (см. приложения 3, 4) в редакторе MS Word, используя схему классификации моделей по способу их реализации.
2. Выявить основные закономерности применения разнотипных моделей в различных циклах дисциплин.
3. Дать предложения по введению в учебный процесс новых моделей.

4.3. Моделирование популяций.

Аналитические модели роста популяций

Одной из простейших моделей роста популяции является модель, заданная дифференциальным уравнением

$$\frac{dN}{dt} = rN$$

где N - численность популяции в момент t , а r - коэффициент рождаемости.

Иными словами скорость увеличения численности популяции пропорциональна самой численности (N) и коэффициенту рождаемости (r). Это же уравнение можно представить в виде экспоненциальной модели роста популяции:

$$N = N_0 e^{rt}$$

где N_0 - численность в момент времени.

Подобные модели применяются в микробиологии, демографии, а также при описании динамики численности одноклеточных водорослей и животных, для которых на некоторых этапах их жизненного цикла действительно можно пренебречь влиянием смертности. Более реально описывает рост популяций логистическая модель:

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right)$$

Упражнение 4.2. Моделирование роста популяций.

Задание. Создать в электронных таблицах MS Excel и на языке Бэйсик модель экспоненциального и логистического роста популяции. В электронных таблицах построить график динамики численности.

Имитационные модели динамики численности популяций

Рассмотрим на простейшей имитационной модели популяции основные этапы моделирования биосистем. В качестве объекта моделирования предлагается использо-

вать популяцию снетка, небольшой по размерам короткоцикловой рыбы из семейства корюшковых. Это обычный обитатель водохранилищ Верхней Волги.

Постановка задачи. Моделирование отдельных популяций рыб проводится, как правило, для исследования их динамики численности. В данном случае модель разработана для анализа динамики численности популяции, определяемой нелинейной зависимостью между количеством отложенной икры (популяционной плодовитостью) и количеством появившейся из нее молоди.

Определение состава и структуры моделируемой системы. Для этой цели удобно использовать блок-схему – графическое изображение основных элементов и функциональных связей оригинала, переносимых на модель. Функциональные связи могут быть прямыми и обратными. Прямая связь увеличивает первоначальное воздействие. Обратная связь уменьшает первоначальное воздействие.

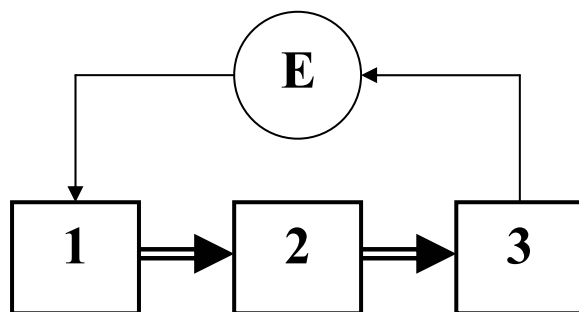


Рис. 20 . Блок-схема популяции короткоцикловой рыбы.

Выбор типа модели. Обычно разрабатываются имитационные модели надорганизменных систем, при этом не последнюю роль играет низкая точность биологических данных и зависимостей, которая иногда позволяет оценить состояние элементов изучаемой системы лишь на качественном уровне. Имитационное моделирование экосистем непосредственно связано с программной реализацией на ЭВМ.

Выбор шага моделирования. Переход имитационной модели из одного состояния в следующее осуществляется в дискретные моменты времени, расстояние между которыми называется временным шагом. Выбор величины временного шага является ответственным моментом при построении модели, так как все процессы, характерные времена которых оказываются меньше величины шага, выпадают из рассмотрения. Например, в модели с шагом 1 год не могут быть учтены сезонные миграции животных.

Задание переменных и констант модели. Идентификаторами являются названия переменных и констант, используемых в алгоритме и программе.

Таблица 13

Список переменных и констант модели популяции сетка.

Названия	Краткая характеристика
N1	Численность первой возрастной группы.
N2	Численность второй возрастной группы.
N3	Численность третьей возрастной группы.
Ks	Индивидуальная плодовитость.
E	Популяционная плодовитость.
a, b	Коэффициенты в формуле Риккера

Разработка алгоритма модели. Алгоритм - это систематизированное описание целенаправленной последовательности действий. В имитационном моделировании алгоритмом называется определенная последовательность применения формул и действий над величинами, позволяющая переходить модели от одного состояния к другому. Ниже приводится алгоритм работы модели популяции сетка:

Таблица 14

Алгоритм модели популяции сетка.

Формулы	Пояснения к алгоритму
$E = N_3^t(1 - m)Ks$	Популяционная плодовитость равна произведению числа оставшихся в живых трехгодовалых особей на индивидуальную плодовитость.
$N_3^{t+1} = N_2^t(1 - m)$	Численность третьей возрастной группы через год равна численности оставшихся в живых двухгодовиков.
$N_2^{t+1} = N_1^t(1 - m)$	Численность второй возрастной группы через год равна численности оставшихся в живых годовиков.
Если $E \geq E_k$	Если количество икры превышает критическое количество
то $N_1^{t+1} = N_k$	то численность первой возрастной группы через год равна критической численности,
иначе $N_1^{t+1} = \alpha E e^{-\beta E}$	иначе расчет по уравнению Риккера.

Программирование. Прежде чем писать свои собственные программы, необходимо научиться "читать" и понимать текст готовых программ, реализующих модели биосистем разной степени сложности.

Ниже приводится программа на Бейсике.

```

10 REM Имитационная модель популяции сетка (файл Popul.bas)
20 REM Автор - В.В.Меншуткин
30 REM *****
40 CLS ' Очистка экрана
50 INPUT " Смертность (проба 0.3) " M
60 INPUT " Плодовитость (проба 1000) " Ks
70 INPUT " Критическое число икринок на нерестилищах (проба 200000) " Ek
80 INPUT " Соответствующее число молодежи Nk (проба 100) " Nk
90 INPUT " Коэффициент 'альфа' в формуле Риккера (проба 0.03) " Alfa
100 PRINT
110 OPEN "Tabl.txt" FOR OUTPUT AS #1
120 Beta=(-1/Ek)*LOG(Nk/(Alfa*Ek))
130 IF Beta > 0 THEN 170
140 Beta=1/Ek
150 Alfa=2.71823*Nk/Ek
160 PRINT " У Вас получилась кривая Бивертон-Холта "
170 PRINT " У Вас получилась кривая Риккера "
180 PRINT
190 INPUT " Нажмите клавишу Enter для продолжения работы модели", Enter
200 CLS
210 REM Начальные численности возрастных групп
220 N1=300
230 N2=200
240 N3=100
250 REM Средний вес особей возрастных групп
260 W1=2
270 W2=4
280 W3=7
290 PRINT "Год  N1  N2  N3 Сумма  В      Р    Р/В "
```

```

300 PRINT #1, "Год  N1  N2  N3 Сумма  B      P    P/B "
310 FOR Y=1 TO 20 ' Цикл по годам
320     E2=Ks*(1-M)*N3
330     N3=N2*(1-M)
340     N2=N1*(1-M)
350     IF E2 > Ek THEN N1=Nk
360     IF E2 <= Ek THEN N1=Alfa*E2*EXP(-Beta*E2)
370     B=N1*W1+N2*W2+N3*W3 ' Биомасса
380     P=(W1/2)*(N1+N2)/2+(W2-W1)*(N2+N3)/2+(W3-W2)*(N3/2)/2 ' Продукция
390     Z$ = "## ####### ##### ##### ##### ##### #.##"
400     PRINT USING Z$; Y,N1,N2,N3,N1+N2+N3,B,P,P/B
410     PRINT #1, USING Z$; Y,N1,N2,N3,N1+N2+N3,B,P,P/B
420 NEXT Y
430 CLOSE #1
440 END

```

Отладка программы.

Идентификация модели. На заключительном этапе моделирования необходимо осуществить проверку соответствия (адекватности) модели объекту-оригиналу.

Использование модели. Постановка экспериментов, прогноз, оптимизация и т.д.

В Бейсике имеются достаточно развитые графические функции, но для анализа результатов экспериментальной работы с моделью лучше воспользоваться графическими возможностями электронных таблиц MS Excel или пакета Statistica. Для этого необходимо организовать экспорт выходной информации модели в виде текстового файла в соответствующий пакет.

Вывод результатов экспериментального изучения модели производится в текстовой файл "Tabl.txt", который в дальнейшем может быть подвергнут анализу в MS Excel или модуле "Анализ временных серий" пакета Statistica.

Таблица 15

Годы	N1	N2	N3	Сумма	В	Р	Р/В
1	501	210	140	851	2822	811	0.29
2	395	351	147	893	3223	981	0.30
3	376	277	246	898	3577	1033	0.29
4	153	263	194	610	2714	810	0.30
5	253	107	184	544	2223	609	0.27
6	277	177	75	529	1787	535	0.30
7	538	194	124	855	2719	776	0.29
8	440	376	136	952	3335	1022	0.31
9	408	308	263	980	3893	1127	0.29
10	127	286	216	628	2906	869	0.30
11	206	89	200	495	2167	586	0.27
12	239	144	62	445	1490	445	0.30
13	536	167	101	804	2447	695	0.28
14	499	375	117	991	3318	1017	0.31
15	459	349	262	1071	4153	1213	0.29
16	128	321	245	694	3253	974	0.30
17	154	90	225	469	2242	605	0.27
18	188	108	63	359	1248	366	0.29
19	536	132	76	744	2129	598	0.28
20	537	375	92	1005	3221	993	0.31

УПРАЖНЕНИЕ 4.3. Экспериментальное исследование имитационной модели популяции.

Задание. Исследовать динамику численности модельной популяции. Меняя значения входных параметров модели получить различные типы динамики численности популяции с сеткой: вымирание, периодические колебания с разной амплитудой, затухающие колебания с переходом в стационарный режим.

Порядок работы.

1. Проанализировать текст программы. Найти блоки ввода исходных данных, обработки информации, вывода результатов.
2. Запустить программу на выполнение. Задать значения входных параметров. Скопировать результаты эксперимента из текстового файла "Tabl.txt" в электронные таблицы.
3. Построить линейные графики динамики численности всех возрастных групп и популяции в целом.
4. Построить столбчатую нормированную диаграмму динамики возрастной структуры популяции.

5. Сделать выводы об адекватности реакции модели на возмущающие воздействия.

4.4. Моделирование сообществ.

Вычленение из экосистемы для анализа отдельных популяций и изучение их путем моделирования методически вполне оправдано. В свое время такой подход позволил сильно продвинуть вперед наши представления о динамической стороне продукционных процессов в экосистемах. Не потерял он актуальности и сейчас. Однако некоторая однобокость такого подхода уже стала вполне очевидной. Суть экосистемного образа мышления состоит в признании функциональной зависимости популяций друг от друга. Эту зависимость и ее влияние на каждую отдельно взятую популяцию невозможно игнорировать.

Закономерности замещения одних видов другими в процессе сукцессии экосистемы невозможно понять, не рассматривая сопряженной динамики популяций в сообществе. Это обстоятельство предопределило возникновение ценотического подхода в изучении динамических процессов, происходящих в водных и наземных экосистемах. Суть его заключается в том, что экосистема рассматривается не просто как взаимодействие множества отдельных популяций, а как взаимодействие подсистем - сообществ, внутри которых популяции связаны наиболее тесно.

Моделирование двухвидовых сообществ (систем "хищник-жертва").

Аналитические модели парных взаимодействий популяций берут свое начало от уравнений Лотки и Вольтерры, предложенных еще в 20-е годы прошлого столетия. Поэтому, естественно, что именно эти уравнения легли в основу предлагаемой модели. Если N_1 - численность жертв и N_2 - численность хищников, то соотношение между ними имеет вид:

$$\frac{dN_1}{dt} = r_1 N_1 - c N_1 N_2 \quad \frac{dN_2}{dt} = k c N_1 N_2 - m_2 N_2$$

где r_1 , c , k , m_2 - коэффициенты.

Биологическую сущность первого уравнения надо понимать следующим образом. Изменение численности жертв складывается из размножения, пропорционального

числу особей, и выедания, пропорционального произведению численностей жертв и хищников. Прирост численности хищников (второе уравнение) полагается пропорциональным числу съеденных жертв, причем учитывается естественная смертность хищников m_2 .

Программа реализует систему дифференциальных уравнений при помощи замены производных конечными разностями. Сделано это самым грубым методом, т.к. точное решение системы нас не интересует. По существу вся модель располагается в строках 170-190 программы, все остальное - это организация ввода исходных данных, установление начальных значений и представление результатов.

```

10  DIM N1(200), N2(200), D$(20)
20  R1 = .1: K = .02: M2 = .01: C = 100
30  PRINT "Скорость роста жертв "; R1
40  INPUT "Изменить (Y/N)"; Y$
50  IF Y$ = "Y" THEN INPUT R1
60  PRINT "Интенсивность хищничества "; K
70  INPUT "Изменить (Y/N)"; Y$
80  IF Y$ = "Y" THEN INPUT K
90  PRINT "Смертность хищников "; M2
100 INPUT "Изменить (Y/N)"; Y$
110 IF Y$ = "Y" THEN INPUT M2
120 IF K > 0 THEN N1(0) = 1.2 * M2 / K ELSE N1 = 100
130 N2(0) = R1 / (K * C)
140 MIN1 = 10000: MIN2 = 10000
150 MAX1 = 0: MAX2 = 0
160 FOR I = 0 TO 199
170  CON = K * N1(I) * N2(I):
180  N1(I + 1) = N1(I) * (1 + R1) - C * CON
190  N2(I + 1) = N2(I) * (1 - M2) + CON
200  IF N1(I + 1) < .0001 THEN N1(I + 1) = .0001
210  IF N2(I + 1) < .0001 THEN N2(I + 1) = .0001
220  IF N1(I + 1) > MAX1 THEN MAX1 = N1(I + 1)
230  IF N2(I + 1) > MAX2 THEN MAX2 = N2(I + 1)
240  IF N1(I + 1) < MIN1 THEN MIN1 = N1(I + 1)

```

```

250 IF N2(I + 1) < MIN2 THEN MIN2 = N2(I + 1)
260 NEXT I
270 FOR J = 1 TO 20
280 D$(J) = STRING$(60, ".")
290 NEXT J
300 FOR I = 1 TO 200 STEP 4
310 S1 = 1 + INT(19 * (N1(I) - MIN1) / (MAX1 - MIN1))
320 F = 1 + INT(I / 4)
330 MID$(D$(S1), F, 1) = "X"
340 S2 = 1 + INT(19 * (N2(I) - MIN2) / (MAX2 - MIN2))
350 MID$(D$(S2), F, 1) = "*"
360 NEXT I
370 FOR J = 20 TO 1 STEP -1
380 K$ = " |": K1$ = "| "
390 IF J = 19 THEN K$ = LEFT$(STR$(MAX1), 6)
400 IF J = 19 THEN K1$ = LEFT$(STR$(MAX2), 6)
410 IF J = 20 THEN K$ = " N1|"
420 IF J = 20 THEN K1$ = "| N2 "
430 PRINT K$ + D$(J) + K1$
440 NEXT J
450 INPUT "PHASE DIAG (Y/N)"; L$
460 FOR J = 1 TO 20
470 D$(J) = STRING$(40, ".")
480 NEXT J
490 FOR I = 1 TO 200 STEP 4
500 F = 1 + INT(I / 4)
510 S1 = 1 + INT(39 * (N1(I) - MIN1) / (MAX1 - MIN1))
520 S2 = 1 + INT(19 * (N2(I) - MIN2) / (MAX2 - MIN2))
530 F$ = STR$(F)
540 IF LEN(F$) < 2 THEN F$ = F$ + "."
550 MID$(D$(S2), S1, 3) = F$
560 NEXT I
570 FOR J = 20 TO 1 STEP -1
580 IF J = 20 THEN K$ = "N2" ELSE K$ = "| "

```



```

590 PRINT K$ + D$(J)
600 NEXT J
610 PRINT "_____ "
620 END

```

При первом знакомстве с программой рекомендуется не изменять ее параметров, а использовать тестовый вариант. Выдача результатов осуществляется в виде линейного графика сопряженной динамики численности двух популяций и фазовой диаграммы (рис. 21). Динамику численности можно представить и в системе Statistica (рис. 22).

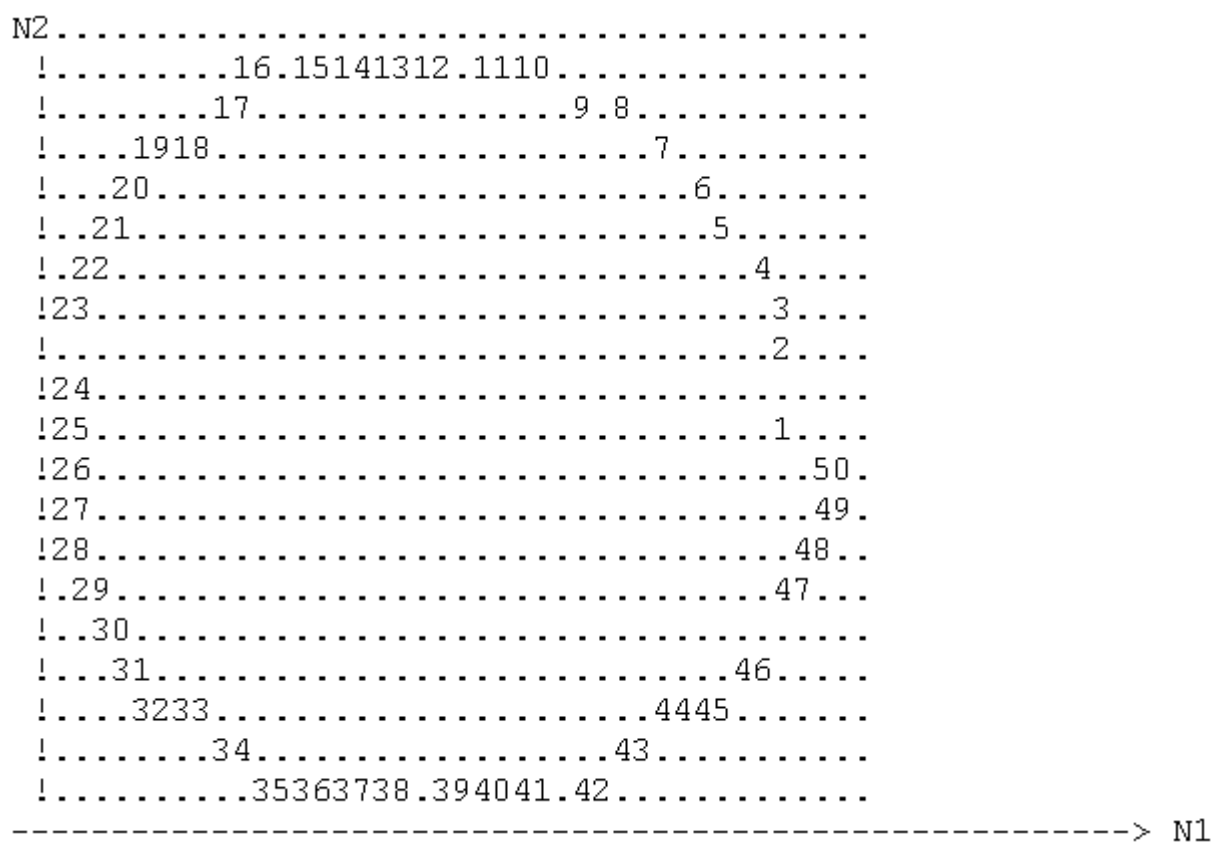


Рис. 21. Фазовая диаграмма динамики системы "хищник-жертва".

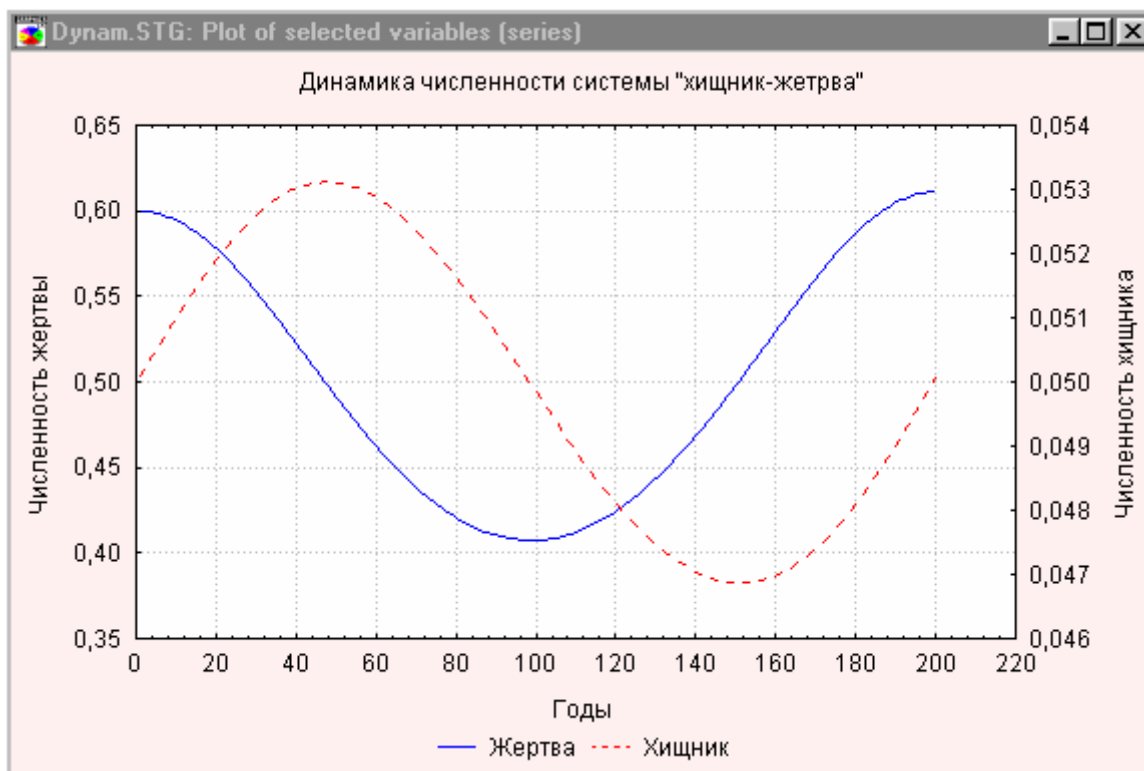


Рис. 22. Динамика численности популяций жертвы и хищника.

УПРАЖНЕНИЕ 4.4. Экспериментальное исследование имитационной модели системы "хищник-жертва".

Задание. Исследовать динамику численности системы двух взаимодействующих популяций. Меняя значения входных параметров модели получить различные варианты сопряженной динамики численности.

Порядок работы.

1. Проанализировать текст программы. Найти блоки ввода исходных данных, обработки информации, вывода результатов.
2. Запустить программу на выполнение. Задать значения входных параметров.
3. Построить линейные графики динамики численности популяций хищника и жертвы.
4. Сделать выводы об адекватности реакции модели на возмущающие воздействия.

Модель многовидового сообщества.

К многовидовым относятся сообщества с количеством видов больше двух. В настоящем разделе рассматривается очень упрощенная модель многовидового сообщества на примере озерного сообщества рыб. В ней не учитывается возрастная структура входящих в сообщество популяций рыб. Каждая популяция относится целиком к одному трофическому уровню. Такое приближение достаточно грубо, но вполне пригодно для практического знакомства с моделированием сообществ.

Каждая популяция, входящая в сообщество, характеризуется величиной биомассы составляющих ее рыб (B). Поступление энергии в популяцию осуществляется в виде суммарного рациона (C), который, как правило, складывается из нескольких частных рационов (C_1), соответствующих различным источникам питания.

В качестве внешних источников энергии для сообщества рассматриваются продукции зоопланктона и зообентоса, а также запасы мертвого органического вещества, поскольку в состав ихтиоценоза входит такой детритофаг, как плотва.

Распределение пищевых потребностей по кормовым объектам предполагается пропорционально биомассам жертв с учетом особенностей спектров питания особей каждой популяции (K). Степень выедания кормового объекта (M_2) связывается с напряженностью пищевых отношений (трофэкологическим коэффициентом) соотношением В.С.Ивлева:

$$M_2 = 1 - \text{EXP}(-K_4 * K_3)$$

где K_4 - коэффициент пропорциональности,

K_3 - отношение потребностей в данном виде корма к их наличному количеству.

Рассматриваемая модель из-за отсутствия в ней возрастных структур популяций не пригодна для имитации динамических процессов, поэтому приходится ограничиваться только стационарными состояниями, при которых приход энергии в популяцию равен энергетическим расходам. Иными словами полагается, что весь прирост биомассы расходуется на выедание, отмирание и изъятие промыслом. Модель по своему существу представляет собой сложную систему нелинейных уравнений, которая решается методом итераций.

```

10 REM Модель многовидового сообщества рыб (файл FishComm.bas)
20 DIM N$(12) 'Названия видов рыб и объектов внешней кормовой базы
30 DIM B(12) 'Биомассы популяций рыб и объектов внешней кормовой базы (кг/га)
40 DIM B0(12) 'Начальная биомасса популяций рыб и объектов
50 'внешней кормовой базы (кг/га)
60 DIM B1(9) 'Кормовая база популяций
70 DIM R1(9) 'Соотношение между максимальным рационом и биомассой
80 DIM M(9) 'Естественная смертность (начальная)
90 DIM M1(12) 'Естественная смертность (не зависящая от хищников) (9)
100 DIM M2(12) 'Степень выедания кормового объекта (смертность от хищников)
110 DIM C(9) 'Суммарный (реальный) рацион отдельный популяций рыб
120 DIM C1(9,12) 'Частные рационы популяций рыб по различным источникам питания
130 DIM V(12) 'Давление хищников
140 DIM K(9,12) 'Спектр питания особей каждой популяции (коэффициент
150 'избирательности питания)
160 DIM K2(9) 'Кормовой коэффициент
170 DIM K3(12) 'Трофозоологический коэффициент (отношение потребностей
180 'в данном виде корма к их наличному количеству)
190 DIM F(12) 'Степень изъятия промыслом рыбы данной популяции
200 DIM Y(12) 'Вылов рыбы данного вида
210 DATA "Язь", "Плотва", "Ерш", "Лещ", "Окунь", "Щука", "Налим"
220 DATA "Уклея", "Снеток", "Бентос", "Детрит", "Зоопланктон"
230 FOR I=1 TO 12 : READ N$(I) : NEXT I
240 NN$ = "Язь Плотва Ерш Лещ Окунь Щука Налим Уклея Снеток"
250 DATA 5, 50, 25, 30, 25, 10, 2, 5, 3, 300, 100, 50
260 FOR I=1 TO 12 : READ B0(I) : NEXT I
270 DATA 2.56, 3.48, 3.45, 3.0, 1.39, 1.54, 1.64, 3.31, 3.6
280 FOR I=1 TO 9 : READ R1(I) : NEXT I
290 DATA 4.5, 2.8, 3.0, 4.8, 3.2, 3.0, 2.8, 4.0, 4.0
300 FOR I=1 TO 9 : READ K2(I) : NEXT I
380 DATA 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.2, 0, 0
390 DATA 0, 0, 0, 0, 0, 0, 0, 0, 0, .54, 1.72, 0
400 DATA 0, 0, 0, 0, 0, 0, 0, 0, 0, .83, 0, 0
410 DATA 0, 0, 0, 0, 0, 0, 0, 0, 0, .84, 0, 0

```

```

420 DATA 3.8, 2.8, 2.8, 0, 0, 0, 0, 2.4, 2.9, .24, 0, 0
430 DATA 1.1, 1.17, .78, .6, .4, 0, 0, 1.1, .92, 0, 0, 0
440 DATA 0, .8, .75, 0, 0, 0, 0, 1.5, .8, 0, 0, 0
450 DATA 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, .8
460 DATA 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, .8
470 FOR I=1 TO 9 : FOR J=1 TO 12
480 READ K(I,J)
490 NEXT J : NEXT I
500 DATA .02, .01, .01, .01, .01, .01, .08, .01, .01
510 FOR I=1 TO 9 : READ M(I) : NEXT I
520 PRINT "Продукция бентоса";B0(10)*.1;" г/кв.м. Изменить? (Y/N)"
530 INPUT Y$
540 IF Y$<>"Y" AND Y$<>"y" THEN GOTO 560
550 INPUT B2 : B0(10)=B2*10 : B0(11)=.3*B0(10)
560 PRINT "Продукция зоопланктона";B0(12);"г/кв.м. Изменить? (Y/N)"
570 INPUT Y$
580 IF Y$<>"Y" AND Y$<>"y" THEN GOTO 600
590 INPUT P : B0(12)=P*10
600 FOR I=1 TO 12
610 B(I)=B0(I)
620 NEXT I
630 PRINT "Промысловое воздействие на отдельные популяции (0-1)"
640 FOR I=1 TO 9
650 PRINT N$(I)
660 INPUT F(I)
670 F(I)=F(I)/10
680 NEXT I
710 FOR T=1 TO 20 'Цикл по годам
720 FOR I=1 TO 9
730 B1(I)=0
740 FOR J=1 TO 12
750 IF K(I,J)>0 THEN B1(I)=B1(I)+B(J)
760 NEXT J
770 NEXT I

```

```

780  FOR I=1 TO 9
790    R=2*R1(I)
800    IF B(I)>0.01 THEN R=2*R1(I)*EXP(-.693*B0(I)/B(I))
810    FOR J=1 TO 12
820      C1(I,J)=0
830      IF B1(I)>0 THEN C1(I,J)=K(I,J)*B(J)*R*B(I)/B1(I)
840    NEXT J
850  NEXT I
860  FOR J=1 TO 12
870    V(J)=0
880    FOR I=1 TO 9
890      V(J)=V(J)+C1(I,J)
900    NEXT I
910    K3(J)=0
920    IF B(J)>0 THEN K3(J)=V(J)/B(J)
930    IF K3(J)>20 THEN K3(J)=20
940    IF K3(J)>0 THEN M2(J)=(1-EXP(-.2*K3(J)))
950  NEXT J
960  Y2=0
970  FOR I=1 TO 9
980    C(I)=0
990    FOR J=1 TO 12
1000   IF V(J)>0 THEN C(I)=C(I)+(C1(I,J)/V(J))*B(J)*M2(J)
1010  NEXT J
1020  M1(I)=M(I)*(1-EXP(-B(I)/B0(I)))
1030  B(I)=B(I)*(1-M1(I))*(1-M2(I))+C(I)/K2(I)
1040  Y(I)=B(I)*F(I)
1050  B(I)=B(I)-Y(I)
1060  Y2=Y2+Y(I)
1070  NEXT I
1080  IF T<>20 THEN GOTO 1260
1090  Z1$=" Вид    Ихтиомасса  Смертность Троф./экол Пром.  Вылов  "
1100  Z2$=" кг/га    коэфф.    усилие кг/га  "
1110  CLS: PRINT Z1$ : PRINT Z2$

```

```

1150 Z$="\      \ #####.## #.### ##.### #.## ###.##"
1160 FOR I=1 TO 12
1180   PRINT USING Z$;N$(I),B(I),M1(I),K3(I),F(I),Y(I)
1190   Z3$="-----"
1200 IF I=9 THEN PRINT Z3$
1230 NEXT I
1240 PRINT "Суммарный вылов =";Y2
1260 NEXT T
1270 PRINT : INPUT "Для продолжения нажмите Enter", H
1280 SCREEN 9 : COLOR 0,7 : CLS
1310 LINE (20,3)-(620,3),1
1320 LINE -(620,320),1
1330 LINE -(20,320),1
1340 LINE -(20,3),1
1350 LINE(20,256)-(620,256),1
1360 BMAX=0
1370 FOR I=1 TO 9
1380 IF B(I)>BMAX THEN BMAX=B(I)
1390 NEXT I
1400 D100=(100/BMAX)*250
1410 X=30 : Y=255
1420 FOR I=1 TO 9
1430 IF I<5 THEN A=2
1440 IF I>4 AND I<8 THEN A=4
1450 IF I>7 THEN A=1
1460 LINE (X,Y)-(X+20,Y),A
1470 D=(B(I)/BMAX)*250
1480 LINE -(X+20,Y-D),A
1490 LINE -(X,Y-D),A
1500 LINE -(X,Y),A
1510 PAINT (X+2,Y-D/2),A
1520 X=X+62
1530 NEXT I
1540 COLOR 8,7 : LOCATE 20,5 : PRINT NNS$

```

```

1560  COLOR 1,7 : LOCATE 2,60 : PRINT " - планктофаги"
1580  COLOR 2,7 : LOCATE 4,60 : PRINT " - бентофаги"
1600  COLOR 4,7 : LOCATE 6,60 : PRINT " - ихтиофаги"
1620  COLOR 1,7 : LOCATE 22,18: PRINT "Трофическая структура сообщества рыб"
1650  LOCATE 24,1 : INPUT "Для продолжения нажмите Enter", Enter : SCREEN 0
1670  END

```

УПРАЖНЕНИЕ 4.5. Экспериментальное исследование имитационной модели многовидового сообщества.

Задание. Исследовать изменение структуры сообщества.

Порядок работы.

1. Проанализировать текст программы. Найти блоки описания переменных, ввода исходных данных, обработки информации, вывода результатов в текстовой и графической форме.
2. На основе матрицы коэффициентов избирательности питания создать блок-схему трофической сети в MS Word.
3. Запустить программу на выполнение. Задать значения входных параметров.
4. Сделать выводы об адекватности реакции модели на возмущающие воздействия.

4.5. Моделирование экосистем

Построение модели экологической системы представляет собой средство для решения той или иной задачи. Вот некоторые из них: прогноз возможных состояний экосистемы при различных вариантах внешних воздействий как природных, так и антропогенных факторов; оптимизация управляющих воздействий для получения максимального полезного продукта (например, вылова рыбы) при сохранении состояния экосистемы в допустимых пределах; прогнозирование действия на экосистему загрязняющих веществ и определение допустимых норм выбросов, гарантирующих сохранение состояния системы в принятых границах; минимизация числа и объема наблюдений над состоянием экологической системы, обеспечивающих адекватное отображение ее состояния с заданной степенью точности (задача мониторинга). Примеры задач моделирования экосистем можно было бы продолжить, но уже приведенное перечисление показывает существенное разнообразие проблем охватываемых разбираемым методом.

Если к этому еще добавить богатство типов природных и искусственных экосистем (пруд, озеро, водохранилище, море, океан, лес и т.п.), то становится ясно, что ни о какой универсальной модели не может быть и речи. Можно говорить только об универсальных принципах построения подобных моделей.

На примере моделирования экосистемы удобно рассмотреть другие аспекты метода: определение границ, входов и выходов моделируемой системы, задание пространственной структуры и параметров модели. Ниже будет рассмотрена лишь специфика моделирования этих систем

Определение границ моделируемой экосистемы

После установления цели построения модели экологической системы встает вопрос о границах моделируемого объекта. В случае пруда, озера или внутреннего моря границы представляются достаточно естественными, но при моделировании экосистемы эстуария, залива, участка реки, открытого участка крупного водоема или лесного массива выбор границ экосистемы представляет большие трудности и от него зависит судьба всей дальнейшей работы. Общий принцип выбора границ моделируемого объекта заключается в том, чтобы через эти границы потоки вещества и энергии были бы минимальны и хорошо известны.

Установлении внешних связей изучаемой экосистемы

Следующий этап построения модели экологической системы заключается в установлении внешних связей изучаемой системы. Естественно, что выбор входов и выходов модели определяется целями моделирования. Целесообразно различать управляемые (т.е. поддающиеся воздействию человека) и неуправляемые входы экосистемы. Примерами управляемых входов, например водной экосистемы, могут служить интенсивность промысла рыбы, регулирование уровня, сброс загрязняющих веществ, вселение новых видов гидробионтов, удобрение, искусственная аэрация и т.п. К неуправляемым входам могут относиться атмосферные осадки, речной сток, солнечная радиация, теплообмен с атмосферой и т.д.

Выходы модели играют существенную роль в том случае, когда разрабатываемая модель экосистемы представляет собой составную часть более сложной модели, например экономического района или целой географической зоны. Примерами выходов

могут служить сток из озера, вылов рыбы или других объектов промысла, качество воды, влияющее на системы водоснабжения и эффективность рекреации.

Другой аспект выходных величин заключается в получении необходимой информации о функционировании модели. Для проверки модели выводятся обычно те величины, которые наиболее надежно измеряются в природе. В исследовательских целях, наоборот, выводятся величины, которые трудно или невозможно определить путем натурных измерений. Примерами выходов 1-го типа могут служить прозрачность воды или биомасса зоопланктона, 2-го типа - напряженность трофических отношений.

Пространственная структура модели

Основополагающей характеристикой модели экосистемы является ее пространственная структура и разбиение на ячейки. Если, исходя из условий задачи и специфики экосистема, можно не учитывать неоднородность ее характеристик по вертикали и горизонтали, то модель называется *точечной* и пространственная структура отсутствует. Примеры таких моделей достаточно многочисленны и относятся или к небольшим экосистемам, или к теоретически обобщенным задачам.

В общем случае пространственной неоднородностью экосистемы пренебрегать нельзя. В связи с этим вводится понятие *ячейки*, т.е. такого объема среды, в котором распределение всех характеристик экосистемы полагается равномерным и может быть заменено средними величинами. Разбиение всего моделируемого объекта на ячейки определяется не только особенностями конкретного объекта, но и формулировкой задачи моделирования. Так, при грубой оценке уровня эвтрофикации озера бывает достаточно точечной модели, но выяснение условий заморов в том же озере потребует выделения эпилимниона и гипolimниона или даже более подробного описания вертикальной стратификации. Взаимное расположение ячеек определяет пространственную структуру экологической системы. Процессы, которые происходят внутри ячейки, условимся называть процессами преобразования веществ, или просто *процессами преобразования*, а процессы, в которых участвуют соседние ячейки, - процессами обмена между ячейками, или просто *процессами обмена*. Процессы преобразования включают в себя такие явления, как фотосинтез, питание, рост, размножение, отмирание, дыхание, выделение организмов всех трофических уровней, химические реакции, растворение веществ и их выделение в твердом или газообразном состоянии и т.п. Процессы обмена представлены в водных экосистемах различными видами течений, турбулентной диффузией, ак-

тивными миграциями гидробионтов, опусканием или всплытием частиц; в наземных экосистемах - воздушным переносом, активными и пассивными миграциями организмов и др.

Параметры состояния

Состояние ячейки и всей экосистемы определяется набором компонент или переменных. В качестве компонент обычно выбираются биомассы различных групп организмов, концентрации растворенных и взвешенных веществ. Примерами могут служить концентрация хлорофилла, биомасса хищного зоопланктона, численность рыб данного вида и данного возраста, концентрация нитратов и растворенных органических веществ. Величины, определяющие состояние экосистемы, обычно являются неотрицательными действительными числами, хотя не исключено применение лингвистических переменных.

Временной шаг

Переход модели экологической системы из одного состояния в следующее осуществляется в дискретные моменты времени, расстояние между которыми называется временным шагом (Δ). Выбор величины временного шага является ответственным моментом при построении модели, так как все процессы, характерные времена которых оказываются меньше величины шага, выпадают из рассмотрения. Например, в модели с суточным шагом не могут быть учтены ночные скопления зоопланктона у поверхности воды или приливно-отливные явления. В пределах одного временного шага процессы преобразования предполагаются разнесенными по времени, т.е. происходят по очереди. Это создает предел скорости передачи возмущений по экосистеме, так как за один временной шаг возмущение может сказаться только на состоянии соседней ячейки. Предлагаемая схема функционирования модели экосистемы основана на независимом протекании процессов преобразования во всех ячейках, после чего происходят обменные процессы между ячейками, причем эта процедура повторяется на каждом временном шаге со сменой внешних воздействий (управляемых и неуправляемых входов), зависящих от времени.

```

REM  Модель экосистемы озера без стратификации
REM  ***** Описание параметров модели *****
DIM B(4)  'Биомасса фитопланктона, зоопланктона, рыбы и бактерий
DIM B1(4) 'Эффективная биомасса
DIM Q(4)  'Дыхание
DIM Q1(4)
DIM M(4)  'Смертность
DIM M1(4) 'Коэффициенты смертности
DIM K(4)  'Коэффициенты трофического давления на кормовые объекты
DIM E(4)  'Элиминация биомассы
DIM T0(24)'Температура по 15-дневным интервалам
DIM Z(24) 'Солнечная радиация
REM  ***** Исходные данные *****
DATA 0, 1.7, 0.9, 2 'Начальные биомассы в мг-С/кв.м
DATA 0.9, 0.05, 0.5, 0.3 'Дыхание
DATA 0.05, 1, 1.3, 0.1 'Смертность
DATA 2, 0 'Начальные фосфор и детрит
REM  Температура воды
REM  Янв Фев Мар Апр Май Июн Июл  Авг  Сен  Окт Ноя Дек
DATA 1,1, 1,1, 1,2, 2,3, 4,5, 6,8, 10,12, 15,16, 14,12, 10,8, 4,3, 2,1
REM  Солнечная радиация в мг-С/сутки*кв.м
DATA 0,0, 0,0, 0,0, 0,0, 4,5, 7,7, 5,4, 3,3, 2,2, 1,1, 0,0, 0,0
REM  ***** Начальное состояние *****
FOR I=1 TO 4
    READ B(I)
    READ Q1(I)
    READ M1(I)
NEXT I
READ N 'Фосфор
READ D 'Детрит
INPUT "Сток фосфора в мг/год/кв.м = ", N1
INPUT "Интенсивность промысла рыбы = ", F
E1=0.05 'Соотношение фосфора и углерода в биомассе
REM  *****

```

```

FOR I=1 TO 24 : READ T0(I) : NEXT I
FOR I=1 TO 24 : READ Z(I) : NEXT I
U=0.8 ' Коэффициент усвоения пищи
S=0.5 ' Коэффициент седиментации
Y1=0
P9=0
REM *****
OPEN "Output.txt" FOR OUTPUT AS #1 '
PRINT #1, "Время ";" Фитопл ";" Зоопл ";" Рыба ";" Бактерии ";" Детрит ";" Фосфор"
REM *****
FOR T=1 TO 24 'Цикл по временным интервалам (0.5 мес.)
  T1=T0(T) 'Текущая температура
  W1=EXP(20.5*(T1-20)/(273+T1)) 'Уравнение Аррениуса
  W2=EXP(-0.172*ABS(T1-15)) 'Температурный коэффициент потребления
  M9=0
  H9=0
  FOR I=1 TO 4
    Q(I)=Q1(I)*W1*B(I) 'Дыхание
    M(I)=M1(I)*B(I)
    H9=H9+E1*Q(I)
    M9=M9+M(I)
    K(I)=0
  NEXT I
  R2=5*W2*B(2)*EXP(-0.11*B(2)) 'Уровень максимального потребление пищи
  R3=W2*B(3)*EXP(-0.4*B(3)) 'Уровень максимального потребление пищи
  P1=15*Z(T)-Q(1) 'P1 - первичная продукция в мг C/15 суток/м²
  IF P1>N/E1 THEN P1=N/E1
  IF P1<0 THEN P1=0
  P4=0.5*D-Q(4)
  IF P4>10*W1 THEN P4=10*W1
  IF P4<0 THEN P4=0 'P4 - бактериальная продукция
  B1(1)=B(1)-M(1)+P1
  B1(2)=B(2)-M(2)-Q(2)
  B1(3)=B(3)-M(3)+Q(3)

```

```

B1(4)=B(4)-M(4)+P4
IF B1(1)+B1(4)>0 THEN K(1)=R2/(B1(1)+B1(4))
IF B1(2)>0 THEN K(2)=R3/B1(2)
IF B1(3)>0 THEN K(3)=F
K(4)=K(1)
C9=0
FOR I=1 TO 4
    E(I)=B1(I)*(1-EXP(-K(I))) 'Е -элиминация биомассы
    IF I<>3 THEN C9=C9+E(I)
NEXT I
REM Баланс вещества в экосистеме
B(1)=B1(1)-E(1) ' Фитопланктон
B(2)=B1(2)-E(2)+U*(E(1)+E(4)) ' Зоопланктон
B(3)=B1(3)-E(3)+U*E(2) ' Рыба
B(4)=B1(4)-E(4) ' Бактерии
D=D*(1-S)+(1-U)*C9+M9-P4 ' Детрит
N=N+H9-E1*P1+N1/24 ' Фосфор
P9=P9+P1
Y1=Y1+E(3)
PRINT #1, T; B(1); B(2); B(3); B(4); D; N
NEXT T
PRINT #1, "Первичная продукция", P9;"мг-С/год/кв.м"
PRINT #1, "Рыбная продукция", Y1;"мг-С/кв.м";Y1*20;"кг/га"
CLOSE
END

```

УПРАЖНЕНИЕ 4.6. Экспериментальное исследование имитационной модели водной экосистемы.

Задание. Исследовать внутригодовую динамику экосистемы.

Порядок работы.

1. Проанализировать текст программы. Найти блоки описания переменных, ввода исходных данных, обработки информации, вывода результатов в текстовой и графической форме.

2. На основе матрицы коэффициентов избирательности питания создать блок-схему экосистемы в MS Word.

3. Запустить программу на выполнение. Задать значения входных параметров. Повторить прогон программы многократно, меняя значения входных параметров. В электронных таблицах построить линейные графики внутригодовой динамики различных параметров экосистемы при различных начальных условиях.

4. Сделать выводы об адекватности реакции модели на изменение значений входных параметров.

Литература

- Боровиков В. Statistica: искусство анализа данных на компьютере. СПб: Питер, 2001. 656 с.
- Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1977. 479 с.
- Гроссман С., Тернер Дж. Математика для биологов. М.: Высшая школа, 1983. 383 с. Джермен М. Количественная биология в задачах и примерах. М.: Мир, 1972. 151 с.
- Джефферс Дж. Введение в системный анализ: применение в экологии. М.: Мир, 1981. 256 с.
- Жаков Л.А., Меншуткин В.В. Практические занятия по ихтиологии. Ярославль, 1982. 112 с.
- Жаков Л.А., Меншуткин В.В., Тихонов С.В. Применение учебных имитационных моделей в курсе "Основы рыбного хозяйства" // Современные проблемы биологии, химии, экологии и экологического образования. Ярославль, 2001. С. 375-377.
- Китаев С.П. Экологические основы биопродуктивности озер разных природных зон. М.: Наука, 1984. 207 с.
- Кулаичев А.П. Методы и средства анализа данных в среде Windows. STADIA 6.0. М.: Информатика и компьютеры, 1999. 340 с.
- Лакин Г.Ф. Биометрия. М.: Высшая школа, 1990. 352 с.
- Меншуткин В.В. Математическое моделирование популяций и сообществ водных животных. Л.: Наука, 1971, 196 с.
- Меншуткин В.В. Имитационное моделирование водных экологических систем. СПб.: Наука, 1993. 160 с.
- Неуймин Я.Г. Модели в науке и технике. Л.: Наука, 1984. 189 с.
- Теннант-Смит Дж. Бейсик для статистиков. М.: Мир, 1988. 208 с.
- Фролов Ю.П. Математические методы в биологии. ЭВМ и программирование. Куйбышев, 1987. 146 с.
- Терентьев П.В., Ростова Н.С. Практикум по биометрии. Л.: Изд-во Ленингр. ун-та, 1977. 152 с.

Приложение 1

Морфо-физиологические показатели детей младшего школьного возраста (фрагмент базы данных)

Пол	Возраст	Рост, см	Вес, кг	ОГК1	ОГК2	ОГК3	Сила правой руки, кг	Сила левой руки, кг
f	6	128,5	24	67	63	59	10	5
f	6	135	25,5	61	57,5	54	15,5	10
f	6	126	20,3	63	59	55	9	9
f	6	124	20,2	60	56	53	15	6
f	6	128,5	20,4	60	56	53	10	10
f	6	124,6	20,5	68	62	57	6	10
f	6	124	19,8	58	57	56	3	4
f	6	124	20	58	57	56	6	4
f	6	130	25,6	62	60	60	3	6
f	6	125	27,2	65	63	62	15	13
f	6	124	21,1	60	58	56	5	9
f	6	124	21,1	59	57,5	57	10	12
f	6	122	23,8	62	60	59	10	12
f	6	112	17	56	54	53	9	5
f	6	129	24	59	57	56	14	11
f	6	122,5	24,6	63	62	61	5	5
f	6	127,5	29	65	63	63	10	9
f	6	116	19,3	60	58	57	5	0
f	6	127	25,5	62,5	61	60	5	4
f	6	117	19	59	58	57	5	4
f	6	118	20,8	58	57	56	5	5
f	6	115	17	53	52	51	5	5
f	6	117,5	23,5	65	60	60	9	10
f	6	127,5	24,4	62	60	58	12	10
f	6	117	19,4	60	57	57	5	3
f	6	120	22,2	60	58	58	6	4
f	6	120,5	19,7	60	56	55	0	0
f	6	116	19,6	60	57	57	5	4
f	6	104	15,5	59	55	55	5	5
f	6	112,5	17,1	56,5	54	54	3	5
f	6	112,5	17,4	59	56	55	2	6
f	6	127,5	25,3	66	63	62,5	13	12
f	6	123,5	22,7	64	60	60	10	11
f	6	124	20,6	60	59	58	2	1
f	6	131	24,7	60	59	59	12	12
f	6	121	23,6	62	58	58	2	2
f	6	125	22,5	64	61	61	1	1
f	6	115	18,5	58	55	55	10	10
f	6	126	26,3	64	59	60	6	3
f	7	126	23	61	58	54,5	5	10

f	7	124	22,6	61,5	60	59	9	5
f	7	123	21	59	58	57	6	10
f	7	123	23,3	62	60	59	10	2
f	7	129	24	61	60	59	10	11
f	7	130,5	26,5	63	61	60	14	10
f	7	115	22,5	65	60	60	0	0
f	7	131	26,9	66	61	60,5	16	15
f	7	129	25,2	66	62	62	10	2
f	7	124,5	24	62	58	58	11	11
f	7	125	23	64	62	62	12	10
f	7	119	20,7	61	57	56	10	11
f	7	118	20	55	53	53	10	10
f	7	118	22	65	62	61,5	10	10
f	7	116,5	22,5	63	60	60	10	10
f	7	122	21	63	59,5	59	8	9
f	7	115	20	61,5	58	58	5	6
f	7	126	26,5	60	55	54	9	9
f	7	118	21,4	60	56	55,5	5	5
f	7	127,5	25,5	63	59	58,5	10	10
f	7	126,5	24,5	66	62,5	62,5	10	9
f	7	133,5	28	66,5	60	60	10	10
f	7	122	21	64,5	60	59	5	11
f	7	121	22,1	62	58	58	9	5
f	7	137	31,5	70	68	66	12	14
f	7	130	21	61	58	56	10	12
f	7	119	20,5	61	58	57	11	12
f	7	121	23,7	60	58	57	8	6
f	7	124	26,8	62	58	57,5	1	9
f	7	121	20	60	58	57	11	10
f	7	125	24,8	60	58	56	7	6
f	7	126	27,3	62	58	57	5	4
f	7	128	23,2	58	56	54	4	3
f	7	122	22,6	56	55	54	6	5
f	7	123	24,3	62	57	55	8	6
f	7	126	23	63	59	57	6	7
f	7	124	24,3	60	58	56	6	4
f	7	127	27,5	68	65	64	13	11
f	7	127	25,3	64	61	60	8	7
f	7	122	23,1	58	55	53	6	5
f	7	127	24	60	58	57	4	3
f	7	128	25,8	62	59	57	10	9
f	7	124	23,8	60	58	57	6	9
f	7	128	25	62	59	58	10	9
f	7	121	21	58	56	54	7	6
f	7	124	22,5	60	58	57	4	3
f	7	123	22	62	58	56	6	5
m	6	116	20,2	61	58	57	8	5

m	6	119	19,8	61	60	59	7	5
m	6	117	19,8	61	58	56	4	3
m	6	116	18,5	59	57	56	9	8
m	6	118	21	59	56	54	10	7
m	6	114	20	54	52	50	6	4
m	6	116	20	61	58	57	10	8
m	6	115	19,7	60	58	56	9	4
m	6	119	21,5	58	56	56	3	0
m	6	118	19,8	61	60	59	8	3
m	6	119	20	60	57	57	9	7
m	6	118	21	59	56	56	10	7
m	6	116	20	58	54	54	8	5
m	6	116	19	60	56	56	5	7
m	6	117	20,5	60	56	54	9	6
m	6	111,5	18	58	56	54	8	6
m	6	116	19,5	60	58	56	11	10
m	6	116	19,8	60	58	56	12	8
m	6	118	23	61	60	57	6	5
m	6	119	21,6	58	56	55	2	1
m	6	118	22	56	55	53	8	5
m	6	118	23	64	60	59	8	7
m	6	116	20,6	58	57	56	7	8
m	6	115	19	57	55	54	9	10
m	6	116	23	61	58	56	8	7
m	6	115	23	62	60	58	6	6
m	6	119	18,2	59	55	54	6	5
m	6	118,5	21	59	57	56	8	7
m	6	118	19,8	61	59	58	6	5
m	6	116	19	60	58	56	10	9
m	6	116	21	62	58	57	9	9
m	6	112	18	59	57	55	1	1
m	6	114	20	60	57	56	11	7
m	6	111	19,5	61	59	57	9,5	8
m	6	112	18	62	59	58	6	7
m	6	125	23,5	62	58	56	10	9
m	6	116	20	61	58	57	12	10
m	6	120	21,4	60	58	57	8	7,5
m	6	115	19,8	60	56	56	8	8,5
m	6	119	22	62	60	59	10	9
m	6	118	20	58	56	54	8	8
m	6	116	20	56	54	52	6	4
m	6	116	25	64	60	58	8	6
m	6	118	23	60	58	56	8	6
m	6	116	20	58	56	54	7	6
m	7	126	27,3	62	60	59	10	8
m	7	127	23,1	59	58	57	7	5
m	7	136	27,5	61	59	58	9	9

m	7	120	23	60	58	56	11	7
m	7	126	22,7	63	59	57	10	10
m	7	124	24,5	62	60	58	10	12
m	7	126	26	62	60	60	10	8
m	7	121	20	61	59	58	11	9
m	7	123	24	59	56	55	11	11
m	7	122	22	59	57	55	10	8
m	7	127	28	68	64	63	9	8
m	7	124	24,5	60	58	57	12	10
m	7	119	22	62	60	58	10	10
m	7	130	26	62	60	58	13	10
m	7	122	22	60	58	57	8	7
m	7	124	25	59	56	54	9	12
m	7	120	20,3	60	58	56	9	7
m	7	121	20	60	59	57	5	3
m	7	120	22,5	58	54	53	9	5
m	7	125	24	62	59	56	13	12
m	7	128	29,5	63	60	58	12	10
m	7	125	22	58	56	55	9	9
m	7	120	23,5	61	58	57	6	10
m	7	125	23	63	56	59	13	10
m	7	120	24,5	61	58	58	7	4
m	7	123	21,8	59	56	55	10	9
m	7	120	22	63	61	60	11	9
m	7	120	23,6	60	55	55	13	11
m	7	124	25	62	58	57	10	7
m	7	120	24,3	59	56	55	15	10
m	7	122	25	64	62	60	10	9
m	7	124	23	61	59	58	4	2
m	7	128	27	63	60	59	12	9
m	7	122	21,5	59	55	55	10	8
m	7	125	23	62	59	58	10	10
m	7	119	21,1	59	57	56	6	5
m	7	112	18,2	57	56	54	10	11
m	7	121	24	62	59	58	15	13
m	7	122	23,4	59	54	54	10	10
m	7	127	28	68	64	64	8,5	8
m	7	121	24,7	62	58	57	10	10
m	7	123	23,5	62	58	57	8	6
m	7	124	22,5	61	60	59	8	8
m	7	123	21	60	59	57	5	5
m	7	120	23,8	63	60	58	10	9
m	7	123	26,8	59	56	55	12	12,5
m	7	126	23,8	60	58	56	12	13
m	7	121	22,3	60	59	58	11	8
m	7	130	26	62	59	57	8	8

Приложение 2

Лимнологические параметры малых и средних озер Вологодской области (фрагмент базы данных).

Название озера	Площадь, га	Средняя глубина, м	Максимальная глубина, м	Прозрачность, м	Температура у дна в летний период	Содержание кислорода у дна	Минеральный фосфор	Зоопланктон, г/м ³	Бентос, кг/га
Карасьозеро	116	1,03	2	1,4	20,5	98	0,0051	4,07	1,3
Кедринское	520	1,3	2,25	1,2	19	100	0,005	1,41	2,16
Кобылье	210	1,17	2,4	1,2	20,8	89	0,0053	0,55	45,16
Панское	106	0,81	2	1,1	19	81	0,006	1,98	2,46
Жабинское	122	0,98	2,1	1,4	20,9	100	0,0046	5,8	7,62
Чагозеро	184,4	1,04	2,5	1,2	20,9	95	0,0048	0,681	2,28
Мегрское	4112	1,64	3,05	1,1	18,5	84	0,0068	1,289	2,25
Котечное	116	1,4	1,8	1,3	19	94	0,004	2,704	24,6
Водлицкое	186	0,88	2	1,1	20	90	0,0057	0,46	11,08
Вихозеро	416	1,61	2,1	1	21	79	0,0038	1,519	1,14
Великое	1094	1,8	2,2	0,9	19	90	0,0032	1,432	2,83
Лужандозеро	347,9	2,55	4,3	2,2	21	105	0,0022	0,372	0,24
Надречозеро	95	2,27	11	0,9	13	83	0,008	17,9	188,13
Качозеро	545,8	1,8	3	0,9	20	74	0,005	4,502	0,02
Онашкозеро	52	0,63	2,2	1,1	16,8	75	0,0035	5,14	9,14
Кудомозеро	68,7	1,94	8,8	1,2	11	20	0,0046	5,274	13,55
Ивачевское	58,7	2,47	5	1,5	20	72	0,0042	3,565	1,55
Сохитское	75,5	1,9	7,1	1,1	16,9	88	0,0036	2,952	13,95
Ежозеро	258	2,27	9,5	1,1	11	92	0,0057	2,98	4,78
Тудозеро	1226	3,03	4,8	1,2	21	89	0,0038	2,882	7,6
Дехозеро	12,1	1,22	1,7	0,6	22	62	0,001	5,138	1,32
Павшозеро	9	1,91	2,15	1	19,7	95	0,005	2,58	10,9
Ундозеро	235,2	2,33	3,1	1	19	89	0,0058	1,05	0,28
Лухтозеро	495,4	1,9	14	0,9	16	56	0,0019	6,346	0,23
Куштозеро	1707,5	2,8	11	1,8	19	83	0,0046	0,993	6
Струкозеро	58,2	1,2	4,5	0,8	18,7	71	0,0027	0,68	3,76
Палозеро	32,4	6,64	15	1,7	18	46	0,003	1,682	23,94
Долгозеро	141,2	3,2	12	1,5	16	51	0,0035	1,119	0,28
Шимозеро	780,6	2,7	21,5	2,1	16	91	0,005	0,442	2,81
Маслозеро	218,8	6,7	20	2,7	16	85	0,0032	0,289	0
Яндозеро	367,5	3,1	8,2	2,8	15,5	22	0,0036	0,757	0,82
Сюргозеро	3,3	10,8	21	1,6	11	13	0,0036	0,778	2,041
Югозеро	78,2	1,3	5,5	1	10	22	0,0022	3,084	2,4
Айнозеро	453	1,1	4,3	1	19,4	80	0,0024	0,514	0,12

Габозеро	88	3,15	5,05	2,6	19,5	92	0,0075	0,43	16,32
Тумбаж	274,2	0,9	1,4	0,8	16,5	82	0,0048	0,662	0,8
Щучье	71	1,6	2,05	1,4	20,5	91	0,0011	3,12	17,32
Серово	45	1,88	2,35	1,6	20,7	92	0,0075	4,89	164,53
Лопозеро	63,4	1,2	2,8	1	19,6	88	0,004	9,845	56,78
Толинское	87,5	0,8	2,7	1,4	19,6	116	0,0022	2,303	25,6
Глубокое	38	1,44	2,1	1,1	20	77	0,0015	2,76	27,97
Пажемское	139,5	3,5	6,2	1,1	19,2	80	0,003	2,58	16,28
Кемское	1451	0,9	1,7	1	20,4	76	0,0024	1,027	5,34
Кукозеро	155,3	1,25	1,6	1	20,7	106	0,002	4,84	29,9
Окштомское	35,5	2,8	7,5	1,2	15,2	56	0,0037	3,563	4
Пичемское	62	1,78	2,5	1,5	19,6	79	0,003	2,21	8,04
Ньюкозеро	94	1,41	2,8	1	19,5	81	0,0014	4,88	27,05
Сарозеро	66,5	1,9	4,5	0,9	17	41	0,003	2,616	3,73
Головное	24,5	1,9	3,8	0,9	18	49	0,0033	1,883	223,72
Ковжское	6248	5,9	15	2,1	19,3	96	0,0024	0,256	0,82
Лунча	21,5	1,4	2,7	1	21	86	0,0032	5,374	6,07
Маленькое	12	1,35	1,8	1,1	19,7	79	0,0015	4,8	0,034
Павшинское	163,9	4,6	14,5	1,6	13,2	48	0,0048	2,287	1,29
Сеза	92	4,31	10	1,2	19,6	82	0,002	0,88	18,28
Корбозеро	224	3,3	7,5	1,4	17,9	55	0,006	1,93	1,26
Поповское	72	2,8	3,5	3,5	19,6	117	0,005	1,32	507,89
Хибальское	76,3	1,4	2,1	1,4	18,3	88	0,0032	0,683	6,11
Калозеро	6	1,84	2	1,2	18,4	84	0,004	2,77	53,96
Грязнозеро	112,5	1	2	1,8	21	97	0,0041	1,831	0,085
Кодозеро	101,1	1,2	1,7	1,2	16	85	0,0032	2,465	0
Вехкозеро	35,2	2,1	7,5	2,5	18,8	97	0,0023	3,35	1,41
Веркозеро	53	0,73	1,2	1	18,5	95	0,0029	0,234	0,16
Пяжозеро	1173,6	1,9	4,5	1,8	17,5	91	0,0029	1,181	0,31
Белое	12	3	8,2	0,9	18	76	0,0025	1,723	0,087
Кленозеро	73,4	3,9	7	2,5	17,5	95	0,005	0,085	0
Линдозеро	843,3	1,7	2,7	1,2	16,2	91	0,002	0,833	1,09
Маслозеро	180	2,3	5,1	1,3	18,4	103	0,0026	1,109	0,41
Мороцкое	648	1,3	2,1	0,5	23,2	98	0,0175	4,765	4,4
Искрицкое	668	1,16	2,5	0,5	21,4	88	0,127	2,921	0,49
Верхнее	57,8	0,6	0,8	0,8	19,8	89	0,0064	0,882	0,35
Слицовское	32	2,39	3,9	2,6	23	93	0,002	2,345	7,6
Яндровское	65	1,36	2,05	0,9	22,1	87	0,0125	3,294	0

Приложение 3

АНКЕТА

**Применение метода моделирования в биологических дисциплинах
на 1-3 курсах факультета биологии и экологии ЯрГУ им. П.Г. Демидова**

Предметы	Модели							
	Материальные		Идеальные					
	Геометрически подобные	Субстратно подобные	Вербальные	Графические иконические	Графические условные	Графоанали- тические	Аналитические	Алгоритмиче- ские (имитационные)
Общая биология								
Ботаника								
Зоология								
Анатомия человека								
Матема- тические методы в биологии								
Цитоло- гия и гис- тология								
Физиоло- гия чело- века и животных								
Физиоло- гия рас- тений								
Микро- биология								
Генетика								
Экология								

Приложение 4

АНКЕТА

**Применение метода моделирования в экологических дисциплинах
на 1-3 курсах факультета биологии и экологии ЯрГУ им. П.Г. Демидова**

Предме- ты	Модели							
	Материальные		Идеальные					
	Геомет- рически подоб- ные	Суб- страт- но подоб- ные	Вербал ьные	Графи- ческие икони- ческие	Графи- ческие услов- ные	Графоана- литические	Анали- тические	Алгорит- мические (имитаци- онные)
Общая биология								
Геогра- фия								
Почво- ведение								
Учение о биосфе- ре								
Ланд- шафто- ведение								
Общая экология								
Матема- тические методы в экологии								
Основы приро- дополь- зования								
Генетика								

Содержание

Введение

1. Программное обеспечение анализа данных
 - 1.1. Электронные таблицы MS Excel
 - 1.2. Статистический пакет Statistica
 - 1.3. Язык программирования Бейсик.
2. Подготовка первичных данных для обработки
 - 2.1. Характер биологических и экологических данных
 - 2.2. Источники данных
 - 2.3. Ввод данных
3. Статистический анализ
 - 3.1. Предварительная обработка выборочных данных
 - 3.2. Основные характеристики выборочной совокупности
 - 3.3. Критерии достоверности статистических оценок
 - 3.4. Дисперсионный анализ
 - 3.5. Корреляционный анализ
 - 3.6. Регрессионный анализ
4. Математическое моделирование биологических и экологических систем
 - 4.1. Моделирование как метод исследования сложных систем
 - 4.2. Классификация моделей
 - 4.3. Моделирование популяций
 - 4.4. Моделирование сообществ
 - 4.5. Моделирование экосистем

Литература

Приложение