

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**федеральное государственное автономное
образовательное учреждение высшего образования_
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Институт биологии и биомедицины

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Рабочая программа дисциплины

Анализ данных в экологии

Уровень высшего образования

Бакалавриат

Направление подготовки / специальность

05.03.06 - Экология и природопользование

Направленность образовательной программы

Экология

Форма обучения

очная

г. Нижний Новгород

2024 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.02 Анализ данных в экологии относится к части, формируемой участниками образовательных отношений образовательной программы.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства	
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	Для текущего контроля успеваемости	Для промежуточной аттестации
УК-1: Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1: Анализирует задачу, выделяя ее базовые составляющие УК-1.2: Находит и критически анализирует необходимую информацию УК-1.3: Критически рассматривает возможные варианты решения задачи УК-1.4: Грамотно, логично, аргументированно формирует собственные суждения и оценки УК-1.5: Определяет и оценивает последствия возможных решений задачи	УК-1.1: Способен определить базовые составляющие задачи по анализу экологических данных УК-1.2: Способен находить необходимые источники информации для анализа экологических данных УК-1.3: Способен критически рассмотреть разные варианты проведения анализа экологических данных УК-1.4: Способен формировать обоснованные выводы по результатам анализа экологических данных УК-1.5: Способен определить и оценить ограничения полученных результатов и выводов	Тест	Экзамен: Тест
ПК-21: Владеет методами геохимических и геофизических исследований, общего и	ПК-21.1: Знает: - геохимических и геофизических исследований, общего и геоэкологического картографирования	ПК-21.1: Знает принципы применения теории вероятностей и математической статистики для описания данных	Расчетно-графическое задание	Экзамен: Расчетно-графическое задание

геоэкологического картографирования, методами обработки, анализа и синтеза полевой и лабораторной экологической информации	<p>ПК-21.2: Умеет: - использовать методы анализа и синтеза полевой и лабораторной геоэкологической информации</p> <p>ПК-21.3: Владеет: - методами обработки, анализа и синтеза полевой и лабораторной экологической информации</p>	<p>экологических исследований и мониторинга</p> <p>ПК-21.2: Умеет применять методы описания, статистического анализа и визуализации данных экологических исследований и мониторинга</p> <p>ПК-21.3: Владеет компьютерными средствами для подготовки, хранения и анализа данных экологических исследований и мониторинга</p>		
--	--	---	--	--

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	6
Часов по учебному плану	216
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	40
- занятия семинарского типа (практические занятия / лабораторные работы)	56
- КСР	4
самостоятельная работа	44
Промежуточная аттестация	72
	Экзамен

3.2. Содержание дисциплины

(структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий)

Наименование разделов и тем дисциплины	Всего (часы)	в том числе			
		Контактная работа (работа во взаимодействии с преподавателем), часы из них			Самостоятельная работа обучающегося, часы
		Занятия лекционного типа	Занятия семинарского типа (практические занятия/лабораторные работы), часы	Всего	
	о ф о	о ф о	о ф о	о ф о	о ф о

Тема 1. Основные понятия количественного анализа данных	6	2	2	4	2
Тема 2. Основы программирования на языке R	16	2	8	10	6
Тема 3. Элементы теории вероятностей	8	4	2	6	2
Тема 4. Основы базовой графической системы R	10	2	4	6	4
Тема 5. Нормальное распределение	8	4	2	6	2
Тема 6. Описательная статистика и визуализация данных в R	8	2	4	6	2
Тема 7. Проверка статистических гипотез	10	4	4	8	2
Тема 8. Нормальное распределение и проверка гипотез	8	4	2	6	2
Тема 9. Анализ качественных данных	8	2	4	6	2
Тема 10. Анализ средних: критерий Стьюдента и дисперсионный анализ	18	4	8	12	6
Тема 11. Регрессионный анализ	10	2	4	6	4
Тема 12. Многофакторный анализ	10	2	4	6	4
Тема 13. Обобщенные линейные модели	8	2	4	6	2
Тема 14. Бутстреп и рандомизация	12	4	4	8	4
Аттестация	72				
КСР	4			4	
Итого	216	40	56	100	44

Содержание разделов и тем дисциплины

Тема 1. Основные понятия количественного анализа данных
Тема 2. Основы программирования на языке R
Тема 3. Элементы теории вероятностей
Тема 4. Основы базовой графической системы R
Тема 5. Нормальное распределение
Тема 6. Описательная статистика и визуализация данных в R
Тема 7. Проверка статистических гипотез
Тема 8. Нормальное распределение и проверка гипотез
Тема 9. Анализ качественных данных
Тема 10. Анализ средних: критерий Стьюдента и дисперсионный анализ
Тема 11. Регрессионный анализ
Тема 12. Многофакторный анализ
Тема 13. Обобщенные линейные модели
Тема 14. Бутстреп и рандомизация

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа обучающихся включает в себя подготовку к контрольным вопросам и заданиям для текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведенным в п. 5.

Самостоятельная работа студентов направлена на закрепление теоретических сведений и отработку конкретных навыков работы в программной среде R, формирование культуры работы с данными.

Цель самостоятельной работы – подготовка современного компетентного специалиста и формирование способностей и навыков к непрерывному самообразованию и

профессиональному совершенствованию.

Виды самостоятельной работы студентов:

- изучение понятийного аппарата дисциплины;
- проработка тем дисциплины и решение задач;
- работа с основной и дополнительной литературой;
- изучение сайтов по темам дисциплины в сети Интернет;
- самоподготовка к занятиям практического типа;
- подготовка к решению задач;
- подготовка к тестам;
- подготовка к экзамену.

Изучение понятийного аппарата дисциплины

Вся система индивидуальной самостоятельной работы должна быть подчинена усвоению понятийного аппарата, поскольку одной из важнейших задач подготовки современного грамотного специалиста является овладение и грамотное применение профессиональной терминологии. Лучшему усвоению и пониманию дисциплины помогут различные энциклопедии, словари, справочники и другие материалы, указанные в списке литературы.

Работа над основной и дополнительной литературой

Изучение рекомендованной литературы следует начинать с учебников и учебных пособий, затем переходить к нормативно-правовым актам, научным монографиям и материалам периодических изданий. Конспектирование – одна из основных форм самостоятельного труда, требующая от студента активно работать с учебной литературой и не ограничиваться конспектом лекций.

Студент должен уметь самостоятельно подбирать необходимую для учебной и научной работы литературу. При этом следует обращаться к предметным каталогам и библиографическим справочникам, которые имеются в библиотеках.

Для аккумуляции информации по изучаемым темам рекомендуется формировать личный архив, а также каталог используемых источников. При этом если уже на первых курсах обучения студент определяет для себя наиболее интересные сферы для изучения, то подобная работа будет весьма продуктивной с точки зрения формирования библиографии для последующего написания дипломного проекта на выпускном курсе.

Изучение сайтов по темам дисциплины в сети Интернет

Ресурсы Интернет являются одним из альтернативных источников быстрого поиска требуемой информации. Их использование возможно для получения основных и дополнительных сведений по изучаемым материалам.

Самостоятельная работа студента при подготовке к экзамену

Контроль выступает формой обратной связи и предусматривает оценку успеваемости студентов и разработку мер по дальнейшему повышению качества подготовки современных специалистов.

В начале семестра рекомендуется внимательно изучить перечень вопросов к экзамену по данной дисциплине, а также использовать в процессе обучения программу, другие методические материалы, разработанные кафедрой по данной дисциплине. Это позволит в процессе изучения тем сформировать более правильное и обобщенное видение студентом существа того или иного вопроса за счет:

- а) уточняющих вопросов преподавателю;

- б) самостоятельного уточнения вопросов на смежных дисциплинах;
- в) углубленного изучения вопросов темы по учебным пособиям.

После изучения соответствующей тематики рекомендуется проверить наличие и формулировки вопроса по этой теме в перечне вопросов к экзамену, а также попытаться изложить ответ на этот вопрос. Если возникают сложности при раскрытии материала, следует вновь обратиться к лекционному материалу, материалам практических занятий, уточнить терминологический аппарат темы, а также проконсультироваться с преподавателем.

5. Фонд оценочных средств для текущего контроля успеваемости и промежуточной аттестации по дисциплине (модулю)

5.1 Типовые задания, необходимые для оценки результатов обучения при проведении текущего контроля успеваемости с указанием критериев их оценивания:

5.1.1 Типовые задания (оценочное средство - Тест) для оценки сформированности компетенции УК-1:

1. Число братьев и сестер у пациента является переменной следующего типа:

- а) альтернативная; б) категориальная; в) дискретная; г) непрерывная.

2. Какая из переменных является категориальной?

- а) рост; б) концентрация гемоглобина;
- в) тип опухоли; г) возраст.

3. Выборочная статистика – это:

- а) среднее и стандартное отклонение;
- б) точечная оценка генерального среднего;
- в) количественный показатель, рассчитываемый на основе выборочных данных;
- г) количественный показатель, используемый для проверки статистической гипотезы.

4. Множество, представляющее интерес для исследователя, осуществляющего статистический анализ, называется:

- а) выборка; б) генеральная совокупность; в) проба; г) описательная статистика.

5. 10 студентов писали тест и получили следующие баллы: 5, 7, 2, 1, 3, 4, 8, 8, 6, 6. Какова медиана этой выборки?

- а) 4.5; б) 5; в) 5.5; г) 6.

Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
превосходно	90–100% правильных ответов
отлично	80–90% правильных ответов
очень хорошо	70–80% правильных ответов
хорошо	60–70% правильных ответов
удовлетворительно	50–60% правильных ответов
неудовлетворительно	25–50% правильных ответов
плохо	менее 25% правильных ответов

5.1.2 Типовые задания (оценочное средство - Расчетно-графическое задание) для оценки сформированности компетенции ПК-21:

1. Вычислите следующее выражение:

$$1.06 + \frac{(3.85 - 2.77)^3}{4.56}$$

1. Вычислите следующее выражение:

$$\sin \frac{(3.68 + 2.93)^4}{1 + 1.71}$$

1. Создайте переменную a и присвойте ей значение 2.82. Создайте переменную b и присвойте ей значение 3.83.
2. Вычислите квадрат суммы значений a и b и присвойте результат переменной d.
3. Округлите d до третьего знака после запятой, результат присвойте той же переменной и выведите в консоль.
4. Вычислите остаток от деления d на 2 и присвойте результат переменной residual.
5. Выведите в консоль список имен всех переменных рабочего пространства.
6. Удалите переменную a из рабочего пространства.
7. Выведите в консоль текущую рабочую директорию.
8. Сохраните рабочее пространство в файл hw1.surname.rda, где surname - фамилия (предпочтительно - латиницей).
9. Сохраните итоговый скрипт в файл hw1.surname.r.
10. Отправьте два файла на адрес преподавателя в сообщении с темой group.surname.hw1, где group - номер группы.

Критерии оценивания (оценочное средство - Расчетно-графическое задание)

Оценка	Критерии оценивания
превосходно	90–100% правильно выполненных задач
отлично	80–90% правильно выполненных задач
очень хорошо	70–80% правильно выполненных задач
хорошо	60–70% правильно выполненных задач
удовлетворительно	50–60% правильно выполненных задач
неудовлетворительно	25–50% правильно выполненных задач
плохо	менее 25% правильно выполненных задач

5.2. Описание шкал оценивания результатов обучения по дисциплине при промежуточной аттестации

Шкала оценивания сформированности компетенций

Уровень сформированности компетенций (индикатора достижения компетенций)	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено		зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки	Минимально допустимый уровень знаний. Допущено много негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Ошибок нет.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки	Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном	Продемонстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания в полном объеме, но некоторые с	Продемонстрированы все основные умения. Решены все основные задачи с отдельными несущественными недочетами	Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов

2. Нулевая и альтернативная гипотезы формулируются относительно:

- а) выборочных данных; б) выборочных статистик;
в) параметров генеральной совокупности; г) параметров выборочной совокупности.

3. Выберите корректную нуль-гипотезу

- а) $H_0: \mu \neq 12$; б) $H_0: \mu = 12$; в) $H_0: \mu > 12$; г) $H_0: \mu = 12$.

4. Уровень значимости – это:

- а) вероятность того, что в условиях нуль-гипотезы будет получено такое же либо более экстремальное значение критериальной статистики;
- б) пороговое значение вероятности, выше которого нуль-гипотеза отвергается;
- в) максимально допустимая вероятность ошибки первого рода;
- г) максимально допустимая вероятность ошибки второго рода.

5. Какие события взаимоисключают друг друга и обязательно происходят в результате любого опыта:

- а) достоверные; б) элементарные; в) невозможные; г) несовместные.

6. Вероятность любого события всегда удовлетворяет следующему условию:

- а) она не меньше 0 и не больше 1; б) может принять любое значение;
- в) всегда строго больше 0; г) может принимать значения, меньшие 0.

Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
превосходно	90–100% правильных ответов
отлично	80–90% правильных ответов
очень хорошо	70–80% правильных ответов
хорошо	60–70% правильных ответов
удовлетворительно	50–60% правильных ответов
неудовлетворительно	25–50% правильных ответов
плохо	менее 25% правильных ответов

5.3.2 Типовые задания (оценочное средство - Расчетно-графическое задание) для оценки сформированности компетенции ПК-21

1. Вычислите следующее выражение:

1. Создайте следующий вектор `p` с помощью функции `rep()` и индексирования:

```
## (-5, 3, 5, -5, 3, 5, 3, 5, -5, 3, 5, -5, 3, 5, -5, 3, 5, -5, 3, 5)
```

1. Следующая команда генерирует выборку из 1000 случайных чисел, подчиняющихся стандартному нормальному распределению.

```
sam <- rnorm(1000)
```

Выполните ее. Вычислите число элементов вектора `sam`, больших 0.135.

1. Вычислите сумму 10 наибольших элементов вектора `sam`.
2. Постройте гистограмму значений элементов вектора `sam` на основе 50 интервалов с отображением не частот, а относительной плотности (используйте аргумент `freq = F`). Столбцы гистограммы должны быть залиты пшеничным цветом. На гистограмму наложите график плотности стандартного нормального распределения в виде штрих-пунктирной линии тройной толщины голубого цвета. Итоговый график снабдите легендой, в которой отображены залитая область, соответствующая гистограмме и линия, соответствующая плотности распределения. Указание. Для наложения плотности создайте последовательность чисел, соответствующих диапазону варьирования данных, рассчитайте для нее значения плотности с помощью функции `dnorm()` и сохраните их в отдельном векторе.

Описание данных к заданиям 6-9. В файле `apis.rda` содержится фрейм с данными по морфологической изменчивости пчел трех пород (переменная `breed`). Набор признаков включает шесть количественных (длина хоботка – `proboscis`, длина третьего тергита – `tergit.length`, ширина третьего тергита – `tergit.width`, длина крыла – `wing.length`, ширина крыла – `wing.width`, кубитальный индекс – `cubital.index`) и один качественный (дискоидальное смещение – `discoidal.shift`). Помимо породной принадлежности указана также семейная принадлежность (переменная `family`).

1. Загрузите данные по пчелам. Выберите в новый фрейм данные по пчелам породы *Apis mellifera caucasica*.
2. Рассчитайте медиану, стандартное отклонение, дисперсию, а также тридцатый процентиль длины крыла пчел породы *Apis mellifera caucasica* и сохраните в новом векторе `'descr'` в указанном порядке.
3. Рассчитайте натуральную частоту и долю пчел породы *Apis mellifera carnica*, имеющих нейтральное дискоидальное смещение, и сохраните их в векторе `'disc'`.
4. Постройте параллельную диаграмму размахов длины хоботка для трех пород пчел в вертикальном исполнении. Отдельные диаграммы размахов должны отличаться цветом (цвет на ваше усмотрение). Оси и график должны быть подписаны.

Описание данных к заданиям 10-12. Группа пациентов, участвующих в изучении нового препарата, характеризуется нормальным распределением роста со средним 175.7 см и стандартным отклонением 30.47 см.

1. Рассчитайте долю пациентов, рост которых менее 165.24 либо более 186.59, и сохраните ее в переменной p10.
2. Рассчитайте такое значение роста, что 55 % пациентов имеют более высокий рост, и сохраните его в переменной p11.
3. Рассчитайте значения роста, ограничивающие диапазон, за пределами которого находится 21 % пациентов по данному признаку, и сохраните их в переменной p12.

Критерии оценивания (оценочное средство - Расчетно-графическое задание)

Оценка	Критерии оценивания
превосходно	90–100% правильно выполненных задач
отлично	80–90% правильно выполненных задач
очень хорошо	70–80% правильно выполненных задач
хорошо	60–70% правильно выполненных задач
удовлетворительно	50–60% правильно выполненных задач
неудовлетворительно	25–50% правильно выполненных задач
плохо	менее 25% правильно выполненных задач

6. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Основная литература:

1. Анализ данных : учебник / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. - Москва : Юрайт, 2023. - 490 с. - (Высшее образование). - ISBN 978-5-534-00616-2. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847079&idb=0>.
2. Энатская Н. Ю. Математическая статистика и случайные процессы : учебное пособие / Н. Ю. Энатская. - Москва : Юрайт, 2023. - 201 с. - (Высшее образование). - ISBN 978-5-9916-9808-5. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=848905&idb=0>.

Дополнительная литература:

1. Гателюк О. В. Практикум по теории вероятностей и математической статистике : Учебное пособие для вузов / Гателюк О. В., Манюкова Н. В.; Манюкова Н. В. - Санкт-Петербург : Лань, 2022. - 132 с. - Книга из коллекции Лань - Математика. - ISBN 978-5-8114-9842-0., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=804174&idb=0>.

2. Прохоров Ю. В. Лекции по теории вероятностей и математической статистике : учебник и практикум / Ю. В. Прохоров, Л. С. Пономаренко. - 3-е изд. ; испр. и доп. - Москва : Юрайт, 2023. - 219 с. - (Высшее образование). - ISBN 978-5-534-10807-1. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=843646&idb=0>.
3. Горленко О. А. Дисперсионный анализ экспериментальных данных : учебное пособие / О. А. Горленко, Н. М. Борбаць, Т. П. Можаяева. - 2-е изд. ; испр. и доп. - Москва : Юрайт, 2023. - 132 с. - (Высшее образование). - ISBN 978-5-534-14677-6. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=840028&idb=0>.

Программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

1. <http://www.r-project.org/> – R Project: The R Project for Statistical Computing – язык программирования и среда разработки для статистических вычислений.
2. <http://cran.r-project.org/> – CRAN: The Comprehensive R Archive – архив пакетов для расширения языка программирования R.
3. <http://www.rstudio.com/> – R Studio – среда разработки для языка программирования R.
4. <http://www.statsoft.ru/home/textbook/> – электронный учебник по статистике и планированию эксперимента.
5. <http://r-analytics.blogspot.ru/> – R: Анализ и визуализация данных.

7. Материально-техническое обеспечение дисциплины (модуля)

Учебные аудитории для проведения учебных занятий, предусмотренных образовательной программой, оснащены мультимедийным оборудованием (проектор, экран), техническими средствами обучения, компьютерами.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ОС ННГУ по направлению подготовки/специальности 05.03.06 - Экология и природопользование.

Автор(ы): Якимов Василий Николаевич, доктор биологических наук, доцент.

Заведующий кафедрой: Якимов Василий Николаевич, доктор биологических наук.

Программа одобрена на заседании методической комиссии от 05.12.2023, протокол № 2.