

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Акчурин Сергей Владимирович

Должность: Заместитель директора Института зоотехнии и биологии

Дата подписания: 24.02.2025 10:55:41

Уникальный программный ключ:

7abcc100773ae7c9cceb4a7a083ff3fbbf160d2a



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –

МСХА имени К.А. ТИМИРЯЗЕВА»

(ФГБОУ ВО РГАУ - МСХА имени К.А. Тимирязева)

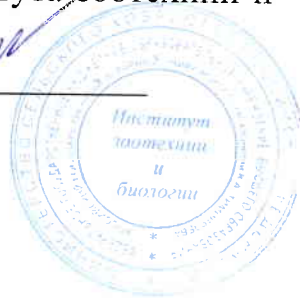
Институт садоводства и ландшафтной архитектуры
Кафедра ботаники, селекции и семеноводства садовых растений

УТВЕРЖДАЮ:

И.о. директора института зоотехнии и
биологии

Акчурин С.В.

“30” августа 2024 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.04 Язык R в биологических исследованиях

для подготовки магистров

ФГОС ВО

Направление 06.04.01 Биология

Направленность (программа) «Биоинформатика»

Курс: 1

Семестр: 2

Форма обучения: очная

Год начала подготовки: 2024

Регистрационный номер _____

Москва, 2024

Разработчики(и): С.Г. Монахос, д.с.-х.н., профессор
Д.Д. Лисовая, ассистент



«28» августа 2024 г.

Рецензент: Монахос Г.Ф., к.с.-х.н., ст.н.с.
(ФИО, ученая степень, ученое звание)

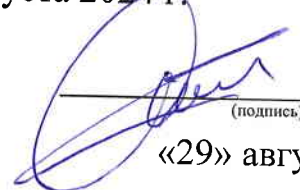

(подпись)

«29» августа 2024 г.

Программа составлена в соответствии с требованиями ФГОС ВО, по направлению подготовки 06.04.01 Биология и учебного плана.

Программа обсуждена на заседании кафедры ботаники, селекции и семеноводства садовых растений, протокол №9.1 от «29» августа 2024 г.

Зав. кафедрой С.Г. Монахос, д.с.-х.н., профессор
(ФИО, ученая степень, ученое звание)


(подпись)

«29» августа 2024 г.

Согласовано:

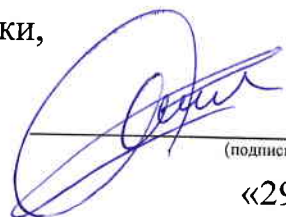
Председатель учебно-методической комиссии института
садоводства и ландшафтной архитектуры
Маланкина Е.Л., д.с.-х.н., профессор
(ФИО, ученая степень, ученое звание)


(подпись)

Протокол №7 от «29» августа

«29» августа 2024 г.

Заведующий выпускающей кафедрой ботаники,
селекции и семеноводства садовых растений
С.Г. Монахос, д.с.-х.н., профессор
(ФИО, ученая степень, ученое звание)


(подпись)

«29» августа 2024 г.

Зав. Отделом комплектования ЦНБ /


(подпись)

Разработчики(и): С.Г. Монахос, д.с.-х.н., профессор
Д.Д. Лисовая, ассистент

«28» августа 2024 г.

Рецензент: Монахос Г.Ф., к.с.-х.н., ст.н.с.
(ФИО, ученая степень, ученое звание)

(подпись)

«29» августа 2024 г.

Программа составлена в соответствии с требованиями ФГОС ВО, по направлению подготовки 06.04.01 Биология и учебного плана.

Программа обсуждена на заседании кафедры ботаники, селекции и семеноводства садовых растений, протокол №9.1 от «29» августа 2024 г.

Зав. кафедрой С.Г. Монахос, д.с.-х.н., профессор
(ФИО, ученая степень, ученое звание)

(подпись)

«29» августа 2024 г.

Согласовано:

Председатель учебно-методической комиссии института
садоводства и ландшафтной архитектуры
Маланкина Е.Л., д.с.-х.н., профессор
(ФИО, ученая степень, ученое звание)

(подпись)

Протокол №7 от «29» августа
«29» августа 2024 г.

Заведующий выпускающей кафедрой ботаники,
селекции и семеноводства садовых растений
С.Г. Монахос, д.с.-х.н., профессор
(ФИО, ученая степень, ученое звание)

(подпись)

«29» августа 2024 г.

Зав. Отделом комплектования ЦНБ

(подпись)

СОДЕРЖАНИЕ

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ	4
2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ	4
3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ «ЯЗЫК R В БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЯХ», СООТВЕТСТВУЮЩИХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	5
4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ	5
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЕМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ ПО СЕМЕСТРАМ	5
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ	9
4.3 ПРАКТИЧЕСКИЕ ЗАНЯТИЯ	11
5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ	14
6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ	15
6.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности	15
7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	20
7.1 Основная литература	20
7.2 Дополнительная литература	20
8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	20
9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ	21
10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ	21
11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ СТУДЕНТАМ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ	21
Виды и формы отработки полученных знаний	22
12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ	22

АННОТАЦИЯ

рабочей программы учебной дисциплины Б1.В.04 «Язык R в биологических исследованиях» для подготовки магистров по направлению 06.04.01 Биология направленности «Биоинформатика»

Цель освоения дисциплины: Овладение навыками программирования на языке R, знаниями об основных пакетах статистического анализа данных и машинного обучения, и умением применять их для решения актуальных задач генетики, биотехнологии и селекции.

Место дисциплины в учебном плане: дисциплина включена в часть, формируемую участниками образовательных отношений, учебного плана по направлению подготовки 06.04.01 Биология

Требования к результатам освоения дисциплины: в результате освоения дисциплины формируются следующие компетенции: 2 профессиональные компетенции ПКос-2, ПКос-3 (ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3).

Краткое содержание дисциплины: Дисциплина "Язык R в биологических исследованиях" представляет собой курс, направленный на обучение студентов использованию программного языка R в биологических исследованиях. Курс охватывает основные концепции и методы анализа данных, включая обработку, визуализацию и статистический анализ биологических данных с использованием языка R.

Общая трудоемкость дисциплины: 144/4 (часы/зач. ед.)

Промежуточный контроль: экзамен

1. Цель освоения дисциплины

Цель данной дисциплины заключается в овладении программирования на языке R, включая синтаксис языка, операторы, базовые функции, импорт и сохранение информации, построение графиков и диаграмм; применении основных методов и критериев математической статистики для анализа молекулярно-биологических и биотехнологических данных, используя R; освоение методов машинного обучения (метод опорных векторов, искусственные нейронные сети, деревья решений, методы кластеризации данных и др.), реализованных в R, и их применение для анализа биологических данных.

2. Место дисциплины в учебном процессе

Дисциплина «Язык R в биологических исследованиях» включена в часть профессионального цикла, формируемую участниками образовательных отношений. Реализация в дисциплине «Язык R в биологических исследованиях» требований ФГОС ВО, ОПОП и Учебного плана по направлению 06.04.01 Биология для подготовки магистров направленности «Биоинформатика».

Предшествующими курсами, на которых непосредственно базируется дисциплина «Язык R в биологических исследованиях», являются «Популяционная

генетика», «Структурная и сравнительная геномика», «Молекулярная биология», «Язык R в биологических исследованиях», «Генерация и анализ омиксных данных».

Дисциплина «Язык R в биологических исследованиях» является основополагающей для изучения следующих дисциплин: «Современная селекция растений», «Современная селекция животных», «Гранскриптомика и протеомика».

Данная дисциплина представляет собой курс, который призван обучить студентов использованию программного языка R в биологических исследованиях. В рамках курса основное внимание уделяется основным концепциям и методам анализа данных, таким как обработка, визуализация и статистический анализ биологических данных с использованием языка R.

Рабочая программа дисциплины «Язык R в биологических исследованиях» для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.

3. Перечень планируемых результатов обучения по дисциплине «Язык R в биологических исследованиях», соотнесенных с планируемыми результатами освоения образовательной программы

Изучение данной учебной дисциплины направлено на формирование у обучающихся компетенций, представленных в таблице 1.

4. Структура и содержание дисциплины

4.1 Распределение трудоемкости дисциплины по видам работ по семестрам

Общая трудоемкость дисциплины составляет 4 зач.ед. (144 часов), их распределение по видам работ семестрам представлено в таблице 2.

Таблица 1

Требования к результатам освоения учебной дисциплины

№ п/п	Код компетенции	Содержание компетенции (или её части)	Индикаторы компетенций	В результате изучения учебной дисциплины обучающиеся должны:		
				знать	уметь	владеть
1.	ПКос-2	Способен самостоятельно в качестве руководителя или члена коллектива организовывать и управлять производственной и научно-исследовательской деятельностью в избранной и смежных предметных областях	ПКос-2.1 правила и методики анализа результатов научных исследований, специализированное программное обеспечение, способы обработки полученных эмпирических данных и их интерпретации	Синтаксис и основные функции языка R. Основные статистические методы и критерии, их реализацию в R. Знать основные методы машинного обучения и их реализацию в R	Использовать R для статистической обработки экспериментальных и клинических данных. Уметь интерпретировать полученные результаты	Написания простых программ, построения графиков и диаграмм в R. Владеть основными методами статистики, методами машинного обучения, навыками по их использованию с помощью среды R
			ПКос-2.2 анализировать получаемую полевою и лабораторную биологическую информацию с использованием современной вычислительной техники; систематизировать экспериментальные данные; обобщать полученные результаты в контексте ранее накопленных в науке знаний; получать новые данные	Основные понятия теории вероятностей и математической статистики, основные методы описания данных	Ориентироваться в основных алгоритмах статистической обработки данных для решения практических задач и графического представления результатов	Соответствующими приемами программирования, компьютерными методами обработки данных биологических исследований

				верные факты на основе наблюдений, опытов, научного анализа эмпирических данных; представлений результатов научных исследований; нести ответственность за качество выполняемых работ				
				ПКос-2.3 навыками подготовки научных публикаций, отчетов, обзоров, патентов и докладов, участия в организации научных семинаров и конференций; статистическими методами сравнения полученных экспериментальных данных и определения закономерностей, способностью формулировать выводы и практические рекомендации на основе репрезентативных и оригинальных результатов исследований	Принципы применения теории вероятностей и математической статистики для описания данных биологических исследований	Применять методы описания, статистического анализа и визуализации данных биологических исследований	Компьютерными средствами для подготовки, хранения и анализа данных биологических исследований	

ОЧНАЯ ФОРМА ОБУЧЕНИЯ

Таблица 2а

Распределение трудоёмкости дисциплины по видам работ по семестрам			
Вид учебной работы		Трудоёмкость	
		час.	
Общая трудоёмкость дисциплины по учебному плану		144	
1. Контактная работа:		40,4	
Аудиторная работа		38	
в том числе:			
лекции (Л)		14	
практические занятия (ПЗ)		24	
консультации перед экзаменом		2	
контактная работа на промежуточном контроле (КРА)		0,4	
2. Самостоятельная работа (СРС)		70	
(подготовка и повторение лекционного материала и материалов учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиумам и т.д.)		36,4	
Подготовка к экзамену (контроль)		33,6	
Вид промежуточного контроля:		Экзамен	

4.2 Содержание дисциплины

Таблица 3

Тематический план учебной дисциплины

Наименование разделов и тем дисциплины (укрупнено)	Всего	Аудиторная работа			Внеаудиторная работа
		Л	ПЗ/С	ПКР	работы СР
Раздел 1 Язык R в биологических исследованиях	144	14	24	2,4	70
Тема 1. Установка R и RStudio. Особенности интерфейса RStudio. Установка библиотек.	7,4	1	2	-	4,4
Тема 2. Форматы данных. Загрузка данных в R. Работа с текстовыми файлами в R.	8	1	2	-	5
Тема 3. Основы работы с базами данных.	8,5	2	2	-	4,5
Тема 4. Базовые графики в R. Визуализация количественных и качественных данных в R.	10,5	2	4	-	4,5
Тема 5. Корреляционный и регрессионный анализ в R. Множественная регрессия в R.	8,5	2	2	-	4,5
Тема 6. Управляющие конструкции в R. Циклы for и while. Функции в R.	10,5	2	4	-	4,5
Тема 7. Парсинг веб-страниц средствами R. Библиотека rvest.	10,5	2	4	-	4,5
Тема 8. Работа с API средствами R.	10,5	2	4	-	4,5
Консультация перед экзаменом	2	-	-	2	-

2. ПКос-3	Способен профессионально работать с исследовательскими и испытательными оборудованием (приборами и установками, специализированными пакетами прикладных программ) в избранной предметной области	ПКос-3.1 научно-методические основы и методы биоинформатики для решения производственных и научно-исследовательских задач в области растениеводства и животноводства	основные термины и понятия биоинформатики, современные концепции биоинформатики, объекты изучения биоинформатики, последовательности нуклеиновых кислот и аминокислот	находить, анализировать, обобщать и систематизировать научные данные, полученные в ходе лабораторных экспериментов, для постановки целей исследования и выбора оптимальных методов их достижения;	предложить расчётов оптимальных параметров проведения анализа, систематизации и интерпретации данных биологических объектов, и их коррелирование
	ПКос-3.2 проводить производственную деятельность в области биоинформатики и смежных дисциплин, самостоятельно использовать современные технологии для решения задач профессиональной деятельности	методы исследования биологических последовательностей растений, их описания, представления структуры и функции белков,	подбирать необходимые условия проведения научного анализа в зависимости от специфики поставленной задачи с применением методов биоинформатики	основными методами, способами и средствами получения, хранения, анализа и систематизации информации применять биологическим объектам	
	ПКос-3.3 современные технологии в области биоинформатики и геномики, применение при решении теоретических и практических задач в селекции растений и животных	особенности, возможности и ограничения специализированных баз данных и способы работы с ними	работать со специализированными серверами и различными базами данных	навыками использования программных средств и работы в компьютерных сетях, использования ресурсов Интернета применительно к биологическим объектам	

Наименование разделов и тем дисциплины (укрупненно)	Всего	Аудиторная работа			Внеаудиторная работа СР
		Л	ПЗ/С	ПКР	
Контактная работа на промежуточном контроле (КРА)	0,4	-	-	0,4	-
Подготовка к экзамену	33,6	-	-	-	33,6
Итого по дисциплине	144	14	24	2,4	70

Раздел 1 Язык R в биологических исследованиях

Тема 1. Установка R и RStudio.

Особенности интерфейса RStudio. Установка библиотек. Базовые объекты R: переменные, векторы, списки, матрицы. Установка R и RStudio. Консоль R. Знакомство с интерфейсом RStudio. Язык разметки Markdown. Создание файлов Rmarkdown. Публикация кода на Rpubs. Установка и загрузка библиотек в R. Знакомство с документацией R. R как калькулятор. Переменные в R. Типы данных: числовой, целочисленный, логический, текстовый. Преобразование типов. Факторы (factor vectors) и уровни. Базовые объекты в R: векторы, списки, матрицы. Векторы: создание, доступ к элементам, изменение, добавление и удаление элементов, выбор элементов, сортировка. Матрицы и списки: создание, доступ к элементам, изменение, добавление и удаление элементов.

Тема 2. Форматы данных. Загрузка данных в R.

Работа с текстовыми файлами в R. Разнообразие форматов данных: таблицы Excel (xls,xlsx), таблицы Stata и SPSS (dta, sav), текстовые файлы (txt, csv), json-файлы. Загрузка данных в R. Открытие, изменение и запись файлов в R. Регулярные выражения. Тема 3. Основы работы с базами данных. Выбор, добавление и удаление переменных. Преобразование типов переменных. Фильтрация, добавление и удаление наблюдений. Создание базы данных. Преобразование матриц и списков в объект data.frame. Объединение баз данных.

Тема 4. Базовые графики в R.

Визуализация количественных и качественных данных в R. Разведывательный анализ данных в R. Базовые графики в R. Функция plot(). Построение графиков математических функций. Генерирование данных в R. (Псевдо)случайные значения. Визуализация количественных данных в R: гистограммы, графики плотности распределения, ящики с усами. Визуализация качественных данных в R: таблицы сопряженности, столбчатые и круговые диаграммы. Описательные статистики: среднее арифметическое, среднеквадратичное отклонение, медиана, квантили, квартили и проценты. Разведывательный анализ данных: типы и распределения данных. Проверка данных на нормальность в R: нормальная вероятностная бумага, критерий КолмогороваСмирнова и критерий Шапиро-Уилка.

Тема 5. Корреляционный и регрессионный анализ в R.

Множественная регрессия в R. Работа с пространственно-временными данными в R. Коэффициенты корреляции Пирсона и Спирмена в R. Корреляционные матрицы в R. Визуализация корреляций между переменными в R: heatmap. Множественная регрессия в R: построение модели, интерпретация выдачи R. Визуализация результатов регрессионного анализа с помощью библиотеки sjPlot. Выгрузка необходимой информации из R в текстовые редакторы. Перекрестные данные (cross section data), временные ряды (time series data) и пространственно-временные данные (time series cross section). Регрессионные модели для пространственно-временных данных: модель с фиксированными эффектами, модель со случайными эффектами.

Тема 6. Управляющие конструкции в R.

Циклы for и while. Функции в R. Операторы в R. Условные операторы if и else. Множественные и разветвленные условия. Циклы в R. Циклы for и while: достоинства и недостатки. Циклы vs векторные операции в R. Устройство функций в R. Основные функции в R. Функция assign(). Написание простейших функций в R.

Тема 7. Парсинг веб-страниц средствами R.

Библиотека rvest. Структура html-файлов. Извлечение информации из html-файлов. Выгрузка текста из html-файлов. Функционал библиотеки rvest.

Тема 8. Работа с API средствами R. Знакомство с API. Работа с API.

4.3 Практические занятия

Таблица 4

Содержание практических занятий и контрольные мероприятия

№ п/п	Название раздела, темы	№ и название лекций/практических/семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во часов
	Раздел 1. Язык R в биологических исследованиях		ПКос-2, ПКос-3	устный опрос контрольная работа 1,2	24
1	Тема 1. Установка R и RStudio. Особенности интерфейса RStudio. Установка библиотек.	Лекционное занятие №1. Установка R и RStudio. Особенности интерфейса RStudio. Установка библиотек.	ПКос-2, ПКос-3		1
		Практическое занятие №1. Базовые объекты R: переменные, векторы, списки, матрицы. Установка R и RStudio. Консоль R. Знакомство с интерфейсом RStudio.	ПКос-2, ПКос-3	устный опрос	1
		Практическое занятие №2. Переменные в R. Типы данных: числовой, целочисленный, логический, текстовый. Преобразование типов.	ПКос-2, ПКос-3	устный опрос	1

№ п/п	Название раздела, темы	№ и название лекций/практических/семинарских занятий	Формы русские компетенции	Вид контрольного мероприятия	Кол-во часов
2	Тема 2. Форматы данных. Загрузка данных в R. Работа с текстовыми файлами в R.	Лекционное занятие №2. Форматы данных. Загрузка данных в R. Работа с текстовыми файлами в R.	ПКос-2, ПКос-3		1
2	Тема 2. Форматы данных. Загрузка данных в R. Работа с текстовыми файлами в R.	Практическое занятие №3. Разное изображение форматов данных: таблицы Excel (xls, xlsx), таблицы Stata и SPSS (dia, sav), текстовые файлы (txt, csv), json-файлы	ПКос-2, ПКос-3	устный опрос	1
		Рубежная контрольная работа по темам 1, 2	ПКос-2, ПКос-3	контрольная работа 1	1
3	Тема 3. Основы работы с базами данных	Лекционное занятие №3. Основы работы с базами данных	ПКос-2, ПКос-3		2
3	Тема 3. Основы работы с базами данных	Практическое занятие №4. Создание базы данных	ПКос-2, ПКос-3	устный опрос	2
		Лекционное занятие №4. Базовые графики в R. Визуализация количественных и качественных данных в R	ПКос-2, ПКос-3		2
4	Тема 4. Базовые графики в R. Визуализация количественных и качественных данных в R	Практическое занятие №5. Разведывательный анализ данных в R. Базовые графики в R	ПКос-2, ПКос-3	устный опрос	4
5	Тема 5. Корреляционный и регрессионный анализ в R. Множественная регрессия в R.	Лекционное занятие №5. Корреляционный и регрессионный анализ в R. Множественная регрессия в R.	ПКос-2, ПКос-3		2
		Практическое занятие №6. Работа с пространственно-временными данными в R	ПКос-2, ПКос-3	устный опрос	1
		Практическое занятие №7. Визуализация корреляций между переменными в R	ПКос-2, ПКос-3	устный опрос	1
6	Тема 6. Управление функциями в R. Циклы for и while. Функции в R	Рубежная контрольная работа по темам 3, 4, 5	ПКос-2, ПКос-3	контрольная работа 2	1
		Лекционное занятие №6. Управление функциями в R. Циклы for и while. Функции в R	ПКос-2, ПКос-3		2
6	Тема 6. Управление функциями в R. Циклы for и while. Функции в R	Практическое занятие №8. Операторы в R. Условные операторы if и else. Множественные и разветвленные условия	ПКос-2, ПКос-3	устный опрос	2

№ п/п	Название раздела, темы	№ и название лекций/практических/семинарских занятий	Формы русские компетенции	Вид контрольного мероприятия	Кол-во часов
		Практическое занятие №9. Циклы в R. Циклы for и while: достоинства и недостатки. Циклы vs векторные операции в R. Устройство функций в R	ПКос-2, ПКос-3	устный опрос	2
7	Тема 7. Парсинг веб-страниц средствами R. Библиотека rvest.	Лекционное занятие №7. Парсинг веб-страниц средствами R. Библиотека rvest.	ПКос-2, ПКос-3		2
		Практическое занятие №10. Структура html-файлов. Извлечение информации из html-файлов.	ПКос-2, ПКос-3	устный опрос	2
7	Тема 7. Парсинг веб-страниц средствами R. Библиотека rvest.	Практическое занятие №11. Выгрузка текста из html-файлов, функционал библиотеки rvest.	ПКос-2, ПКос-3	устный опрос	2
		Лекционное занятие №8. Работа с API средствами R.	ПКос-2, ПКос-3	устный опрос	2
8	Тема 8. Работа с API средствами R.	Практическое занятие №12. Знакомство с API	ПКос-2, ПКос-3	устный опрос	2
		Практическое занятие №13. Работа с API, библиотеки для работы с API	ПКос-2, ПКос-3	устный опрос	2

Таблица 5

Перечень вопросов для самостоятельного изучения дисциплины			
№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения	
Раздел 1. Язык R в биологических исследованиях			
1.	Тема 1. Установка R и RStudio. Основы работы с интерфейсом RStudio. Язык разметки интерактивной среды RStudio. Установка библиотек.	Базовые объекты R: переменные, векторы, списки, матрицы. Установка R и RStudio. Консоль R. Знакомство с интерфейсом RStudio. Язык разметки Markdown. Создание файлов Rmarkdown. Публикация кода на Rpubs. Установка и загрузка библиотек в R. Знакомство с документацией R. R как калькулятор. Переменные в R. Типы данных: числовой, логический, текстовый. Преобразование типов. Факторы (factor vectors) и уровни. Базовые объекты в R: векторы, списки, матрицы. Векторы: создание, доступ к элементам, изменение, добавление и удаление элементов, выбор элементов, сортировка. Матрицы и списки: создание, доступ к элементам, изменение, добавление и удаление элементов. ПКос-2, ПКос-3	
2.	Тема 2. Форматы данных. Загрузка данных в R. Работа с текстовыми файлами в R.	Разное изображение форматов данных: таблицы Excel (xls, xlsx), таблицы Stata и SPSS (dia, sav), текстовые файлы (txt, csv), json-файлы. Загрузка данных в R. Открытие, изменение и запись файлов в R. Регрессионные выражения. ПКос-2, ПКос-3	

№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения
3.	Тема 3. Основы работы с базами данных.	Выбор, добавление и удаление переменных. Преобразование типов переменных. Фильтрация, добавление и удаление наблюдений. Содержание базы данных. Преобразование матриц и списков в объект data.frame. Объединение баз данных.
4.	Тема 4. Базовые графики в R. Визуализация количественных и качественных данных в R.	Построение графиков математических функций. Генерирование данных в R. (Псевдо)случайные значения. Визуализация количественных данных в R: гистограммы, графики плотности распределения, ящики с усами. Визуализация качественных данных в R: таблицы сопряженности, столбчатые и круговые диаграммы. Описательные статистики: среднее арифметическое, среднеквадратичное отклонение, медиана, квантили, квартили и проценты. Разведывательный анализ данных: типы и распределения данных. Проверка данных на нормальность в R: нормальная вероятностная бумага, критерий Колмогорова-Смирнова и критерий Шапиро-Уилка. ПКос-2, ПКос-3
5.	Тема 5. Корреляционный и регрессионный анализ в R. Множественная регрессия в R.	Работа с пространственно-временными данными в R. Коэффициенты корреляции Пирсона и Спирмена в R. Корреляционные матрицы в R. Визуализация корреляций между переменными в R: heatmaps. Множественная регрессия в R: построение модели, интерпретация выводов R. Визуализация результатов регрессионного анализа с помощью библиотеки sjPlot. Выгрузка необходимых информации из R в текстовые редакторы. Перекрестные данные (cross section data), временные ряды (time series data) и пространственные данные (time series cross section). Регрессионные модели для пространственно-временных данных: модель с фиксированными эффектами, модель со случайными эффектами. ПКос-2, ПКос-3
6.	Тема 6. Управляющие конструкции в R. Циклы for и while. Функции в R.	Операторы в R. Условные операторы if и else. Множественные и разветвленные условия. Циклы в R. Циклы for и while: достоинства и недостатки. Циклы vs векторные операции в R. Устройство функций в R. Основные функции в R. Функция assign(). Написание простейших функций в R. ПКос-2, ПКос-3
7.	Тема 7. Парсинг веб-страниц средствами R. Библиотека gvest.	Структура html-файлов. Извлечение информации из html-файлов. Выгрузка текста из html-файлов. Функционал библиотеки gvest. ПКос-2, ПКос-3
8	Тема 8. Работа с API средствами R.	Знакомство с API. Работа с API. ПКос-2, ПКос-3

5. Образовательные технологии

Таблица 6
Применение активных и интерактивных образовательных технологий

№ п/п	Тема и форма занятия	Наименование используемых активных и интерактивных образова- тельных технологий (форм обучения)
1.	Тема 1. Установка R и RStudio. Особенности интер- фейса RStudio. Установка библиотек	ПЗ Интерактивная форма: мастер-класс
1.	Тема 3. Основы работы с ба- зами данных	ПЗ Интерактивная форма: мастер-класс
2.	Тема 5. Корреляционный и регрессионный анализ в R. Множественная регрессия в R.	ПЗ Круглый стол
3.	Тема 7. Парсинг веб-страниц средствами R. Библиотека gvest	ПЗ Интерактивная форма: мастер-класс

6. Текущий контроль успеваемости и промежуточная аттестация по итомам освоения дисциплины

6.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

Вопросы для подготовки к контрольным мероприятиям (текущий кон- троль)

Устный опрос

1. Общая структура научного исследования. Наблюдения и экспери-
менты.
2. Принципы планирования эксперимента.
3. Типы переменных. Матрица данных.
4. Отношения между переменными.
5. Выборочный метод.
6. Репрезентативность.
7. Описание количественных данных: меры положения.
8. Описание количественных данных: меры разброса.
9. Принципы построения гистограммы.
10. Виды распределений: модальность, скошенность.
11. Проценты и квантили.
12. Диаграмма размахов.
13. Описание качественных данных.
14. Случайные события. Алгебра событий. Теорема сложения вероят-
ностей.
15. Условная вероятность. Теорема умножения вероятностей.
16. Повторение испытаний. Формула Бернулли.
17. Случайные величины. Распределение вероятностей.
18. Общая характеристика нормального распределения.
19. z-преобразование.

20. Плотность распределения.
21. Функции распределения.
22. Диагностика нормального распределения: визуальные средства.
23. Диагностика нормального распределения: формальные критерии.
24. Структура статистического критерия. Гипотезы.
25. Процедура проверки гипотез.
26. Критериальная статистика. p-значение.
27. Ошибки статистических критериев.
28. Односторонние и двусторонние критерии.
29. Расчет вероятности попадания в интервал.
30. Расчет вероятностей отклонения от среднего.
31. Стандартная ошибка.
32. Доверительный интервал.
33. Доверительная вероятность.
34. Использование нормального распределения для проверки гипотез.
35. Распределение выборочной доли.
36. Распределение выборочного среднего.
37. Сравнение доли с эталонным значением.
38. Сравнение двух выборочных долей.
39. Сравнение распределения категориальной переменной с теоретическим.
40. Анализ зависимости между двумя категориальными переменными.
41. Анализ категориальных переменных без использования непрерывных распределений.
42. Сравнение среднего с эталоном.
43. Сравнение средних в независимых выборках.
44. Сравнение средних в связанных выборках.
45. Использование доверительных интервалов для сравнения средних.
46. Критерий Уилкоксона (Манна-Уитни).
47. Однофакторный дисперсионный анализ.
48. Проблема множественных сравнений.
49. Коэффициент корреляции и его анализ.
50. Основы регрессионного анализа.

Контрольная работа №1 Вариант 1

1. R – это свободно распространяемая версия:
- а) языка программирования C;
 - б) языка программирования S;
 - в) программы Statistica;
 - г) среды вычислений Matlab.

2. Результат выражения 19 %/% 5 * 3:

- а) 4;
- б) 12;
- в) 9;
- г) NA.

3. Какая из функций осуществляет округление?

- а) avenge();
- б) sqrt();
- в) abs();
- г) ceiling().

4. Какая из перечисленных операций присваивания не может быть выполнена?

- а) a = b = 5;
- б) 7 -> d -> f;
- в) 6 = r = s;
- г) w <- 5 -> z.

5. Какая команда не создает вектор (5, 5, 5, 3, 3, 3, 1, 1, 1)?

- а) c(c(5,5,5), c(3,3,3), c(1,1,1));
- б) rep(5:1, by = -2, each = 3);
- в) c(rep(5,3), rep(3,3), rep(1,3));
- г) все создано

Контрольная работа №2 Вариант 1

1. Какова площадь под кривой стандартного нормального распределения левее

$z = -1.4$:

- а) 0.04;
- б) 0.08;
- в) 0.16;
- г) 0.8.

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) $H_0: \mu = 12$

Перечень вопросов, выносимых на промежуточную аттестацию (экзамен)

1. Общая структура научного исследования. Наблюдения и эксперименты.
2. Принципы планирования эксперимента.
3. Типы переменных. Матрица данных.
4. Отношения между переменными.
5. Выборочный метод.
6. Репрезентативность.
7. Описание количественных данных: меры положения.
8. Описание количественных данных: меры разброса.
9. Принципы построения гистограммы.
10. Виды распределений: модальность, скошенность.
11. Процентили и квантили.
12. Диаграмма размахов.
13. Описание качественных данных.
14. Случайные события. Алгебра событий. Теорема сложения вероятностей.
15. Условная вероятность. Теорема умножения вероятностей.
16. Повторение испытаний. Формула Бернулли.
17. Случайные величины. Распределение вероятностей.
18. Общая характеристика нормального распределения.
19. z-преобразование.
20. Плотность распределения.
21. Функция распределения.
22. Диагностика нормального распределения: визуальные средства.
23. Диагностика нормального распределения: формальные критерии.
24. Структура статистического критерия. Гипотезы.
25. Процедура проверки гипотез.
26. Критериальная статистика. p-значение.
27. Ошибки статистических критериев.
28. Односторонние и двусторонние критерии.
29. Расчет вероятности попадания в интервал.
30. Расчет вероятностей отклонения от среднего.
31. Стандартная ошибка.
32. Доверительный интервал.
33. Доверительная вероятность.
34. Использование нормального распределения для проверки гипотез.
35. Распределение нормальной доли.
36. Распределение выборочного среднего.
37. Сравнение доли с эталонным значением.
38. Сравнение двух выборочных долей.
39. Сравнение распределения категориальной переменной с теоретическим.
40. Анализ зависимости между двумя категориальными переменными.

41. Анализ категориальных переменных без использования непрерывных распределений.
42. Сравнение среднего с эталоном.
43. Сравнение средних в независимых выборках.
44. Сравнение средних в связанных выборках.
45. Использование доверительных интервалов для сравнения средних.
46. Критерий Уилкоксона (Манна-Уитни).
47. Однофакторный дисперсионный анализ.
48. Проблема множественных сравнений.
49. Коэффициент корреляции и его анализ.
50. Основы регрессионного анализа.

6.2. Описание показателей и критериев контроля успеваемости, описание шкал оценивания

Балльно-рейтинговая система оценки

Объем рейтинга составляет: за текущий контроль - 30% от нормативного рейтинга дисциплины, за рубежный контроль - 30% от нормативного рейтинга дисциплины и за итоговый контроль - 40% от нормативного рейтинга дисциплины.

Таблица 7

Система рейтинговой оценки

Оценочные средства	Баллы			
	0	2	4	5
Устный опрос	0	2	4	5
Контрольная работа	0-4	5-6	7-8	9-10
Экзамен	0-8	9-13	14-17	18-20
Оценка	Неуд.	Удовл.	Хорошо	Отлично
Посещение лекций и практических занятий				
Посещаемость	≤85%	86-88%	89-91%	92-100%
Баллы	0	10	20	30

Посещаемость рассчитывается, как отношение числа пропущенных занятий к общему числу занятий.

Максимальное число баллов – 100

Для допуска к сдаче зачета с оценкой по дисциплине необходимо:

- фактический рейтинг семестрового контроля должен составлять более 50% от нормативного рейтинга семестрового контроля для дисциплины (Рфакт.сем > 50%Rнорм семестр), т.е. должен быть достигнут пороговый рейтинг;
- должен быть выполнен объем аудиторных занятий (включая посещение лекций), предусмотренный учебным планом.

Рейтинговый балл, выставляемый студенту

Рейтинговый балл

(в % от макс. балла за дисциплину)

85,1-100%

65,1 – 85 %

60,1 – 65 %

Менее 60 %

Оценка по традиционной шкале

Отлично

Хорошо

Удовлетворительно

Неудовлетворительно

7. Учебно-методическое и информационное обеспечение дисциплины

7.1 Основная литература

1. Буховец, А. Г. Алгоритмы вычислительной статистики в системе R: учебное пособие / А. Г. Буховец, П. В. Москалев. — 2-е изд., перераб. и доп. — Санкт-Петербург : Лань, 2022. — 160 с. — ISBN 978-5-8114-1802-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/212195>
2. Галиновский, Н. Г. Введение в программирование на языке R: учебное пособие / Н. Г. Галиновский. — Гомель : ГТУ имени Ф. Скорины, 2022. — 222 с. — ISBN 978-985-577-826-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/320945>

7.2 Дополнительная литература

1. Буре, В. М. Методы прикладной статистики в R и Excel / В. М. Буре, Е. М. Паршина, А. А. Седяков. — 4-е изд., стер. — Санкт-Петербург : Лань, 2024. — 152 с. — ISBN 978-5-507-46766-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/319424>
2. Гришин, В. А. Основы программирования на языке R : учебно-методическое пособие / 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>
3. Каледин, В. О. Теория R-функций : учебное пособие / В. О. Каледин. — 2-е изд., перераб. и доп. — Новокузнецк : КГПИ КемГУ, 2017. — 119 с. — ISBN 978-5-8353-1952-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/169599>
4. Кузьмин, В. И. Методы анализа данных : учебное пособие / В. И. Кузьмин, А. Ф. Газзаев. — 2-е изд., перераб. и доп. — Москва : РТУ МИРЭА, 2020. — 155 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/171433>
5. Пыльнев, В. В. Основы селекции и семеноводства / В. В. Пыльнев, А. Н. Березкин ; Под ред.: Пыльнев В. В. — 2-е изд., стер. — Санкт-Петербург : Лань, 2024. — 216 с. — ISBN 978-5-507-45402-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/267383>

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

1. <https://www.r-project.org/> (основной сайт R)
2. <https://cran.r-project.org/> (архив пакетов для статистического анализа данных в R)
3. <http://www.biostatistics.org/> (Biostatistics – архив пакетов R, предназначенных для анализа биологических данных).

9. Перечень программного обеспечения и информационных справочных систем

Таблица 8

Перечень программного обеспечения				
№ п/п	Наименование раздела учебной дисциплины	Наименование программы	Тип программы	Автор разра-ботки
1	Коммерческие программное обеспечение и информационно справочные системы не используются			

10. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Процесс изучения дисциплины обеспечен аудиторией, оборудованной персональными компьютерами, мультимедийными средствами для демонстрации презентаций и доступом к информационно-телекоммуникационной сети «Интернет».

Таблица 9

Сведения об обеспеченности специализированными аудиториями, кабинетами, лабораториями

Наименование специальных помещений и помещений для самостоятельной работы (№ учебного курса, № аудитории)	Оснащенность специальных помещений и помещений для самостоятельной работы**
Центральная научная библиотека имени Н.И. Жезезнова. Читальные залы библиотек	Столы, стулья, учебная литература
Общедоступные №5 Комната для самоподготовки	Столы, стулья

11. Методические рекомендации студентам по освоению дисциплины

Основной объем материала студенту необходимо освоить самостоятельно в соответствии с темами для самостоятельной подготовки из таблицы 5. Студенты должны соблюдать дисциплину, вовремя приходить на занятия, пред-

ставлять на проверку домашнюю работу, готовиться к проверочным и контрольным работам, предусмотренным курсом, проявлять активность на занятиях. Важное место в образовательном процессе занимает самостоятельная работа студентов. Для организации самостоятельной работы студентов по курсу используются современные информационные технологии: размещенные в сетевом доступе комплексы учебных и учебно-методических материалов (программа, список рекомендованной литературы и информационные ресурсы, задания для самоконтроля), свободный доступ к сети «Интернет» для работы с молекулярными базами данных.

Виды и формы отработки пропущенных занятий

Студент, пропустивший занятия обязан предоставить и защитить реферат по пропущенной теме.

12. Методические рекомендации преподавателям по организации обучения по дисциплине

Педагог, проводящий занятия, должен обладать высокой квалификацией и опытом. Необходимо разбираться в нюансах работы, чтобы при необходимости была возможность исправить ошибку студента. Для успешного освоения предмета необходимо периодически организовывать обсуждения и дискуссии по темам дисциплины.

Все практические работы носят строго профессиональный характер. Навыки, полученные при выполнении этих работ, пригодятся студенту на всех этапах обучения, при подготовке выпускной работы магистра и в профессиональной деятельности.

При преподавании курса необходимо ориентироваться на современные образовательные технологии путем использования группового способа обучения на практических занятиях, разбора конкретных ситуаций и интерактивного обсуждения результатов исследовательских учебных работ. Реализация современного подхода должна обеспечиваться широким использованием активных интерактивных форм проведения занятий, посещение профильных научно-исследовательских учреждений и повышение интереса к изучению дисциплины. Задачей преподавателя является приведение максимального количества позитивных примеров учреждений и специалистов добившихся высоких результатов в своих отраслях биотехнологии, для стимулирования интереса студентов к углубленному изучению данных дисциплин.

Программу разработал (и):

Лисовая Д.Д., ассистент

Монахов С.Г., д.с.-х.н., профессор

РЕЦЕНЗИЯ

на рабочую программу дисциплины «Язык R в биологических исследованиях» ОПОП ВО по направлению 06.04.01 Биология, направленность «Биоинформатика» (квалификация выпускника – магистр).

Монахосом Григорием Федоровичем, генеральным директором ООО «Селекционная станция им. Н.Н.Тимофеева», кандидатом сельскохозяйственных наук, старшим научным сотрудником (далее по тексту рецензент), проведена рецензия рабочей программы дисциплины «Язык R в биологических исследованиях» ОПОП ВО по направлению 06.04.01 Биология, направленность «Биоинформатика» (магистратура) разработанной в ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», на кафедре ботаники, селекции и семеноводства садовых растений (*разработчики – Лисовая Дарья Дмитриевна, ассистент, Монахос Сократ Григорьевич, заведующий кафедрой, д. с.-х.н., профессор*).

Рассмотрев представленные на рецензию материалы, рецензент пришел к следующим выводам:

1. Предъявленная рабочая программа дисциплины «Язык R в биологических исследованиях» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 06.04.01 Биология. Программа содержит все основные разделы, соответствует требованиям к нормативно-методическим документам.

2. Представленная в Программе **актуальность** учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к части, формируемой участниками образовательных отношений – Б1.

3. Представленные в Программе цели дисциплины соответствуют требованиям ФГОС ВО направления 06.04.01 Биология.

4. В соответствии с Программой за дисциплиной «Язык R в биологических исследованиях» закреплена **2 компетенции**. Дисциплина «Язык R в биологических исследованиях» и представленная Программа способна реализовать их в объявленных требованиях.

Результаты обучения, представленные в Программе в категориях знать, уметь, владеть соответствуют специфике и содержанию дисциплины и демонстрируют возможность получения заявленных результатов.

5. Общая трудоёмкость дисциплины «Язык R в биологических исследованиях» составляет **4 зачётных единицы (144 часов)**.

6. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Язык R в биологических исследованиях» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 06.04.01 Биология и возможность дублирования в содержании отсутствует.

7. Представленная Программа предполагает использование современных образовательных технологий, используемые при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

8. Программа дисциплины «Язык R в биологических исследованиях» предполагает **18 часов** занятий в интерактивной форме.

9. Виды, содержание и трудоёмкость самостоятельной работы студентов, представленные в Программе, соответствуют требованиям к подготовке выпускников, содержащимся во ФГОС ВО направления 06.04.01 Биология.

10. Представленные и описанные в Программе формы *текущей* оценки знаний (опрос, как в форме обсуждения отдельных вопросов, так и выступления и участие в дискуссиях, диспутах и аудиторных заданиях), соответствуют специфике дисциплины и требованиям к выпускникам.

Форма промежуточного контроля знаний студентов, предусмотренная Программой, осуществляется в форме **экзамена**, что соответствует статусу дисциплины, как дисциплины учебного цикла – Б1 ФГОС ВО направления 06.04.01 Биология.

11. Формы оценки знаний, представленные в Программе, соответствуют специфике дисциплины и требованиям к выпускникам.

12. Учебно-методическое обеспечение дисциплины представлено: основной литературой – **2 источника** (базовый учебник), дополнительной литературой – **5 наименования**, Интернет-ресурсы – **3 источников** и соответствует требованиям ФГОС ВО направления 06.04.01 Биология.

13. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины «Язык R в биологических исследованиях» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

14. Методические рекомендации студентам и методические рекомендации преподавателям по организации обучения по дисциплине дают представление о специфике обучения по дисциплине «Язык R в биологических исследованиях».

ОБЩИЕ ВЫВОДЫ

На основании проведенной рецензии можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины «Язык R в биологических исследованиях» ОПОП ВО по направлению 06.04.01 Биология, направленность «Язык R в биологических исследованиях» (квалификация выпускника – магистр), разработанная Лисовой Дарьей Дмитриевной, ассистентом и Монахосом Сократом Григорьевичем, заведующим кафедрой, д.с.-х.н., профессором соответствует требованиям ФГОС ВО, современным требованиям экономики, рынка труда и позволит при её реализации успешно обеспечить формирование заявленных компетенций.

Рецензент: Монахос Григорий Федорович, генеральный директор ООО «Селекционная станция им. Н.Н.Тимофеева», кандидат сельскохозяйственных наук, старший научный сотрудник _____ «20» августа 2024 г.
(подпись)