

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. Таюрский Д.А.



_____ 20__ г.

подписано электронно-цифровой подписью

Программа дисциплины

Статистическая обработка и анализ данных Б1.В.ДВ.9

Направление подготовки: 21.03.02 - Землеустройство и кадастры

Профиль подготовки: Землеустройство

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 244217

Казань
2017

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 " Б1.В.ДВ.9 Дисциплины (модули)" основной образовательной программы 21.03.02 Землеустройство и кадастры 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-12 (общекультурные компетенции)	понимать сущность и значение информации в развитии современного информационного общества, сознавать опасности и угрозы, возникающие в этом процессе, соблюдать основные требования информационной безопасности, в том числе защиты государственной тайны
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-1 (профессиональные компетенции)	обладать базовыми знаниями в области фундаментальных разделов математики в объеме, необходимом для владения математическим аппаратом экологических наук, для обработки информации и анализа данных по экологии и природопользованию
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

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 4 зачетных(ые) единиц(ы) 144 часа(ов).

Форма промежуточного контроля дисциплины экзамен 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: оценка параметров и проверка статистических гипотез Статистическое описание. Определение и вычисление статистик случайной выборки Типовые распределения вероятностей Оценки параметров	5	1-7	6	12	0	Письменное домашнее задание
2.	Тема 2. Работа с пакетом статистических исследований Statgraphics	5	8-15	6	12	0	Письменное домашнее задание
3.	Тема 3. Основы языка R для статистической обработки данных и работы с графикой	5	16-18	6	12	0	Письменное домашнее задание
	Тема . Итоговая форма контроля	5		0	0	0	Экзамен
	Итого			18	36	0	

4.2 Содержание дисциплины

Тема 1. Методы математической статистики: оценка параметров и проверка статистических гипотез
Статистическое описание. Определение и вычисление статистик случайной выборки
Типовые распределения вероятностей
Оценки параметров

лекционное занятие (6 часа(ов)):

Методы математической статистики: оценка параметров и проверка статистических гипотез
 Статистическое описание. Определение и вычисление статистик случайной выборки. Типовые распределения вероятностей Введение в статистические методы. Статистическое описание. Определение и вычисление статистик случайной выборки. Типовые распределения вероятностей. Оценки параметров. Выборочные распределения. Проверка статистических гипотез. Некоторые статистики, выборочные распределения и критерии для многомерных распределений.

практическое занятие (12 часа(ов)):

Методы математической статистики: оценка параметров и проверка статистических гипотез

Тема 2. Работа с пакетом статистических исследований Statgraphics

лекционное занятие (6 часа(ов)):

Работа с пакетом статистических исследований Statgraphics

практическое занятие (12 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 (6 часа(ов)):

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

практическое занятие (12 часа(ов)):

Основы языка R для статистической обработки данных и работы с графикой Оператор <-.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: оценка параметров и проверка статистических гипотез Статистическое описание. Определение и вычисление статистик случайной выборки Типовые распределения вероятностей Оценки параметров	5	1-7	подготовка домашнего задания	8	домашнее задание
				подготовка к отчету	2	отчет

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	8-15	подготовка домашнего задания	18	домашнее задание
				подготовка к отчету	2	отчет
3.	Тема 3. Основы языка R для статистической обработки данных и работы с графикой	5	16-18	подготовка домашнего задания	22	домашнее задание
				подготовка к отчету	2	отчет
	Итого				54	

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

домашнее задание , 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.

отчет , примерные вопросы:

по результатам работы оформляется отчет

Тема 2. Работа с пакетом статистических исследований Statgraphics

домашнее задание , примерные вопросы:

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 математического ожидания, 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.

отчет , примерные вопросы:

по результатам работы оформляется отчет

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

домашнее задание , примерные вопросы:

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.

отчет , примерные вопросы:

по результатам работы оформляется отчет

Тема . Итоговая форма контроля

Примерные вопросы к экзамену:

Для аттестации по итогам освоения дисциплины проводится экзамен. 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 к экзамену:

1. Типы данных.
2. Параметрические и непараметрические методы.
3. Параметрические показатели связи.
4. Непараметрические показатели связи количественных и порядковых признаков.
5. Корреляционный анализ номинальных признаков.
6. Типовые распределения вероятностей
7. Нулевая и конкурирующая гипотезы.
8. Сравнение непрерывных распределений
9. Сравнение выборочной средней с гипотетической средней генеральной совокупности
10. Сравнение двух средних нормальных генеральных совокупностей
11. Сравнение двух средних нормальных генеральных совокупностей с неизвестными дисперсиями.
12. Сравнение двух средних нормальных генеральных совокупностей с известными дисперсиями.
13. Сравнение дисперсий двух нормальных генеральных совокупностей
14. Дисперсионный анализ
15. Меры связи номинальных признаков.
16. Сравнительная оценка силы взаимосвязи признаков.
17. Общая линейная модель.
18. Метод главных компонент.

7.1. Основная литература:

- 1.Использование языка R для статистической обработки данных : учебно-методическое пособие / Казан. гос. ун-т, Фак. географии и экологии ; сост. А. А. Савельев, С. С. Мухарамова, А. Г. Пилюгин .-Казань : Казанский государственный университет, 2007 .- 28 с.
2. Использование языка R для статистической обработки данных : учебно-методическое пособие / Казан. гос. ун-т, Фак. географии и экологии ; сост. А. А. Савельев, С. С. Мухарамова, А. Г. Пилюгин .-Казань : Казанский государственный университет, 2007 .- 28 с. - в свободном доступе в каталоге учебных ресурсов КФУ - URL: http://kpfu.ru/docs/F407025247/metodichka_R_2.pdf
- 3.Статистический анализ данных в MS Excel: Учебное пособие / А.Ю. Козлов, В.С. Мхитарян, В.Ф. Шишов. - М.: ИНФРА-М, 2014. - 320 с.: <http://znanium.com/bookread.php?book=429722>

7.2. Дополнительная литература:

- 1.Общая теория статистики: Учебное пособие / С.Н. Лысенко, И.А. Дмитриева. - Изд., испр. и доп. - М.: Вузовский учебник: НИЦ ИНФРА-М, 2014. - 219 с.: 70x100 1/16. (переплет) ISBN 978-5-9558-0115-5 -URL:<http://www.znanium.com/catalog.php?bookinfo=397795> - Загл. с экрана.
- 2.Основные понятия языка R : учебно-методическое пособие / Казан. гос. ун-т, Фак. географии и экологии ; сост. А. А. Савельев, С. С. Мухарамова, А. Г. Пилюгин, Е. А. Алексеева .- Казань : Казанский государственный университет, 2007 .- 29 с.
- 3.Основные понятия языка R : учебно-методическое пособие / Казан. гос. ун-т, Фак. географии и экологии ; сост. А. А. Савельев, С. С. Мухарамова, А. Г. Пилюгин, Е. А. Алексеева .- Казань : Казанский государственный университет, 2007 .- 29 с. - в свободном доступе в каталоге учебных ресурсов КФУ - URL: http://kpfu.ru/docs/F568269105/metodichka_R_1.pdf

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:

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 21.03.02 "Землеустройство и кадастры" и профилю подготовки Землеустройство

Автор(ы):

Пилюгин А.Г. _____

"__" _____ 201__ г.

Рецензент(ы):

Мухарамова С.С. _____

"__" _____ 201__ г.