

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное учреждение
высшего профессионального образования
"Казанский (Приволжский) федеральный университет"
Институт экологии и природопользования



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ

Проф. Талорский Д.А.

ДЕПАРТАМЕНТ
ОБРАЗОВАНИЯ
(ДО КФУ)

_____ 20__ г.

подписано электронно-цифровой подписью

Программа дисциплины

Статистическая обработка и анализ данных Б2.ДВ.1

Направление подготовки: 022000.62 - Экология и природопользование

Профиль подготовки: Моделирование в экологии

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Автор(ы):

Пилюгин А.Г.

Рецензент(ы):

Мухарамова С.С.

СОГЛАСОВАНО:

Заведующий(ая) кафедрой: Зарипов Ш. Х.

Протокол заседания кафедры No ____ от " ____ " _____ 201__ г

Учебно-методическая комиссия Института экологии и природопользования:

Протокол заседания УМК No ____ от " ____ " _____ 201__ г

Регистрационный No 241015

Казань
2015

Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре основной образовательной программы
3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля
4. Структура и содержание дисциплины/ модуля
5. Образовательные технологии, включая интерактивные формы обучения
6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов
7. Литература
8. Интернет-ресурсы
9. Материально-техническое обеспечение дисциплины/модуля согласно утвержденному учебному плану

Программу дисциплины разработал(а)(и) старший преподаватель, б/с Пилюгин А.Г. кафедра моделирования экологических систем отделение экологии , Alexander.Piliouguine@kpfu.ru

1. Цели освоения дисциплины

дать представление о способах и методах статистической обработки и анализ данных используя современные математические пакеты статистической обработки и специализированные языки программирования; сформировать навыки по использованию программного обеспечения в своей профессиональной деятельности; научить применению компьютерных методов статистической обработки данных экологических исследований и мониторинга.

2. Место дисциплины в структуре основной образовательной программы высшего профессионального образования

Данная учебная дисциплина включена в раздел " Б2.ДВ.1 Общепрофессиональный" основной образовательной программы 022000.62 Экология и природопользование и относится к дисциплинам по выбору. Осваивается на 3 курсе, 5 семестр.

Дисциплина относится к разделу Б2.ДВ1 ООП и развивает представления о об общих понятиях и принципах организации статистической обработки и анализ данных. Освоение данной дисциплины способствует развитию навыков использования статистической обработки и анализ данных экологических исследований и мониторинга. Для освоения курса необходимо прослушать курсы "Информатика" и "Теория вероятности и математическая статистика". Материал данного курса необходим при выполнении курсовых и дипломных работ, связанных с обработкой экологической информации. Дисциплина осваивается на 3-ем курсе (5 семестр).

3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОК-1 (общекультурные компетенции)	владеть культурой мышления, способностью к обобщению, анализу, восприятию информации, постановке цели и выбору путей ее достижения
ОК-13 (общекультурные компетенции)	владеть основными методами, способами и средствами получения, хранения, переработки информации, иметь навыки работы с компьютером как средством управления информацией
ОК-2 (общекультурные компетенции)	уметь логически верно, аргументированно и ясно строить устную и письменную речь
ОК-6 (общекультурные компетенции)	иметь базовые знания в области информатики и современных геоинформационных технологий, владеть навыками использования программных средств и работы в компьютерных сетях, умением создавать базы данных и использовать ресурсы Интернета, владеть ГИС-технологиями; уметь работать с информацией из различных источников для решения профессиональных и социальных задач

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-11 (профессиональные компетенции)	владеть методами экологического проектирования и экспертизы, экологического менеджмента и аудита, экологического и компьютерного картографирования; владеть методами классификации и статистической обработки, анализа и синтеза полевой и лабораторной экологической информации и использовать теоретические знания на практике
ПК-2 (профессиональные компетенции)	обладать базовыми знаниями фундаментальных разделов физики, химии и биологии в объеме необходимом для освоения физических, химических и биологических основ в экологии и природопользовании, иметь представление о современной естественнонаучной картине мира, владеть методами химического анализа, а также методами отбора и анализа геологических и биологических проб, иметь навыки идентификации и описания биологического разнообразия, его оценки современными методами количественной обработки информации, включая использование методов прикладной статистики и геоинформационных технологий
ПК-4 (профессиональные компетенции)	иметь базовые общепрофессиональные (общеекологические) представления о теоретических основах общей экологии, геоэкологии, экологии человека, социальной экологии, охраны окружающей среды, о системном анализе и моделировании в экологии и природопользовании
ПК-6 (профессиональные компетенции)	знать основы природопользования, экономики природопользования, устойчивого развития, методов оценки и прогнозирования воздействия на окружающую среду, правовых основ природопользования и охраны окружающей среды, быть способным понимать, излагать и критически анализировать базовую информацию в области экологии и природопользования, в том числе, с использованием методов математической статистики, геоинформационных технологий и математического моделирования
ПК-9 (профессиональные компетенции)	владеть методами прикладной экологии, экологического и компьютерного картографирования, экологической экспертизы и мониторинга; владеть методами классификации и статистической обработки, анализа и синтеза полевой и лабораторной экологической информации и использовать теоретические знания на практике

В результате освоения дисциплины студент:

1. должен знать:

основы языка R для статистической обработки данных и работы с графикой, основными алгоритмами обработки данных и способами их внешнего представления.

2. должен уметь:

ориентироваться в основных алгоритмах статистической обработки данных для решения практических задач и графического представления результатов.

3. должен владеть:

соответствующими приемами программирования, компьютерными методами обработки данных экологических исследований и мониторинга.

4. должен демонстрировать способность и готовность:

Студент должен знать

Основы моделирования данных и их анализа.

Студент должен уметь

Применять современные технологии для решения практических задач.

Студент должен владеть

Соответствующими навыками применения геоинформационных инструментов для решения практических задач.

Студент должен демонстрировать способность и готовность

применять полученные знания и навыки на практике, нести ответственность за результат выполнения работы.

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 2 зачетных(ые) единиц(ы) 72 часа(ов).

Форма промежуточного контроля дисциплины зачет в 5 семестре.

Суммарно по дисциплине можно получить 100 баллов, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов. Минимальное количество для допуска к зачету 28 баллов.

86 баллов и более - "отлично" (отл.);

71-85 баллов - "хорошо" (хор.);

55-70 баллов - "удовлетворительно" (удов.);

54 балла и менее - "неудовлетворительно" (неуд.).

4.1 Структура и содержание аудиторной работы по дисциплине/ модулю

Тематический план дисциплины/модуля

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Методы математической статистики.	5	1-5	10	0	0	устный опрос
2.	Тема 2. Работа с пакетом статистических исследований Statgraphics	5	6-10	4	0	10	контрольная работа

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
3.	Тема 3. Основы языка R для статистической обработки данных и работы с графикой	5	10-18	8	0	8	контрольная работа
	Тема . Итоговая форма контроля	5		0	0	0	зачет
	Итого			22	0	18	

4.2 Содержание дисциплины

Тема 1. Методы математической статистики.

лекционное занятие (10 часа(ов)):

Оценка параметров и проверка статистических гипотез. Статистическое описание. Определение и вычисление статистик случайной выборки. Типовые распределения вероятностей. Оценки параметров.

Тема 2. Работа с пакетом статистических исследований Statgraphics

лекционное занятие (4 часа(ов)):

Работа с пакетом статистических исследований Statgraphics. Меню программы. Поддерживаемые статистические методы. Формы вывода результатов.

лабораторная работа (10 часа(ов)):

Работа с пакетом статистических исследований Statgraphics 1.Создание таблиц при помощи программы Statgraphics. 2.Графический анализ распределения параметров, построение гистограмм. 3.Анализ функций распределения: нормальное и логнормальное распределения. 4.Простая корреляция, таблица парных коэффициентов. 5.Уравнение регрессии: графическое и цифровое представление материала и его анализ.

Тема 3. Основы языка R для статистической обработки данных и работы с графикой

лекционное занятие (8 часа(ов)):

Основы языка R для статистической обработки данных и работы с графикой. Основные операторы. Данные в R. Графики в R. Статистические методы в R.

лабораторная работа (8 часа(ов)):

Отображение и удаление объектов в памяти. Вызов справки. Данные в R. Чтение данных из текстового файла. Сохранение текстовых данных. Таблицы произвольной структуры. Загрузка и запись данных. Генерация (создание) данных. Управление объектами. Управление графическими окнами. Графические функции. Команды управления графиками нижнего уровня. Графические параметры.

4.3 Структура и содержание самостоятельной работы дисциплины (модуля)

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
1.	Тема 1. Методы математической статистики.	5	1-5	подготовка к устному опросу	6	устный опрос

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
2.	Тема 2. Работа с пакетом статистических исследований Statgraphics	5	6-10	подготовка к контрольной работе	12	контрольная работа
3.	Тема 3. Основы языка R для статистической обработки данных и работы с графикой	5	10-18	подготовка к контрольной работе	14	контрольная работа
	Итого				32	

5. Образовательные технологии, включая интерактивные формы обучения

Для проведения лекции используется возможности мультимедийного класса с проектированием части лекций в виде презентации на экран, с включением в лекцию видео материалов. В интерактивной форме проходят все практические занятия, во время которых предполагаются вопросы студентов, ответы на них и рекомендации со стороны преподавателя. Интерактивные формы проведения занятий составляют 45% аудиторной нагрузки.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

Тема 1. Методы математической статистики.

устный опрос , примерные вопросы:

Повторить методы проверки статистических гипотез, методы статистического описания выборки, оценки параметров, типовые распределения вероятностей. Установить на рабочий компьютер программу R, нарабатывать практических навыки работы с программой. Использовать знания, полученные на лекциях и лабораторных занятиях по данной дисциплине, а также на занятиях по "Теории вероятностей и мат.статистке". Использовать основную и дополнительную литературу, Интернет-ресурсы.

Тема 2. Работа с пакетом статистических исследований Statgraphics

контрольная работа , примерные вопросы:

1. Имеется выборка наблюдаемых значений объема 60. Задание: Провести статистическое описание выборки: построить таблицу частот и гистограмму; оценить среднее, медиану, моду, дисперсию, среднеквадратическое отклонение; получить доверительные интервалы для матожидания, дисперсии. Проверить гипотезу о том, что выборка извлечена из нормально распределенной генеральной совокупности. Доверительная вероятность 0.95; уровень значимости 0.05. 2. Разработан прибор новой модификации. Для сравнения с предшествующей модификацией были произведены 20 одновременных измерений обоими приборами при одинаковых эксплуатационных условиях. Вопрос: Обеспечивает ли новый прибор большую точность измерений по сравнению с ранее введенным в эксплуатацию, если в среднем показания приборов одинаковы. Уровень значимости 0.05. 3. Были проведены измерения концентрации ртути в воде реки выше и ниже места сброса с очистных сооружений (по 20 проб выше и ниже сброса). Вопрос: Влияет ли сброс на концентрацию ртути? Уровень значимости 0.05.

Тема 3. Основы языка R для статистической обработки данных и работы с графикой

контрольная работа , примерные вопросы:

1. Имеются 3 расчетные методики по определению максимальной концентрации ВЗВ. Произведено n_1 , n_2 , n_3 , расчетов по каждой из методик при одинаковых условиях эксперимента. В среднем результаты расчетов одинаковы. Вопрос: Можно ли отдать предпочтение одной из методик? Примечание: Использовать критерии Бартлетта или Кочрена в зависимости от числа экспериментов. 2. Производится оценка влияния загрязнения придорожной растительности выхлопами автотранспорта. Проведено 20 замеров концентрации свинца в вегетативной части растений на расстоянии 50, 100 и 150 см от дороги. Вопрос: Влияет ли фактор удаленности от дороги на концентрацию свинца в растениях? 3. В течении 20 лет проводились совместные исследования заболеваемости (на 1000 населения) ОРЗ и уровня загазованности. Задание: Вычислить корреляцию между этими величинами, построить линейную регрессионную модель, проверить ее адекватность и выдать прогноз заболеваемости в случае превышения среднего уровня загазованности в 3 раза.

Тема . Итоговая форма контроля

Примерные вопросы к зачету:

Для аттестации по итогам освоения дисциплины проводится зачет. Текущий контроль успеваемости и контроль выполнения самостоятельной работы осуществляются путем устного опроса студентов на лекциях, проверки решений на практических занятиях и по результатам выполнения контрольных работ. Самостоятельная работа студентов предполагает проработку теоретического материала и развитие практических навыков с использованием основной и дополнительной литературы и по самостоятельно найденным Интернет-источникам, а также выполнение контрольных заданий .

Вопросы на зачет

1. Типы данных. Параметрические и непараметрические методы.
2. Параметрические показатели связи.
3. Непараметрические показатели связи количественных и порядковых признаков.
4. Корреляционный анализ номинальных признаков.
5. Меры связи номинальных признаков.
6. Сравнительная оценка силы взаимосвязи признаков.
7. Общая линейная модель.
8. Метод главных компонент.

7.1. Основная литература:

- Теория вероятностей и математическая статистика, Гмурман, Владимир Ефимович, 2007г.
2. Буре В. М., Парилина Е. М. Теория вероятностей и математическая статистика. Лань, 2013. - 416 с. (http://e.lanbook.com/books/element.php?pl1_cid=25&pl1_id=10249)
 3. Гусева, Е. Н. Теория вероятностей и математическая статистика : Уч. пособ. / Е. Н. Гусева. - 5-е изд., стереотип. - М. : Флинта, 2011. - 220 с. (<http://www.znaniium.com/catalog.php?bookinfo=406064>)
 4. Савельев А.А., Мухарамова С.С., Пилюгин А.Г. Использование языка R для статистической обработки данных. Учебно-методическая разработка.- Казань: Казанский ун-т, 2007. - 28 с. - http://kpfu.ru/publication?p_id=59689, http://shelly.ksu.ru/e-ksu/docs/F1594376599/%EC%E5%F2%EE%E4%E8%F7%EA%E0_R_2.pdf
 5. Савельев А.А. , Мухарамова С.С., Пилюгин А.Г. Основные понятия языка R. Учебно-методическая разработка. Казань: Казанский ун-т, 2007. - 29 с. -http://kpfu.ru/publication?p_id=59687, http://shelly.ksu.ru/e-ksu/docs/F912705413/%EC%E5%F2%EE%E4%E8%F7%EA%E0_R_1.pdf

7.2. Дополнительная литература:

Статистический анализ данных в экологии и природопользовании с использованием программы STATGRAPHICS Plus, Мальцев, Кирилл Александрович; Мухарамова, Светлана Саясовна, 2011г.

2. Общая теория статистики: Учебное пособие / С.Н. Лысенко, И.А. Дмитриева. - Изд., испр. и доп. - М.: Вузовский учебник: НИЦ ИНФРА-М, 2014. - 219 с.
-<http://www.znaniium.com/catalog.php?bookinfo=397795>

7.3. Интернет-ресурсы:

математический образовательный сайт - Allmath.ru

математический образовательный сайт, - www.exponenta.ru

методическая литература по дисциплине. - <http://www.edu.ru>

Портал "Машинное обучение" - <http://www.machinelearning.ru/>

электронная база научной литературы - Allmath.ru

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Статистическая обработка и анализ данных" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, вместимостью более 60 человек. Мультимедийная аудитория состоит из интегрированных инженерных систем с единой системой управления, оснащенная современными средствами воспроизведения и визуализации любой видео и аудио информации, получения и передачи электронных документов. Типовая комплектация мультимедийной аудитории состоит из: мультимедийного проектора, автоматизированного проекционного экрана, акустической системы, а также интерактивной трибуны преподавателя, включающей тач-скрин монитор с диагональю не менее 22 дюймов, персональный компьютер (с техническими характеристиками не ниже Intel Core i3-2100, DDR3 4096Mb, 500Gb), конференц-микрофон, беспроводной микрофон, блок управления оборудованием, интерфейсы подключения: USB, audio, HDMI. Интерактивная трибуна преподавателя является ключевым элементом управления, объединяющим все устройства в единую систему, и служит полноценным рабочим местом преподавателя. Преподаватель имеет возможность легко управлять всей системой, не отходя от трибуны, что позволяет проводить лекции, практические занятия, презентации, вебинары, конференции и другие виды аудиторной нагрузки обучающихся в удобной и доступной для них форме с применением современных интерактивных средств обучения, в том числе с использованием в процессе обучения всех корпоративных ресурсов. Мультимедийная аудитория также оснащена широкополосным доступом в сеть интернет. Компьютерное оборудование имеет соответствующее лицензионное программное обеспечение.

Компьютерный класс, представляющий собой рабочее место преподавателя и не менее 15 рабочих мест студентов, включающих компьютерный стол, стул, персональный компьютер, лицензионное программное обеспечение. Каждый компьютер имеет широкополосный доступ в сеть Интернет. Все компьютеры подключены к корпоративной компьютерной сети КФУ и находятся в едином домене.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен студентам. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, УМК, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего профессионального образования (ФГОС ВПО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе Издательства "Лань", доступ к которой предоставлен студентам. ЭБС Издательства "Лань" включает в себя электронные версии книг издательства "Лань" и других ведущих издательств учебной литературы, а также электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. ЭБС Издательства "Лань" обеспечивает доступ к научной, учебной литературе и научным периодическим изданиям по максимальному количеству профильных направлений с соблюдением всех авторских и смежных прав.

Компьютеры с установленными на них интегрированными математическими пакетами.

www.sciencedirect.com - электронная база научной литературы.

Мультимедийные средства.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 022000.62 "Экология и природопользование" и профилю подготовки Моделирование в экологии .

Автор(ы):

Пилюгин А.Г. _____

"__" _____ 201__ г.

Рецензент(ы):

Мухарамова С.С. _____

"__" _____ 201__ г.