

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



Программа дисциплины

Компьютерные технологии и статистические методы в экологии и природопользовании М1.Б.3

Направление подготовки: 022000.68 - Экология и природопользование

Профиль подготовки: Системная экология и моделирование

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Автор(ы):

Мухарамова С.С. , Савельев А.А.

Рецензент(ы):

-

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 2113

Казань
2014

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) доцент, к.н. (доцент) Мухарамова С.С. кафедра моделирования экологических систем отделение экологии , Svetlana.Mukharamova@kpfu.ru ; профессор, д.н. (профессор) Савельев А.А. кафедра моделирования экологических систем отделение экологии , Anatoly.Saveliev@kpfu.ru

1. Цели освоения дисциплины

Целями освоения дисциплины "Компьютерные технологии и статистические методы в экологии и природопользовании" являются подготовка к научно-исследовательской деятельности, связанной с изучением и численным описанием явлений в области экологии, природопользования, а также в смежных областях. В результате освоения данной дисциплины должны быть изучены современные компьютерные технологии обработки информации методами математической статистики, приобретены навыки работы с программными средствами анализа и моделирования данных.

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

Данная учебная дисциплина включена в раздел " М1.Б.3 Общенаучный" основной образовательной программы 022000.68 Экология и природопользование и относится к базовой (общепрофессиональной) части. Осваивается на 1 курсе, 1 семестр.

Дисциплина "Компьютерные технологии и статистические методы в экологии и природопользовании" относится к общенаучной части учебной программы магистратуры "Системная экология и моделирование". Ее изучение базируется на знаниях и умениях, полученных при изучении общих курсов по теории вероятностей и математической статистики в рамках бакалавриата или специалитета.

Знания и умения, полученные при изучении данной дисциплины, необходимы для освоения дисциплин "Современные статистические модели", "Моделирование в экологии", а также при выполнении магистерских работ, связанных с обработкой и анализом данных в области экологии и природопользования.

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

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОК-2 (общекультурные компетенции)	знание современных компьютерных технологий, применяемых при сборе, хранении, обработке, анализе и передаче географической информации; способностью самостоятельно использовать современные компьютерные технологии для решения научно-исследовательских и производственно-технологических задач профессиональной деятельности (ОК-2)
ОК-4 (общекультурные компетенции)	владение методами оценки репрезентативности материала, объема выборок при проведении количественных исследований, статистическими методами сравнения полученных данных и определения закономерностей (ОК-4)

В результате освоения дисциплины студент:

1. должен знать:

современные компьютерные технологии, применяемые при обработке и анализе информации в области экологии и природопользования

2. должен уметь:

самостоятельно выбирать подходящие методы статистического анализа и моделирования, проверять выполнение условий их применения, выбирать программные средства, реализующие эти методы

3. должен владеть:

навыками работы с компьютерными программами статистической обработки данных и моделирования

4. должен демонстрировать способность и готовность:

-

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет зачетных(ые) единиц(ы) 108 часа(ов).

Форма промежуточного контроля дисциплины экзамен в 1 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 описательной статистики. Проверка гипотез. Исследование связи количественных признаков. Корреляционный анализ и регрессионный анализ. Статистическое программное обеспечение. Язык и среда статистического программирования R.	1	1-2	6	0	2	домашнее задание

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
2.	Тема 2. Программы на языке R. Создание на языке R скрипта для статистического описания набора одномерных выборок. Реализация на языке R проверки статистических гипотез.	1	3-4	0	0	8	домашнее задание
3.	Тема 3. Реализация на языке R исследования связи количественных признаков. Построение линейной регрессии. Реализация на языке R обобщенной линейной модели.	1	5-6	0	0	8	домашнее задание
4.	Тема 4. Реализация на языке R моделирования нелинейных зависимостей. Сравнение моделей.	1	7-8	0	0	8	домашнее задание
	Тема . Итоговая форма контроля	1		0	0	0	экзамен
	Итого			6	0	26	

4.2 Содержание дисциплины

Тема 1. Методы описательной статистики. Проверка гипотез. Исследование связи количественных признаков. Корреляционный анализ и регрессионный анализ. Статистическое программное обеспечение. Язык и среда статистического программирования R.

лекционное занятие (6 часа(ов)):

Лекция 1. Классификация типов данных. Шкалы измерений. Методы описательной статистики (таблицы частот, гистограммы, описательные статистики). Параметрические и непараметрические методы, гипотеза о нормальности распределения. Проверка гипотез. Дисперсионный анализ. Статистическое программное обеспечение. Лекция 2. Язык и среда статистического программирования R. Проект GNU, программная среда с открытым исходным кодом, свободно распространяемый продукт, сайт разработчиков (The Comprehensive R Archive Network). Библиотеки R, поддерживаемые методы. Установка программы, загрузка и установка пакетов. Основные операторы. Данные в R (объекты, типы данных, чтение данных из текстовых файлов, сохранение данных, таблицы произвольной структуры, генерация данных). Управление объектами. Арифметика и простые функции, операции с матрицами. Графики в R. Лекция 3. Исследование связи количественных признаков. Корреляционный анализ. Простая линейная регрессия. Множественная линейная регрессия. Общая линейная модель. Обобщенная линейная модель. Обобщенная аддитивная модель. Анализ остатков. Коэффициент детерминации. Выбор наилучшего подмножества регрессоров. Информационные критерии (AIC, BIC), коэффициент возрастания дисперсии (VIF). Сравнение моделей. Интерпретация моделей.

лабораторная работа (2 часа(ов)):

Лабораторное занятие 1. Реализация на языке R статистического описания одномерной выборки: Чтение данных из текстового файла. Выборка. Объем выборки. Вычисление описательных статистик. Вывод на экран. Доверительные интервалы. Бокс-плот (ящик с усами). Построение таблицы частот и гистограммы. Проверка на нормальность (q-q графики, критерий хи-квадрат, критерий Колмогорова-Смирнова). Вывод результатов в текстовый файл. Вывод графических результатов в файл.

Тема 2. Программы на языке R. Создание на языке R скрипта для статистического описания набора одномерных выборок. Реализация на языке R проверки статистических гипотез.

лабораторная работа (8 часа(ов)):

Лабораторное занятие 2, 3. Написание на языке R скрипта для статистического описания набора одномерных выборок: Программы на языке R. Редактор Tinn-R. Операторы цикла. Пользовательские функции. Загрузка данных. Последовательное формирование выборок и их анализ. Вывод результатов в текстовые и графические файлы. Интерпретация результатов статистического описания. Лабораторное занятие 4, 5. Реализация на языке R проверки статистических гипотез: Генерация данных. Проверка гипотез. Проверка гипотезы о равенстве дисперсий (критерий Фишера). Проверка гипотезы о равенстве матожиданий (критерий Стьюдента). Вывод результатов с использованием условного оператора. Дисперсионный анализ. Реализация на языке R исследования связи двух количественных признаков. Корреляционный анализ.

Тема 3. Реализация на языке R исследования связи количественных признаков. Построение линейной регрессии. Реализация на языке R обобщенной линейной модели.

лабораторная работа (8 часа(ов)):

Лабораторное занятие 6, 7. Реализация на языке R исследования связи количественных признаков. Построение простой линейной регрессии. Функция $lm()$, параметры и их использование. Анализ остатков. Диаграмма рассеяния, визуализация построенной модели, стандартный графический вывод, построение дополнительных графиков. Прогноз, функция $predict()$. Преобразование переменных. Интерпретация полученных результатов. Построение множественной линейной регрессии и ее анализ. Отклонения от предположений линейной модели (коллинеарность, ее влияние на результаты моделирования, VIF, выбор регрессоров). Анализ зависимостей и нелинейностей в остатках, гетерогенности (зависимость дисперсии остатков от модельного матожидания), пространственной и временной корреляции). Лабораторное занятие 8, 9. Реализация на языке R обобщенной линейной модели. Скалярные переменные и факторы, преобразование. Запись линейного предиктора, знаки арифметических операций и пр., операция $I()$. Возможность динамического формирования формул. Функция $glm()$. Параметры функции, анализ остатков, overdispersion, прогноз. Выбор наилучшего подмножества регрессоров. Информационные критерии (AIC, BIC), коэффициент возрастания дисперсии (VIF). Интерпретация результатов.

Тема 4. Реализация на языке R моделирования нелинейных зависимостей. Сравнение моделей.

лабораторная работа (8 часа(ов)):

Лабораторное занятие 10, 11. Реализация на языке R моделирования нелинейных зависимостей. Обобщенная аддитивная модель, функция gam(). Оценка оптимальной степени сглаживания с использованием перекрестной проверки. Оценка нелинейности, эффективное число степеней свободы. Управление сглаживанием, выбор базиса. Анализ остатков. Выбор наилучшего подмножества регрессоров. Информационные критерии, коэффициент VIF. Интерпретация результатов. Лабораторное занятие 12, 13. Реализация на языке R сравнения моделей, функцияanova(), сравнение моделей с использованием информационных критериев, функция AIC().

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 описательной статистики. Проверка гипотез. Исследование связи количественных признаков. Корреляционный анализ и регрессионный анализ. Статистическое программное обеспечение. Язык и среда статистического программирования R.	1	1-2	подготовка домашнего задания	10	домашнее задание
2.	Тема 2. Программы на языке R. Создание на языке R скрипта для статистического описания набора одномерных выборок. Реализация на языке R проверки статистических гипотез.	1	3-4	подготовка домашнего задания	10	домашнее задание
3.	Тема 3. Реализация на языке R исследования связи количественных признаков. Построение линейной регрессии. Реализация на языке R обобщенной линейной модели.	1	5-6	подготовка домашнего задания	10	домашнее задание

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
4.	Тема 4. Реализация на языке R моделирования нелинейных зависимостей. Сравнение моделей.	1	7-8	подготовка домашнего задания	10	домашнее задание
	Итого				40	

5. Образовательные технологии, включая интерактивные формы обучения

- компьютерные презентации лекций;
- лекционное изложение включает разбор конкретных ситуаций;
- для решения практических задач используются компьютерные симуляции.

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

Тема 1. Методы описательной статистики. Проверка гипотез. Исследование связи количественных признаков. Корреляционный анализ и регрессионный анализ. Статистическое программное обеспечение. Язык и среда статистического программирования R.

домашнее задание , примерные вопросы:

Использовать знания, полученные на лекциях и лабораторном занятии по данной дисциплине. Повторить материалы курса бакалавриата "Теория вероятностей и мат. статистика". Использовать основную и дополнительную литературу, Интернет-ресурсы. Загрузить установочный файл R-*-.win.exe с сайта разработчиков языка статистического программирования R (<http://cran.r-project.org/>). Установить программу R на компьютер. Загрузить и установить пакеты. Ознакомиться со справочной документацией. Выполнить статистическое описание одномерной выборки на языке R. Вычислить описательные статистики, доверительные интервалы. Построить таблицу частот и гистограмму, бокс-плот. Проверить гипотезу о нормальности распределения. Вывести результаты в текстовый файл и графический файл.

Тема 2. Программы на языке R. Создание на языке R скрипта для статистического описания набора одномерных выборок. Реализация на языке R проверки статистических гипотез.

домашнее задание , примерные вопросы:

Использовать знания, полученные на лекциях и лабораторных занятиях по данной дисциплине. Использовать основную и дополнительную литературу, Интернет-ресурсы, справочную документацию пакета R. Установить на компьютер редактор Tinn-R. Создать программу (скрипт) на языке R для статистического описания набора одномерных выборок и проверки гипотез о равенстве дисперсий и равенстве матожиданий с использованием операторов цикла, пользовательских функций. Создать программу (скрипт) на языке R для исследования связи двух количественных признаков.

Тема 3. Реализация на языке R исследования связи количественных признаков. Построение линейной регрессии. Реализация на языке R обобщенной линейной модели.

домашнее задание , примерные вопросы:

Использовать знания, полученные на лекциях и лабораторных занятиях по данной дисциплине. Использовать основную и дополнительную литературу, Интернет-ресурсы, справочную документацию пакета R. Создать программу (скрипт) на языке R для построения простой линейной регрессии. Создать программу (скрипт) на языке R для построения множественной линейной регрессии. Создать программу (скрипт) на языке R для построения обобщенной линейной модели.

Тема 4. Реализация на языке R моделирования нелинейных зависимостей. Сравнение моделей.

домашнее задание , примерные вопросы:

Использовать знания, полученные на лекциях и лабораторных занятиях по данной дисциплине. Использовать основную и дополнительную литературу, Интернет-ресурсы, справочную документацию пакета R. Создать программу (скрипт) на языке R для построения обобщенной аддитивной (нелинейной) модели. Провести сравнение моделей. Написать отчет, содержащий интерпретацию полученных результатов.

Тема . Итоговая форма контроля

Примерные вопросы к экзамену:

Примеры вопросов в экзаменационных билетах

1. Предположения линейной модели.
2. Анализ остатков.
3. Выбор наилучшего подмножества регрессоров.
4. Сравнение моделей с использованием информационных критериев.

7.1. Основная литература:

1. Гмурман В.Е. Теория вероятностей и математическая статистика: учебное пособие для студентов вузов / В. Е. Гмурман.-12-е изд., перераб.-Москва: Высш. образование: Юрайт-Издат, 2009.-478с.
2. Чашкин Ю.Р. Математическая статистика: анализ и обработка данных: учебное пособие для студентов высших учебных заведений / Ю. Р. Чашкин.-Изд. 2-е, перераб. и доп..-Ростов-на-Дону: Феникс, 2010.-236с.
3. Савельев А.А., Мухарамова С.С., Пилюгин А.Г., Чижикова Н.А. Геоэкономический анализ данных в экологии и природопользовании (с применением пакета R). - Казань: Казанский ун-т, 2012. - 120с.http://kpfu.ru/publication?p_id=59674,
<http://shelly.ksu.ru/e-ksu/docs/F1387902447/geostat-2012.pdf>
4. Савельев А.А., Мухарамова С.С., Пилюгин А.Г. Использование языка R для статистической обработки данных. Учебно-методическая разработка.- 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. - 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:

1. А.Б. Шипунов, Е.М. Балдин, П.А. Волкова, А.И. Коробейников, С. А. Назарова, С.В. Петров, В.Г. Суфиянов. Наглядная статистика. Используем R! - М.: ДМК Пресс, 2012. - 298 с.: ил. - ISBN 978-5-94074-785-828-1
2. Wood S.N. Generalized Additive Models: An Introduction with R / S.N. Wood. - London: Chapman and Hall, 2006. - 410 p.

7.3. Интернет-ресурсы:

- 1 Сайт разработчиков языка статистического программирования R - <http://cran.r-project.org/>
- 2 Портал Машинное обучение - <http://www.machinelearning.ru/>
- 3 Математический образовательный сайт - www.exponenta.ru
- 4 Вероятностно-статистические основы современных статистических методов - <http://ru.wikibooks.org/wiki/%D0%9C%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%>
- 5 Сайт по биоинформатике - <http://bioinformatics.ru/>
- 6 Гланц С. Медико-биологическая статистика. Электронная книга. - <http://ru.scribd.com/doc/19937791/SGlantz-MedBViolStatistics>
- 7 Уч.-метод. пособие, применение стат.пакета SPSS - www.matlab.mgppu.ru/work/0022.htm
- 8 Обучающий сайт по статобработке - <http://stattrek.com/>

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Компьютерные технологии и статистические методы в экологии и природопользовании" 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.

нет

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 022000.68 "Экология и природопользование" и магистерской программе Системная экология и моделирование .

Автор(ы):

Мухарамова С.С. _____

Савельев А.А. _____

"__" _____ 201__ г.

Рецензент(ы):

"__" _____ 201__ г.