

ФГБОУ ВО Новосибирский ГАУ
Кафедра прикладной биоинформатики

Рег. № ПБ.04-19

«10» 07 2024 г.



ФГОС 2017 г.
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.ДВ.02.01 Биометрический анализ на языке R

Шифр и наименование дисциплины

36.04.02 Зоотехния

Код и наименование направления подготовки

Прикладная биоинформатика

Направленность (профиль)

Курс: 1
Институт цифровых технологий

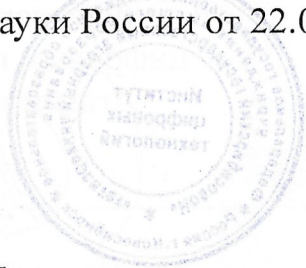
Семестр: 2
очная
очная, заочная, очно-заочная

Объем дисциплины (модуля)

Вид занятий	Объем занятий [зачетных ед./часов]			Семестр
	очная	заочная	очно-заочная	
Общая трудоемкость по учебному плану	5/180	—	—	2
В том числе,		—	—	
Контактная работа	52	—	—	
Занятия лекционного типа	16	—	—	
Занятия семинарского типа	36	—	—	
Самостоятельная работа, всего	128	—	—	
В том числе:				
Курсовой проект (курсовая работа)	—	—	—	
Контрольная работа / реферат / РГР	КР	—	—	2
Форма контроля				
Форма контроля экзамен / зачет / зачет с оценкой	3(О)	—	—	2

Новосибирск 2024

Рабочая программа составлена на основании требований Федерального государственного образовательного стандарта высшего образования – магистратура по направлению подготовки 36.04.02 Зоотехния, утвержденного приказом Минобрнауки России от 22.09.2017 №973.



Программу разработали:

Ст. преподаватель

(должность)

подпись

Петров А.Ф.

ФИО

Доцент

(должность)

подпись

Шатохин К.С.

ФИО

1. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с результатами освоения образовательной программы

Дисциплина «Биометрический анализ на языке R» в соответствии с требованиями ФГОС ВО и с учетом ПООП (при наличии) направлена на формирование компетенций, представленных в табл. 1.

Таблица 1. Связь результатов обучения с приобретаемыми компетенциями

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Запланированные результаты обучения
ПК-3 Способен к управлению производственной деятельностью в организации в соответствии с перспективным и текущим планами развития животноводства	ИПК-3.2: Использует информационные технологии в животноводстве при управлении технологическими процессами	знать: программные инструменты принятия решений для проведения генетической оценки уметь: пользоваться инструментами линейного программирования в R владеть: навыками моделирования и программирования с целью внедрения цифровых решений в технологические цепочки и процессы
	ИПК-3.3. Способен составлять алгоритмы, писать и отлаживать коды, применять конструкции языка программирования на практике	знать: основы матричной алгебры и положения о регрессионном анализе уметь: использовать методы программирования на языке R применительно к имеющимся исходным данным владеть: способностью решать нестандартные задачи в области зоотехнии
ПК-5 Способен обосновывать и использовать методы управления качеством продукции животноводства	ИПК-5.1 Обосновывает использование методов управления качеством продукции животноводства	знать: основные методы и стандарты управления качеством продукции животноводства уметь: применять методы управления качеством на практике, анализируя данные и результаты исследований, чтобы выявлять проблемы и разрабатывать стратегии для повышения качества продукции владеть: навыками разработки и внедрения систем управления качеством в процессе производства животноводческой продукции, используя инструменты биоинформатики для мониторинга и анализа данных о качестве.

2. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Биометрический анализ на языке R» относится к части, формируемой участниками образовательных отношений.

Данная дисциплина опирается на курсы дисциплин: «Проектирование баз данных», «Прикладное программирование на языке R».

3. Содержание дисциплины (модуля)

Распределение часов по темам и видам занятий представляется в таблице 2 по очной форме обучения:

Таблица 2. Очная форма

№ п/п	Наименование разделов и тем	Количество часов				Формируе- мые компетенции
		Лекции (Л)	Вид занятия (ЛР)	Самост. работа (СР)	Всего по теме	
1	2	3	4	5	6	7
1	Описательная статистика					ПК-3; ПК-5
1.1	Оценка выборочных параметров	1	3	12	16	
1.2	Распределения данных и выборок	1	3	12	16	
1.3	Критерии согласия	1	4	12	17	
2	Моделирование					
2.1	Линейные модели в дисперсионном анализе	2	2	13	17	
2.2	Регрессия и предсказания	3	6	13	22	
2.3	Обобщённые линейные модели	3	6	12	21	
3	Кластеризация					
3.1	Классификация	2	6	15	23	
3.2	Машинное обучение	3	6	15	24	
	Контрольная работа			12	12	
	Зачет с оценкой			12	12	
Итого		16	36	128	180	

Учебная деятельность состоит из лекций, лабораторных занятий, самостоятельной работы и контрольной работы

3.1.Содержание отдельных разделов и тем

Раздел 1. Описательная статистика

Тема 1.1 Оценка выборочных параметров

- Разведочный анализ, группировка данных, анализ с учетом пропущенных значений в таблицах данных, типы данных
- Средние величины, показатели вариации, асимметрия и эксцесс, стандартная ошибка, бутстрэп
- Анализ выбросов, доверительные интервалы, усечение данных
- Визуализация данных `ggplot()`, `hist()`

Тема 1.2 Распределения данных и выборки

- Основы теории распределений
- Случайный отбор и смещенная выборка
- Виды распределений, оценка характера распределений, критерии нормальности

Тема 1.3 Критерии согласия

- Оценка различий между двумя выборками для признаков с непрерывной и дискретной изменчивостью. Критерий Стьюдента, одновыборочный t-критерий, поправка Уэлча, сравнение парных выборок, критерий Манну-Уитни
- множественность испытаний, поправка Бонферони
- оценка различий частот, критерий хи-квадрат
- оценка разности распределений выборок, критерий Колмогорова-Смирнова
- оценка мощности критерия, рандомизация, бутстреп
- корреляционный анализ
- критерии проверки равенства дисперсий

Раздел 2. Моделирование

Тема 2.1. Дисперсионный анализ

- Влияние группирующего фактора на изменчивость зависимой переменной. Дисперсионный анализ, критерий Краскела-Уоллиса
- Двух- и многофакторный дисперсионный анализ
- Моделирование

Тема 2.2 Регрессия и предсказания

- Проверка допущений: диагностика регрессии
- Простая линейная регрессия, метод наименьших квадратов
- Множественная линейная регрессия
- Модели регрессии при различных видах функции потерь
- Внутренние и внешние критерии качества моделей

- Кроссвалидация
- Критерии выбора моделей оптимальной сложности
- Виды нелинейной регрессии, полиномиальная регрессия
- Интерпретация уравнений регрессии
- Факторные переменные регрессии
- Проверка качества прогноза модели

Тема 2.3 Обобщённые линейные модели

- Модели сглаживания
- Обобщенные модели регрессии
- Модели пробит- и логит-регрессии
- Оценка выживаемости
- Ковариационный анализ
- Модели со смешанными эффектами для иерархически организованных данных
- Индуктивные модели (метод группового учета аргументов)

Раздел 3. Кластеризация

Тема 3.1 Классификация

- Байесовский алгоритм
- Дискриминантный анализ
- Логистическая регрессия
- Метод главных компонент
- Факторный анализ

Тема 3.2 Машинное обучение

- К ближайших соседей
- Древовидные модели
- Бэггинг и случайный лес
- Бустинг
- Анализ главных компонент
- Иерархическая кластеризация
- Переобучение моделей

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

4.1. Список основной литературы

1. Методы обработки экспериментальных данных и математического моделирования процессов: учебное пособие / Новосибирский гос. аграр. ун-т; сост.: Е.В. Камалдинов, С.Г. Куликова, М.Л. Кочнева, К.Н. Нарожных. - 3-е изд., доп. - Новосибирск: Изд-во НГАУ, 2021. - 153 с. - Электронная программа

(визуальная). Электронные данные: электронные. - URL: <https://edubiotech.ru/file/1995401>

2. Митина, О. А. Языки программирования для статистической обработки данных (R) : учебное пособие/ О. А. Митина. — Москва: РТУ МИРЭА, 2020. — 191 с. — Текст: электронный// Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/163912>.

4.2. Список дополнительной литературы

1. Гришин, В. А. Основы программирования на языке R / В. А. Гришин. — Нижний Новгород: ННГУ им. Н. И. Лобачевского, 2021. — 67 с. — Текст: электронный// Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/191498>

2. Перекатов, А. С. Статистическая обработка экспериментальных данных. Полный факторный эксперимент в языке R: учебное пособие / А. С. Перекатов, М. Б. Никифоров. — Рязань: РГРТУ, 2019. — 48 с.— Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/168309>

4.3. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Таблица 3. Перечень информационных ресурсов

№ п/п	Наименование	Адрес
1.	Биологическая статистика с "R" (KamaldinovE.V.)	http://rbiostats.blogspot.com/
2.	Официальный перечень рекомендуемых к ознакомлению статей от «TheRDevelopmentCoreTeam. TheRJournal»	http://www.r-project.org/doc/Rnews/bib/Rnewsbib.html
3.	Графический метод решения задач линейного программирования	http://ru.wikipedia.org/wiki/Графический_метод_решения_задачи_линейного_программирования
4.	Информационный портал по языкам программирования и машинному обучению «Statology»	https://www.statology.org/
5.	Официальный ресурс документации к функциям и пакетам языка программирования R	https://www.rdocumentation.org/

4.4. Методические указания для обучающихся по освоению дисциплины (модулю) и самостоятельной работы

Методы обработки экспериментальных данных и математического моделирования процессов: учебное пособие / Новосибирский гос. аграр. ун-т; сост.: Е.В. Камалдинов, С.Г. Куликова, М.Л. Кочнева, К.Н. Нарожных. - 3-е изд., доп. - Новосибирск: Изд-во НГАУ, 2021. - 153 с. - Электронная программа (визуальная). Электронные данные: электронные. - URL: <https://edubiotech.ru/file/1995401>

4.5. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень лицензионного и свободно распространяемого программного обеспечения и информационных справочных систем, наглядных пособий

1. Применение интерактивной доски и проектора для демонстрации презентаций и программного кода.
2. Применение сенсорного дисплея (не менее 100") для вывода нескольких пользовательских экранов одновременно от обучающихся для анализа, обсуждения и сравнения подходов к решению задач.
3. Применение вычислительного центра для демонстрации решения задач, требующих распараллеливания и применения алгоритмов BigData.

Таблица 4. Перечень лицензионного и свободно распространяемого программного обеспечения

№ п/п	Наименование	Тип лицензии или правообладатель
1.	Astra Linux	Русбитех-Астра
2.	LibreOffice	The Document Foundation
3.	R	Свободно-распространяемая
4.	RStudio Free	Свободно-распространяемая
5.	Notepad++	Свободно-распространяемая

Таблица 5. Перечень плакатов (по темам), карт, стендов, макетов, презентаций, фильмов и т.д.

№ п/п	Тип	Наименование	Примечание
1.	Видеофильм	<i>Элементы программирования в R. Описательные статистики</i>	<i>2 час. 21 мин.</i>

5. Описание материально-технической базы

Таблица 6. Перечень используемых помещений:

№ аудитории	Тип аудитории	Перечень оборудования
НК-302	Компьютерный класс: аудитория для занятий семинарского типа, самостоятельной работы, дипломного и курсового проектирования (выполнения курсовых работ).	15 персональных компьютеров, видеопроектор, интерактивная доска, доска учебная, учебная мебель.
НК-307	Компьютерный класс: аудитория для занятий семинарского типа, самостоятельной работы, дипломного и курсового проектирования (выполнения курсовых работ)	11 персональных компьютеров, видеопроектор, проекционный экран, доска учебная, веб-камера колонки, учебная мебель.

6. Порядок аттестации студентов по дисциплине

Для аттестации студентов по дисциплине используется традиционная система контроля и оценки успеваемости обучающихся.

Условием допуска к экзамену является посещение не менее 50% академических часов в рамках контактной работы. Для получения оценки «отлично» необходимо правильно решить практическую задачу с использованием ЭВМ и ответить на два теоретических вопроса, «хорошо» - решить практическую задачу и ответить на один теоретический вопрос, «удовлетворительно» - решить практическую задачу. При отсутствии решения практической задачи выставляется отметка «удовлетворительно».

Промежуточный контроль проводится с целью установления уровня освоения материала по самостоятельным разделам в виде контрольных работ и выполнения заданий на семинарских занятиях.

Итоговый контроль — оценка уровня освоения дисциплины по окончании её изучения в форме экзамена в устной форме.

Описание шкалы оценивания:

Критерии оценивания устного ответа на вопросы к зачёту с оценкой:

«5» (отлично) — дан полный, развёрнутый ответ на поставленный вопрос, показана совокупность осознанных знаний об объекте, доказательно раскрыты основные положения темы; в ответе прослеживается чёткая структура, логическая последовательность, отражающая сущность раскрываемых понятий, теорий, явлений. Знание об объекте демонстрируется на фоне понимания его в системе данной науки, и междисциплинарных связей. Ответ изложен литературным языком в терминах науки, показана способность быстро реагировать на уточняющие вопросы;

«4» (хорошо) — дан полный, развёрнутый ответ на поставленный вопрос, показано умение выделить существенные и несущественные признаки,

причинно-следственные связи. Ответ чётко структурирован, логичен, изложен в терминах науки, Однако допущены незначительные ошибки ли недочёты, исправленные студентом с помощью «наводящих» вопросов;

«3» (удовлетворительно) — дан неполный ответ, логика и последовательность изложения имеют существенные нарушения. Допущены грубые ошибки при определении сущности раскрываемых понятий, теорий явлений, в следствии непонимания студентом их существенных и несущественных признаков и связей. В ответе отсутствуют выводы. Умение раскрыть конкретные проявления обобщённых знаний не показано. Речевое оформление требует поправок, коррекций.

«2» (неудовлетворительно) — студент демонстрирует незнание теоретических основ предмета, не умеет выделять аргументированные выводы и приводить примеры, показывает слабое владение монологической речью, не владеет терминологией, проявляет отсутствие логичности и последовательности изложения, делает ошибки, которые не может исправить даже при коррекции преподавателем, отказывается отвечать на занятия.

7. Согласование рабочей программы

Соответствует учебному плану, утвержденному Ученым советом ФГБОУ ВО Новосибирского ГАУ, протокол от 27.06.2024 № 6

Рабочая программа обсуждена и утверждена на заседании кафедры прикладной биоинформатики протокол от 28.06.2024 № 1

Заведующий кафедрой
(должность)


подпись

Камалдинов Е.В.
ФИО

Председатель учебно-методического совета
(должность)


подпись

Андронов А.Ю.
ФИО

Рабочая программа обсуждена и соответствует учебному плану, утвержденному Ученым советом ФГБОУ ВО Новосибирского ГАУ, протокол от « » 20 г. № _____

Изменений не требуется/изменения внесены в раздел(-ы): _____
нужное подчеркнуть

Председатель учебно-методического совета

Рабочая программа обсуждена и соответствует учебному плану, утвержденному
Ученым советом ФГБОУ ВО Новосибирского ГАУ, протокол от «____» _____
20__ г. №_____

Изменений не требуется/изменения внесены в раздел(-ы): _____
нужное подчеркнуть

Председатель методического совета

(должность)

подпись

ФИО