



Основная профессиональная образовательная программа  
06.04.01 - Биология  
(Фундаментальная биология)

---

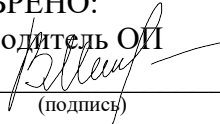
Министерство науки и высшего образования Российской Федерации

**ИВАНОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

Кафедра биологии

ОДОБРЕНО:

Руководитель ОП

  
(подпись)

В.Н. Мельников

« 30 » августа 20 21 г.

**Рабочая программа дисциплины**

Компьютерные технологии  
анализа биологических данных

Уровень высшего образования:	магистратура
Квалификация выпускника:	магистр
Направление подготовки:	06.04.01 Биология
Направленность (профиль) образовательной программы:	Фундаментальная биология
Тип образовательной программы:	программа академической магистратуры

Иваново



### 1. Цели освоения дисциплины

**Целью освоения дисциплины (модуля)** является формирование у студентов фундаментальных теоретических знаний по вопросам методики и практики применения методов многомерного статистического анализа биологических данных, а также обучение студентов современным программным средствам в которых реализованы модули, осуществляющие решение задач многомерного анализа. Курс позволит слушателям подготовиться к практическому использованию современных технологий анализа данных, оценить их значение для биолога.

Основными задачами курса выступают:

- изучение теоретических основ по спектру наиболее распространенных статистических методов анализа данных и условий их применения;
- изучение концепции и технологии современного анализа данных на компьютере, принципов работы программных средств, предназначенных для многомерного статистического анализа данных;
- выработка умения самостоятельного решения задач по выбору методов анализа в практических ситуациях;
- формирование навыков применения программных систем; предназначенных для многомерного статистического анализа данных, а также тестировании программных модулей на модельных данных.

### 2. Место дисциплины в структуре ОП

Курс «Компьютерные технологии анализа биологической информации» является важной составной частью в системе подготовки магистра биологии. Данная дисциплина относится к базовой части ОП по направлению подготовки 06.04.01, читается на 1 курсе, в 2 семестре обучения. Для изучения дисциплины необходимы знания и навыки, полученные в результате освоения дисциплин ОП подготовки 06.03.01 бакалавра биологии «Информатика, новые информационные технологии», «Математика и математические методы в биологии».

### 3. Планируемые результаты обучения по дисциплине

**3.1. Компетенции, формированию которых способствует дисциплина** (согласно матрице соответствия компетенций и составляющих ОП)

При освоении дисциплины формируются следующие компетенции в соответствии с ФГОС ВО по данному направлению подготовки:

а) универсальные (УК):

б) общепрофессиональные (ОПК):

ОПК-6 (способность творчески применять и модифицировать современные компьютерные технологии, работать с профессиональными базами данных, профессионально оформлять и представлять результаты новых разработок);

ОПК-8 (способность использовать современную исследовательскую аппаратуру и вычислительную технику для решения инновационных задач в профессиональной деятельности).

в) профессиональные (ПК):

**3.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с формируемыми компетенциями**

В результате освоения дисциплины обучающийся должен:

**Знать:** принципы хранения, обработки и представления биологической информации; основные термины, понятия, принципы и методы биостатистики, роль статистического анализа в современном исследовательском процессе (ОПК-6); основные методы многомерного



Основная профессиональная образовательная программа  
06.04.01 - Биология  
(Фундаментальная биология)

статистического анализа данных и особенности их применения в различных областях биологии (ОПК-8).

*Уметь:* использовать программные средства общего и специального назначения, уметь применять пакеты прикладных статистических программ для решения научно-исследовательских задач; использовать в познавательной и профессиональной деятельности базовые знания в области компьютерных технологий (ОПК-6); использовать полученные теоретические и практические навыки для организации научно-исследовательской и преподавательской деятельности, формулировать биологические задачи в виде, удобном для их решения с применением математических методов и современной компьютерной техники (ОПК-8).

*Иметь:* опыт самостоятельного применения методов анализа результатов биологического исследования с использованием компьютерных пакетов статистической обработки данных (ОПК-8).

#### 4. Объем и содержание дисциплины

Объем дисциплины составляет 3 зачетные единицы (108 часов).

##### 4.1. Содержание дисциплины по разделам (темам), соотнесенное с видами и трудоемкостью учебных занятий

№ п/п	Разделы (темы) дисциплины	Семестр	Виды учебных занятий, их объем (в часах при очной/заочной формах обучения)		Формы текущего контроля успеваемости (при очной/заочной формах обучения). Формы промежуточной аттестации (с указанием объема контактной работы в часах при очной/заочной формах обучения)
			Занятия лекционного типа	Занятия семинарского типа	
1.	Вводный. Введение в проблематику дисциплины, представление рабочей программы, осмысление требований к организации процесса обучения, самостоятельной работы и форм аттестации	2	-	2 семинар	Список вопросов, интересующих студента по содержанию дисциплины (сдается в письменном виде)
2.	Содержание и назначение прикладного многомерного статистического анализа.	2	-	2 семинар	Выступления на семинаре
3.	Кластерный анализ.	2	-	4 практ. занятие	Опорный конспект Отчет
4.	Дискриминантный анализ.	2	-	4 практ. занятие	Опорный конспект Отчет
5.	Компонентный анализ.	2	-	4 практ. занятие	Опорный конспект Отчет
6.	Факторный анализ.	2	-	4 практ. занятие	Опорный конспект Отчет
7.	Многомерное шкалирование.	2	-	4 практ. занятие	Опорный конспект Отчет
8.	Заключительный. Подведение и анализ промежуточных результатов освоения дисциплины	2	-	1 практ. занятие	
ИТОГО за семестр:			-	25	Зачет



#### **4.2. Развернутое описание содержания дисциплины по разделам (темам)**

Тема 1. СОДЕРЖАНИЕ И НАЗНАЧЕНИЕ ПРИКЛАДНОГО МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА. Сущность и назначение многомерного статистического анализа (МСА). Три базовые проблемы прикладного МСА: статистическое исследование зависимостей, классификация объектов (признаков) и снижение размерности исследуемого признакового пространства. Основные типы задач, решаемых с привлечением аппарата МСА: разведочный анализ и визуализация (формирование рабочих гипотез о природе анализируемых данных); сжатие больших массивов информации; замена многокритериальной оптимизационной схемы ее однокритериальной аппроксимацией; прогнозирование, нормирование; отбор наиболее информативных переменных. Программное обеспечение многомерного статистического анализа данных.

Тема 2. СПОСОБЫ ОПИСАНИЯ ПОВЕДЕНИЯ МНОГОМЕРНОГО ПРИЗНАКА. Понятия многомерного признака и многомерного наблюдения. Смешанная природа компонент многомерного признака. Определение количественных и категоризованных (ординальных и номинальных) переменных. Основные формы записи массива исходных данных в многомерном статистическом анализе (МСА).

Тема 3. КЛАСТЕРНЫЙ АНАЛИЗ. Задача многомерной классификации объектов исследования. Классификация без обучения. Виды кластерного анализа. Меры расстояния и меры сходства. Иерархические кластерные процедуры. Метод К-средних. Функционалы качества разбиения на классы. Зависимость выбора метода классификации от цели исследования. Классификация объектов (субъектов) в биологических исследованиях).

Тема 4. ДИСКРИМИНАНТНЫЙ АНАЛИЗ. Постановка задачи дискриминантного анализа. Метод дискриминантного анализа. Дискриминантная функция. Построение дискриминантной функции. Интерпретация дискриминантной функции. Коэффициенты дискриминантной функции. Оценивание вероятности ошибочной классификации при использовании построенной дискриминантной функции.

Тема 5. КОМПОНЕНТНЫЙ АНАЛИЗ (Модель, математическое обоснование и алгоритм метода главных компонент. Собственные векторы и собственные значения корреляционной матрицы и их использование для получения матрицы весовых коэффициентов. Получение и использование матрицы индивидуальных значений главных компонент. Особенности регрессии на главные компоненты. Использование метода главных компонент для анализа биологических данных).

Тема 6. ФАКТОРНЫЙ АНАЛИЗ. Метод факторного анализа. Область применения и принципы факторного анализа. Факторное отображение и факторная структура. Метод главных факторов. Использование методов вращения. Проведение факторного анализа и интерпретация результатов факторного анализа.

Тема 7. МНОГОМЕРНОЕ ШКАЛИРОВАНИЕ. Основы метода многомерного шкалирования. Матрица расстояний, построение шкал. Интерпретация результатов.

#### **5. Образовательные технологии**

С целью повышения познавательной активности студентов на практических занятиях используются технологии смешанного обучения, интерактивные методы обучения: опрос, демонстрация учебных материалов (видеозаписей, протоколов наблюдения, протоколов обследования и т.п.), моделирование учебных ситуаций, групповые дискуссии, обсуждение результатов и опыта, полученных студентами в ходе выполнения практических заданий.



## **6. Учебно-методическое обеспечение самостоятельной работы обучающихся**

Важная составляющая курса – это практические задания, выполняя которые студенты должны применить полученные теоретические знания в модельной ситуации биологического исследования. Эти задания выполняются с использованием компьютерных методов обработки данных. Тема учебного проекта по дисциплине согласовывается с преподавателем либо назначается им. Методический материал по обеспечению самостоятельной работы студентов приводится в приложении 1 к РП.

## **7. Характеристика оценочных средств для текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине**

Оценочные средства составляются преподавателем самостоятельно при ежегодном обновлении банка средств. Количество вариантов зависит от числа обучающихся. Текущий (промежуточный) контроль успеваемости проводится в виде контрольных и тестовые работы по ключевым темам курса. Итоговый контроль в виде зачета, который выставляется студентам выполнившим программу лабораторного практикума и успешно выполнивших тестовые задания промежуточного контроля с учетом рейтинговых показателей. Типовые варианты тестовой работы находятся в приложении к РП в разделе «Фонд оценочных средств».

## **8. Учебно-методическое и информационное обеспечение дисциплины**

Основная учебная литература:

1. *Калаева Е. А., Артюхов В. Г., Калаев В. Н.* Теоретические основы и практическое применение математической статистики в биологических исследованиях и образовании: учебник [Электронный ресурс] / Воронеж:Издательский дом ВГУ., -284с. - 978-5-9273-2241-1 - URL: <http://biblioclub.ru/index.php?page=book&id=441590>
2. *Чубукова И. А.* Data Mining [Электронный ресурс] / Москва:Интернет-Университет Информационных Технологий, 2008. -383с. - 978-5-94774-819-2 URL:<http://biblioclub.ru/index.php?page=book&id=233055>
3. *Шорохова И. С., Кисляк Н. В., Мариев О. С.* Статистические методы анализа: учебное пособие [Электронный ресурс] / Москва:ФЛИНТА|УрФУ,2017. -301с. - 978-5-9765-3279-3. - ISBN 978-5-7996-1633-5 (Изд-во Урал. ун-та) - URL: <http://biblioclub.ru/index.php?page=book&id=482354>

Дополнительная учебная литература

1. *Астафьев Н. В., Михалев В. И., Безмельницын Н. Г.* Математико-статистический анализ количественных данных физкультурно-педагогических исследований средствами Microsoft Excel: учебное пособие [Электронный ресурс] / Омск:Издательство СибГУФК,2004. -59с. - URL: <http://biblioclub.ru/index.php?page=book&id=274514>
2. *Бельчик, Т.А.* Основы математической обработки информации с помощью SPSS : учебное пособие / Т.А. Бельчик. - Кемерово : Кемеровский государственный университет, 2013. - 232 с. - ISBN 978-5-8353-1265-8 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=232214> (Стр. 168-187).
3. *Крыштановский, А.О.* Анализ социологических данных с помощью пакета SPSS : учебное пособие / А.О. Крыштановский ; Национальный исследовательский университет – Высшая школа экономики. - 2-е изд. - Москва : Издательский дом Высшей школы экономики, 2007. - 283 с. : ил. - (Учебники Высшей школы экономики). - Библиогр. в кн. - ISBN 978-5-7598-0486-4 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=445561> (Стр. 205 — 217).



Основная профессиональная образовательная программа  
06.04.01 - Биология  
(Фундаментальная биология)

4. Мхитарян, С.В. Применение SPSS в маркетинговых проектах : учебно-практическое пособие / С.В. Мхитарян. - Москва : Евразийский открытый институт, 2011. - 272 с. - ISBN 978-5-374-00315-4 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=90454> (Стр. 18-58; 76-94).
5. Остапенко, Р.И. Математические основы психологии : учебно-методическое пособие / Р.И. Остапенко. - Воронеж : ВГПУ, 2010. - 76 с. - ISBN 978-5-88519-680-2 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=120777> (Стр. 4-12).
6. Патронова, Н.Н. Статистические методы в психолого-педагогических исследованиях : учебное пособие / Н.Н. Патронова, М.В. Шабанова ; Министерство образования и науки Российской Федерации, Федеральное государственное автономное образовательное учреждение высшего профессионального образования Северный (Арктический) федеральный университет им. М.В. Ломоносова. - Архангельск : ИПЦ САФУ, 2013. - 203 с. : табл., граф., ил. - Библиогр. в кн. - ISBN 978-5-261-00847-7 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=436382> (Стр. 3-38).

Ресурсы информационно-телекоммуникационной сети «Интернет»:

1. Система электронной поддержки образовательного процесса «Мой университет» <https://uni.ivanovo.ac.ru>
2. Сайт компании Predictive Solutions (до 2012 г. – SPSS Russia) – (официальный поставщик программных продуктов IBM SPSS и сертифицированный провайдер технической поддержки продуктов IBM SPSS на территории России, СНГ и Прибалтики). - URL: <http://www.predictivesolutions.ru/>
3. Сайт компании StatSoft Russia (официальный представитель правообладателя программных продуктов серии Statistica компании TIBCO). - URL: <http://statsoft.ru/>
4. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>
5. StatSoft, Inc. (2012). Электронный учебник по статистике. Москва, StatSoft. URL: <http://www.statsoft.ru/home/textbook/default.htm> .

Профессиональные базы данных и информационно-справочные системы:

- ЭБС «Университетская библиотека онлайн» [www.biblioclub.ru](http://www.biblioclub.ru)
- Электронная библиотека ИвГУ <http://lib.ivanovo.ac.ru>
- Электронный каталог НБ ИвГУ <http://lib.ivanovo.ac.ru/index.php/ek>
- Сайт по методам машинного обучения - <http://archive.ics.uci.edu/ml/>
- Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных - <http://www.machinelearning.ru>

Программное обеспечение и информационно-справочные системы: Операционная система Linux, пакет офисных приложений OpenOffice.org, интернет-браузер Firefox.

## 9. Материально-техническое обеспечение дисциплины

Учебные аудитории:

- для проведения занятий семинарского типа, консультаций, текущего контроля и промежуточной аттестации с комплектом специализированной учебной мебели и техническими средствами обучения (Компьютерный класс).

Помещение для самостоятельной работы, оснащенное комплектом специализированной учебной мебели, компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в ЭИОС.



Основная профессиональная образовательная программа  
06.04.01 - Биология  
(Фундаментальная биология)

---

**Автор рабочей программы дисциплины:** к.б.н., доцент Лукьянов И.Ю.

Программа рассмотрена на заседании кафедры биологии

« \_\_\_\_ » \_\_\_\_\_ 20\_\_ г., протокол № \_\_\_\_

Программа обновлена

протокол заседания кафедры № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Согласовано:

Руководитель ОП \_\_\_\_\_ В.Н. Мельников  
(подпись)