

ФГБОУ ВО Университет биотехнологий
Кафедра прикладной биоинформатики

Рег. № МТИ 26-20

«20» января 2026г.



И.о. директора Института
цифровых технологий
Агафонов О.В.

ФГОС 2017 г.
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

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

Шифр и наименование дисциплины

09.04.03 Прикладная информатика

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	54	—	—	
Занятия лекционного типа	16	—	—	
Занятия семинарского типа	38	—	—	
Самостоятельная работа, всего	126	—	—	
В том числе:				
Курсовой проект (курсовая работа)	—	—	—	
Контрольная работа / реферат / РГР	КР	—	—	2
Форма контроля				
Форма контроля экзамен / зачет / зачет с оценкой	3(0)	—	—	2

Новосибирск 2026

Рабочая программа составлена на основании требований Федерального государственного образовательного стандарта высшего образования – магистратура по направлению подготовки 09.04.03 Прикладная информатика, утвержденного приказом Минобрнауки России от 19.09.2017 № 916.

Программу разработали:

Доцент, к.б.н.



Шатохин К.С.

Ст. преподаватель



Краснов С.Д.

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 проводить статистический анализ биологических данных с использованием специализированных программных пакетов и интерпретировать полученные результаты для решения производственных задач	ИПК-3.2: Самостоятельно выбирает наиболее подходящие статистические методы для анализа конкретных типов биологических данных и обосновывает свой выбор применительно к задачам в животноводстве	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. Интерпретирует результаты статистического анализа биологических данных, формулирует выводы и рекомендации для решения производственных задач в животноводстве, основываясь на статистически значимых показателях.	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

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	Базовая статистика					ПК-3
1.1	Описательная статистика	1	3	12	16	
1.2	Анализ распределения.	1	3	12	16	
1.3	Критерии согласия	1	5	11	17	
2	Корреляция, регрессия, дисперсионный анализ					
2.1	Дисперсионный анализ	3	6	14	23	
2.2	Корреляционный анализ	2	5	9	16	
2.3	Регрессия	3	6	14	23	
3	Кластеризация и метод главных компонент					
3.1	Кластерный анализ	2	5	15	22	
3.2	Метод главных компонент	3	5	15	23	
	Контрольная работа			12	12	
	Зачет с оценкой			12	12	
Итого		16	38	126	180	

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

3.1.Содержание отдельных разделов и тем

Раздел 1. Базовая статистика

Тема 1.1 Описательная статистика

Классификация признаков. Средние величины: средняя арифметическая, средняя взвешанная, усечённая средняя, средняя геометрическая, средняя квадратическая. Методы их вычисления в R, применение в животноводстве. Ошибки средних величин. Описательная статистика признаков с альтернативной изменчивостью: встречаемость, ошибка встречаемости, дисперсия для качественных признаков. Показатели вариации: дисперсия, стандартное отклонение, коэффициент вариации. Комплексные функции расчёта описательной статистики, пакет psych.

Тема 1.2 Анализ распределения

Гауссовское или нормальное распределение. Ассиметрия и эксцесс. Биномиальное распределение. Другие виды распределения. Коэффициент трансгрессии. Визуализация соответствия фактического и нормального распределения при помощи базовых функций и библиотеки ggplot2.

Тема 1.3 Критерии согласия

Нулевая и альтернативная гипотеза. Общие принципы критериев согласия, уровень статистической значимости. Оценка соответствия закону нормального распределения. Понятие параметрических и непараметрических критериев. Критерии Шапиро-Уилка и Андерсона-Дарлинга, библиотека nortest. Сравнение распределения двух выборок, критерии хи-квадрат и Колмогорова-Смирнова. Оценка достоверности различий между двумя выборками, критерии Стьюдента и Манну-Уитни. Оценка равенства дисперсий в двух группах.

Раздел 2. Корреляция, регрессия, дисперсионный анализ

Тема 2.1. Дисперсионный анализ

От попарного сравнения к эффекту группирующего признака. Поправка Бонферрони при множественных сравнениях. Сущность и ограничения метода однофакторного дисперсионного анализа, межгрупповая и внутригрупповая дисперсия, модель дисперсионного анализа в R и её интерпретация. Критерий Краскела-Уоллеса как непараметрический метод анализа изменчивости зависимой переменной от градаций группирующего фактора. Апостериорные методы, поправка Тьюки, метод Коновера-Уимана.

Тема 2.2 Корреляционный анализ

Сущность корреляции. Метод Пирсона как способ оценки сопряженности количественных признаков неотличимых от нормального распределения. Ранговые коэффициенты корреляции Спирмена и Кендалла. Корреляционные

матрицы с использованием пакетов *rcorr* и *corrplot*. Корреляция между качественными признаками, тетракорический и поликорический коэффициенты, библиотека *polychor*.

Тема 2.3 Регрессия

Понятие простой линейной регрессии, структура модели с одним предиктором, построение модели в R, функция `lm()`. Метод наименьших квадратов (МНК). Анализ результатов, вывод сводки по модели, коэффициенты (Estimate, Std. Error, p-value), R^2 (коэффициент детерминации), F-статистика (общая значимость модели). Понятие остатков модели. Диагностика модели, проверка допущений: нормальность остатков, гомоскедастичность (в частности тест Бреуша-Пагана (`lmtest::bptest()`), тест на автокорреляцию остатков (в частности тест Дарбина-Уотсона (`lmtest::dwtest()`), оценка остатков модели на наличие выбросов, дистанции Кука, `plot(model, which = 4 # визуализация дистанций Кука)`, `qqnorm(residuals(model))` # график квантиль-квантиль, показывающий распределение остатков модели). Доверительный интервал для регрессионной модели. Использование регрессионной модели как инструмента предсказания, оценка качества предсказания. Оценка качества модели критериями Акаике и Шварца. Визуализация регрессии с использованием библиотеки *ggplot2*.

Раздел 3. Классификация и метод главных компонент

Тема 3.1. Кластерный анализ

Введение в анализ данных: задачи описания, классификации, кластеризации. Определение кластеризации: цели (обнаружение структуры, группировка, сжатие данных). Типичные приложения (сегментация клиентов, биология, социальные науки, анализ изображений, обнаружение аномалий). Основные термины: объекты, признаки, метрика расстояния/сходства, кластер, центроид. Тест на кластеризуемость выборок, статистика Хопкинса. Дистанции Эвклида, Манхэттена, Миньковского, Чебышева, Канберра, Махаланобиса. Методы кластеризации. Оценка правдоподобности кластерного анализа, метод внешней валидации. Иерархическая Кластеризация (Hierarchical). Принцип: агломеративный (снизу вверх) vs. дивизимный (сверху вниз) подходы. Связь между кластерами: одиночная, полная, средняя, Уорда. Дендрограмма. Определение числа кластеров по дендрограмме. Метод К-средних (K-Means).

Тема 3.2. Метод главных компонент

Концепция метода главных компонент (PCA), понятие «Главная компонента». Линейное преобразование данных для снижения размерности. PCA как средство борьбы с мультиколлинеарностью. Декорреляция признаков, Визуализация многомерных данных. Ковариационная матрица. Собственные векторы (направления компонент). Собственные значения (объяснённая дисперсия). Подавление шума Библиотеки: *factoextra*, *ggplot2*

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

4.1. Список основной литературы

- ✓1. Золотарюк, А. В. Язык и среда программирования R : учебное пособие / А.В. Золотарюк. — Москва : ИНФРА-М, 2026. — 162 с. — (Высшее образование). — DOI 10.12737/textbook_5b8fdb0bd795c4.69435980. - ISBN 978-5-16-018723-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2221665>.
- ✓2. Статистика : учебное пособие для вузов / К. Н. Горпинченко, Е. В. Кремьянская, А. М. Ляховецкий [и др.]. — 2-е изд., стер. — Санкт-Петербург : Лань, 2024. — 156 с. — ISBN 978-5-507-47762-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/415379>

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. 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>.

4.3. 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.	ANOVA in R	https://statsandr.com/blog/anova-in-r/
3.	R Programming for Data Science	https://www.geeksforgeeks.org/r-programming-for-data-science/
4.	R - Statistics	https://www.geeksforgeeks.org/r-language/r-statistics/
5.	Cluster Analysis	https://rpubs.com/MdAhsanul/cluster_analysis

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

Биометрические методы обработки экспериментальных данных и моделирования в сельском хозяйстве и биологии: учебное пособие. Е.В. Камалдинов, С.Г. Куликова, М.Л. Кочнева, К.Н. Нарожных, В.В. Гарт, А.Ф. Петров; Новосиб. гос. аграр. ун-т. – Новосибирск, 2024. – 165 с.

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.	Презентации	В соответствии с темами лекций	Количество слайдов различное в каждой презентации

5. Описание материально-технической базы

Таблица 6. Перечень используемых помещений:

№ аудитории	Тип аудитории	Перечень оборудования
НК-302	Компьютерный класс: аудитория для групповых и индивидуальных	Компьютер - 15 шт.; проектор; доска ученическая; мебель учебная - 16 шт.;

	консультаций, дипломного и курсового проектирования (выполнения курсовых работ), занятий семинарского типа, текущего контроля, промежуточной аттестации, самостоятельной работы.	доска интерактивная; веб-камера с микрофоном; колонки акустические.
НК-307	Компьютерный класс: аудитория для групповых и индивидуальных консультаций, дипломного и курсового проектирования (выполнения курсовых работ), занятий семинарского типа, текущего контроля, промежуточной аттестации, самостоятельной работы.	Компьютер - 11 шт.; проектор; доска ученическая; экран проекционный; мебель учебная - 12 шт.; веб-камера с микрофоном; колонки акустические

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 Университет биотехнологий, протокол от 25.12.2025 № 10.

Рабочая программа обсуждена и утверждена на заседании кафедры протокол от «14» января 2026 г. № 5

Заведующий кафедрой
(должность)


ПОДПИСЬ

Камалдинов Е.В.
ФИО

Председатель
методического совета
(должность)

учебно-

Д. 42 -

ПОДПИСЬ

Чечушкова М.А.

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

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

Председатель учебно-методического совета
(должность)

ПОДПИСЬ

ФИО