

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт биологии, экологии, почвоведения, сельского и лесного хозяйства
(БИОЛОГИЧЕСКИЙ ИНСТИТУТ)

УТВЕРЖДАЮ:
Директор Биологического института

Д.С. Воробьев

« 29 » 06 20 23 г.

Рабочая программа дисциплины

Информационные технологии в естественных науках

по направлению подготовки

06.04.01 Биология

Направленность (профиль) подготовки / специализация:
«Фундаментальная и прикладная биология»

Форма обучения

Очная

Квалификация

Магистр

Год приема

2023

Код дисциплины в учебном плане: Б1.О.08

СОГЛАСОВАНО:

Руководитель ОП

Д.С. Воробьев

Председатель УМК

А.Л. Борисенко

Томск – 2023

1. Цель и планируемые результаты освоения дисциплины (модуля)

Целью освоения дисциплины является формирование следующих компетенций:

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

ОПК-8 – Способен использовать современную исследовательскую аппаратуру и вычислительную технику для решения инновационных задач в профессиональной деятельности.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-6.2 – Использует компьютерные технологии и профессиональные базы данных при планировании профессиональной деятельности, обосновывает их выбор;

ИОПК-6.3 – Профессионально оформляет и представляет результаты новых разработок;

ИОПК-8.1 – Демонстрирует понимание методических принципов полевых и лабораторных биологических исследований и типов используемой современной исследовательской аппаратуры;

ИОПК-8.2 – Применяет современную исследовательскую аппаратуру и вычислительную технику при решении стандартных и инновационных задач в профессиональной деятельности.

2. Задачи освоения дисциплины

– Освоение аппарата концепций, понятий, терминов, определяющих современное состояние и направления развития информационных технологий.

– Умение применять понятийный аппарат современных методов и практических приемов обработки естественнонаучных данных и представления результатов с использованием компьютерных технологий для решения практических задач профессиональной деятельности.

– Использование принципов доказательности в естественнонаучных исследованиях и объективизации содержательных выводов на основе статистического анализа данных.

– Получение навыков самостоятельной научной деятельности, связанной с выполнением законченного естественнонаучного исследования: от стадии планирования эксперимента, через получение, обработку и анализ данных, до этапа интерпретации и презентации результатов.

3. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина относится к обязательной части образовательной программы.

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Семестр 2, зачет.

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования.

Для успешного освоения дисциплины необходимо иметь достаточные знания в области математики, информатики и математических методов в биологии на уровне, соответствующем программам бакалавриата. Необходимым условием является успешное прохождение курса "Б1.О.03 Компьютерные технологии в естественных науках" из обязательной части программ по профилю подготовки «Фундаментальная и прикладная биология» (первый семестр первого года обучения).

6. Язык реализации

Русский

7. Объем дисциплины (модуля)

Общая трудоемкость дисциплины составляет 2 з.е., 72 часа, из которых:

– лекции: 4 ч.;

– практические занятия: 32 ч.;

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины (модуля), структурированное по темам

Тема 1. *Функциональная характеристика компьютерных программ и пакетов статистической обработки данных общего назначения. Обзор основных статистических программ. Программы StatSoft Statistica и PAST. Принципы внутренней организации и импорт данных в StatSoft Statistica и PAST. Графические возможности StatSoft Statistica и PAST.*

1. Функциональная характеристика программ статистической обработки данных общего назначения.
2. Обзор основных статистических программ общего назначения (Statsoft Statistica, SPSS, Systat, Statgraphics)
3. Использование аналитического пакета AtteStat в составе табличного процессора MS Excel
4. Типы документов в программе StatSoft Statistica: таблицы SpreadSheet и ScrollSheet, графики (STG), файлы протокола в формате RTF.
5. Принципы внутренней организации данных и особенности работы в Statistica. Рабочие книги.
6. Два варианта пользовательского интерфейса в StatSoft Statistica: классические меню и "ленточный" интерфейс.
7. Требования к подготовке данных и варианты импорта данных в Statsoft Statistica из других форматов.
8. Типы организации данных для различных видов статистического анализа. Особенности презентации исходных данных, результатов анализа и построения графиков.
9. Возможности Statistica по селекции данных.
10. Работа с признаками, описанными в нецифровой форме: альтернативные и номинальные данные.
11. Взвешивание признаков, использование критериев для исключения части наблюдений из анализа. Работа с пропущенными значениями. Приемы обработки пропущенных значений.
12. Программа PAST (Palaeontological Statistics). Интерфейс пользователя и организация наборов данных.

Особенности подготовки табличных данных для импорта в программу PAST

Тема 2. *Методы статистического анализа данных. Deskриптивные (описательные) статистики и таблицы частот, их вычисление в модуле Descriptive Statistics and Tables (StatSoft Statistica). Статистические сравнения в программе StatSoft Statistica. Основы дисперсионного анализа. Корреляция (изучение взаимной связи признаков), ранговые корреляции. Регрессия (изучение зависимости изменения одного признака от изменений другого). Основные регрессионные модели, применяемые в биологии и почвоведении. Общие вопросы классификации в естественных науках. Кластерный анализ и его место среди других методов анализа многомерных данных.*

1. Дескриптивные (описательные) статистики и таблицы частот, их вычисление в модуле Descriptive Statistics and Tables (в программе StatSoft Statistica).
2. Тестирование выборочных совокупностей на нормальность распределения данных: критерии Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk's
3. Особенности работы со взвешенными вариационными рядами.
4. Статистические гипотезы и их проверка в программе StatSoft Statistica
5. Параметрические и непараметрические критерии сравнения, реализованные в программе StatSoft Statistica.
6. Дисперсионный анализ. Схемы дисперсионного анализа и его реализация в программе StatSoft Statistica: ANOVA/MANOVA
7. Изучение взаимной связи признаков. Корреляции. Понятие о корреляционной связи. Корреляционная решетка.
8. Причины скоррелированности признаков, применяемых для описания объектов в естественных науках.
9. Коэффициент линейной корреляции Пирсона r . Коэффициенты детерминации и алиенации.
10. Нелинейная корреляция.
11. Ранговые корреляции. Область применения, характер данных.
12. Корреляционный анализ и его реализация в программе StatSoft Statistica.
13. Изучение зависимости изменения одного признака от изменений другого. Понятие регрессии.
14. Эмпирические линии регрессии. Регрессионный анализ.
15. Связь линейных регрессионных коэффициентов и коэффициента корреляции. Нелинейная регрессия.
16. Основные регрессионные модели: различия и условия применения. Прогностическая ценность модели.
17. Общие вопросы теории классификации в естественных науках.
18. Кластерный анализ и его место среди других методов анализа многомерных данных.
19. Алгоритмы классификации (древовидная иерархическая кластеризация, двухходовое объединение и метод К-средних) и их реализация в модуле Cluster Analysis в программе StatSoft Statistica
20. Кластерный анализ в программе PAST

Тема 3. *Современные информационно-коммуникационные технологии: локальные сети, глобальная сеть Интернет, протоколы передачи данных. Поиск научной информации в сети Интернет. Компьютерная безопасность. Охрана авторских прав на программное обеспечение и базы данных.*

1. Современные коммуникационные технологии: локальные сети, глобальная сеть Интернет, протоколы передачи данных.
2. Модем (типы принцип работы), модемное подключение. Беспроводные технологии подключения к сети. Мобильный Интернет.
3. Электронная почта (E-mail). Локальные и браузерные программы – почтовые клиенты
4. Получение несанкционированных информационных рассылок ("спам") и борьба с ним.
5. Использование протокола FTP.
6. Мировая паутина (WWW). Протоколы HTTP, TCP/IP. IP-адреса.
7. Язык разметки HTML. Расширения языка HTML.
8. Использование программ-браузеров (MS Internet Explorer, MS Edge, Google Chrome, Safari, Opera, Mozilla FireFox)

9. Информационные ресурсы сети Internet. Принципы и приемы поиска информации в Internet. Поисковые и мета-поисковые серверы.
10. Особенности поиска научной информации. Презентация своей информации в WWW.
11. Компьютерная безопасность. "Стандартные" (классические) компьютерные вирусы, "тройные кони", "почтовые" (HTML) вирусы, макро-вирусы (Word, Excel). Принципы распознавания и борьбы с ними.
12. Программы-антивирусы: сканирующие полифаги, резидентные программы.
13. Безопасность передачи информации в сети Internet и при пересылке электронной почты.

Тема 4. *Типы и форматы графической информации. Принципы обработки графической информации. Основы сканирования. Типы и принципы устройства сканирующих устройств. Оптическое распознавание текста. Основные приемы презентации результатов исследований. Комплексные документы на основе OLE-технологии в среде Windows.*

1. Векторная и растровая графика. Модели цветоделения.
2. Основные форматы растровой графики (BMP, PCX, TIFF, GIF, JPEG, PNG, Targa), способы сжатия графической информации.
3. Графические редакторы и графические процессоры.
4. Обзор наиболее популярных программ обработки графической информации (Corel Draw, Corel Painter, PhotoFinish, Paint, Adobe Illustrator).
5. Основы работы в Adobe Photoshop. Использование инструментальных панелей. Работа со слоями. Трансформация и коррекция изображений, применение стандартных и дополнительных фильтров.
6. Конвертация графических форматов.
7. Принципы устройства сканеров. Сканеры ручные, планшетные, рулонные, слайд-сканеры.
8. Сканирование непрозрачных графических изображений. Типы получаемых изображений, глубина цвета. Выбор разрешения сканирования, его связь с устройствами вывода на печать и экран. Сканирование для передачи факсимильных сообщений.
9. Сканирование в приложения: TWAIN-модули и программы, поддерживающие их.
10. Оптическое распознавание текста (OCR). Работа с программой FineReader.
11. Презентация результатов и подготовка отчетов. Графическое представление результатов анализа.
12. Включение графиков, диаграмм и изображений, созданных в других программах, в документы MS Word и MS Excel.
13. Использование программы MS PowerPoint для презентация своих данных.
14. Основы компьютерной верстки, краткая история и обзор издательских систем (Adobe PageMaker, MS Publisher, Corel Ventura, QuarkXpress). Преобразование информации из электронного представления в печатную форму.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, опросов по лекционному материалу, проверке выполнения домашних тестовых заданий, рефератам студентов (по необходимости) и фиксируется в форме контрольной точки не менее одного раза в семестр.

Успешность выполнения практических занятий в компьютерном классе по статистическому анализу наборов данных (учебные примеры) контролируется магистрантами самостоятельно на основе проверки результатов своих действий в программном обеспечении на индивидуальных рабочих местах по параллельной

демонстрации правильных действий преподавателем на мультимедийном экране. При необходимости выхода из затруднительных ситуаций и исправления ошибок, корректирующие действия выполняются в интерактивном режиме.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет во втором семестре проводится в виде стандартного теста или альтернативного творческого задания.

Стандартный тест предназначен для демонстрации основных практических навыков по работе с программным обеспечением для базовой статистической обработки данных. Тест состоит из трех частей:

1. импорт набора данных из текстового формата, вычисление основных дескриптивных характеристик исходных и трансформированных выборок;
2. проведение статистических сравнений с помощью параметрических и непараметрических критериев; подтверждение или опровержение результатов сравнения с помощью однофакторного дисперсионного анализа; построение и аннотация графиков распределения данных Box Plots (Box & Whiskers);
3. выполнение транспонирования матрицы данных, проведение агломеративного кластерного анализа исходных и стандартизированных данных с построением горизонтальной дендрограммы.

Каждая часть при необходимости уточнить качество знаний и уровень владения методами анализа сопровождается вопросами теоретического характера, ответы на которые требуют освоения лекционного материала и самостоятельной работы с учебными пособиями в течение семестра.

Время на выполнение теста – 60–80 минут. Результаты зачета определяются оценками «зачтено»/«не зачтено». Зачет проводится на зачетной неделе по расписанию.

Альтернативное задание носит творческий характер: перед магистрантом заранее (не менее 2-х недель) ставится задача создания логической компьютерной игры, требующая углубленных знаний в информационных технологиях и навыков владения программными средствами. Описываются выходные характеристики и параметры функциональности игры, однако не предлагаются конкретные компьютерные программы, алгоритмы и способы реализации. Задание рассчитано на активный самостоятельный поиск магистрантом решения с использованием всех доступных инструментов и информационных ресурсов. Оценка производится по неформализуемым критериям: полнота решения, корректность алгоритмов, механизмы обработки ошибочных игровых ситуаций, степень оригинальности наличие дополнительной функциональности.

Итоговая оценка промежуточной аттестации «зачтено» выставляется по совокупности оценки «зачтено» на тесте или при проверке альтернативного творческого задания и результатов посещаемости: при необходимости каждые три пропущенных занятия должны компенсироваться написанием оформленного по ГОСТУ реферата по пропущенному материалу.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» – <https://moodle.tsu.ru/course/view.php?id=18929>

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

в) Материалы для практических занятий по дисциплине: учебные наборы данных

г) Методические указания по организации самостоятельной работы студентов.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

- Гашев С.Н., Бетляева Ф.Х., Лупинос М.Ю.* Математические методы в биологии: анализ биологических данных в системе Statistica: Учебно-методическое пособие. Тюмень: Изд-во ТюмГУ, 2014. 208 с.
- Ивченко Г. И., Медведев Ю. И.* Введение в математическую статистику. М.: ЛКИ, 2015. 599 с.
- Миркин Б.Г.* Введение в анализ данных: учебник и практикум для бакалавриата и магистратуры. М.: Юрайт, 2014. 174 с.
- Песенко Ю.А.* Принципы и методы количественного анализа в фаунистических исследованиях. М.: Наука, 1982. 287 с.
- Шмидт В.М.* Математические методы в ботанике: Учебное пособие. Л.: Изд-во ЛГУ, 1984. 288 с.

б) дополнительная литература:

- Алексахин С.В. и др.* Прикладной статистический анализ данных. Теория. Компьютерная обработка. Области применения. В 2-х т. М.: ПРИОР, 2002. 688 с.
- Берк К., Кэйри П.* Анализ данных с помощью Microsoft Excel: Пер. с англ. М.: ИД "Вильямс", 2005. 560 с.
- Боровиков В.П.* STATISTICA: искусство анализа данных на компьютере. Для профессионалов. СПб.: Питер, 2001. 656 с.
- Гайдышев И.* Анализ и обработка данных: специальный справочник СПб: Питер, 2001. 752 с.
- Гланц С.* Медико-биологическая статистика. Пер. с англ. М.: Практика, 1998. 459 с.
- Дмитриев Е.А.* Математическая статистика в почвоведении: Учебник. М.: Либроком, 2009. 328 с.
- Дубров А.М., Мхитарян В.С., Трошин Л.И.* Многомерные статистические методы: Учебник. М.: Финансы и статистика, 2000. 352 с.
- Зайцев Г.Н.* Математический анализ биологических данных М.: Наука, 1991. 182 с.
- Ивантер Э.В., Коросов А.В.* Введение в количественную биологию. Учебное пособие. 2-е издание. Петрозаводск: Изд-во ПетрГУ, 2011. 304 с.
- Ким Дж.-О., Мьюллер Ч.У., Клекка У.Р., Олдендерфер М.С., Блэшфилд Р.К.* Факторный, дискриминантный и кластерный анализ. М.: Финансы и статистика, 1989. 215 с.
- Коросов А.В., Горбач В.В.* Компьютерная обработка биологических данных. Методическое пособие. Петрозаводск: Изд-во ПетрГУ, 2007. 76 с.
- Лакин С.Ф.* Биометрия: Учебное пособие для биол. вузов. М: Высш. школа, 1990. 352 с.
- Лялин В.С., Зверева И.Г., Никифорова Н.Г.* Статистика: теория и практика в Excel. М.: "Финансы и статистика", ИД "Инфра-М", 2010. 448 с.
- Мандель И.Д.* Кластерный анализ. М.: Финансы и статистика, 1988. 176 с.
- Ростова Н.С.* Корреляции: структура и изменчивость. СПб.: Изд-во С-Петербур. ун-та, 2002. 308 с.
- Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере. 3-е изд. / Под ред. В.Э.Фигурнова М.: ИНФРА-М, Финансы и статистика, 2003. 384 с.
- Урбах В.Ю.* Биометрические методы. Статистическая обработка опытных данных в биологии, сельском хозяйстве и медицине. М.: Наука, 1964. 416 с.
- Шитиков В.К., Зинченко Т.Д., Розенберг Г.С.* Макроэкология речных сообществ: концепции, методы, модели. Тольятти: Кассандра, 2011. 255 с.
- Dytham C.* Choosing and Using Statistics: A Biologist's Guide. Blackwell Science, 2003. 248 p.
- Sokal R.R., Rohlf F.J.* Biometry: the principles and practice of statistics in biological research. 3rd ed. NY, W.H. Freeman, 1995. 850 p.

в) ресурсы сети Интернет:

Биометрика: журнал для медиков и биологов, сторонников доказательной биомедицины [Электрон. ресурс]. URL: <http://www.biometrica.tomsk.ru/>

Ефимов В.М., Ковалева В.Ю. Многомерный анализ биологических данных : учебно-методический комплекс. Томск : ИДО ТГУ, 2008. [Электрон. ресурс] URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/vtls:000341281>

Additional Books on STATISTICA [Электрон. ресурс]. URL: <http://www.statsoft.com/support/books-on-statistica/>

Сравнительный анализ в биологии [Электрон. ресурс]. URL: <http://biocomparison.ucoz.ru/>

Biological Software (Software useful for biological studies) [Электрон. ресурс]. URL: <http://www.bioexplorer.net/Software/>

Software for Biological Collection Management [Электрон. ресурс]. URL: <http://www.bgbm.org/tdwg/acc/Software.htm>

List of statistical software. Wikipedia [Электрон. ресурс]. URL: https://en.wikipedia.org/wiki/List_of_statistical_software

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office OneNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook); Statsoft Statistica 10; PAST v.4.09; Punto Switcher 4.2 (2.9, 3.0)

– публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

б) информационные справочные системы:

– Электронный каталог Научной библиотеки ТГУ – <http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>

– Электронная библиотека (репозиторий) ТГУ – <http://vital.lib.tsu.ru/vital/access/manager/Index>

– ЭБС Лань – <http://e.lanbook.com/>

– ЭБС Консультант студента – <http://www.studentlibrary.ru/>

– Образовательная платформа Юрайт – <https://urait.ru/>

– ЭБС ZNANIUM.com – <https://znanium.com/>

– ЭБС IPRbooks – <http://www.iprbookshop.ru/>

14. Материально-техническое обеспечение

Аудитории для проведения занятий лекционного типа с мультимедийным оборудованием.

Аудитории для проведения практических занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации, оснащенные мультимедийным оборудованием, компьютерной техникой (настольные персональные компьютеры) и доступом в сеть Интернет, в электронную информационно-образовательную среду и к информационным справочным системам (компьютерные классы).

15. Информация о разработчике

Зверев Андрей Анатольевич, к.б.н., доцент, доц. кафедры ботаники ТГУ