

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Раджабов А.К. Мухоморов

Должность: И.о. директора института садоводства и ландшафтной архитектуры

Дата подписания: 2023.08.21:06:03

Уникальный программный ключ:

088d9d84706d89075c4a3aa1678d7c4c996222db



**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:

Директор института садоводства и  
ландшафтной архитектуры

Раджабов А.К.

“21” августа 2023 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Б1.В.ДВ.02.02 Программирование и анализ биоданных R**

для подготовки магистров

ФГОС ВО

Направление 35.04.05 Садоводство

Направленность (программа) «Биотехнология и селекция растений»

Курс: 1

Семестр: 1

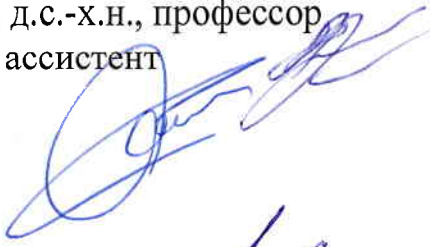
Форма обучения: очная

Год начала подготовки: 2023

Регистрационный номер \_\_\_\_\_

Москва, 2023

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



«21» августа 2023 г.

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

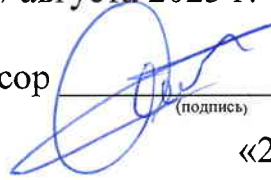


«21» августа 2023 г.

Программа составлена в соответствии с требованиями ФГОС ВО, по направлению подготовки 35.04.05 Садоводство и учебного плана.

Программа обсуждена на заседании кафедры ботаники, селекции и семеноводства садовых растений, протокол №15 от «21» августа 2023 г.

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



(подпись)

«21» августа 2023 г.

**Согласовано:**

Председатель учебно-методической  
комиссии факультета Маланкина Е.Л., д.с.-х.н.

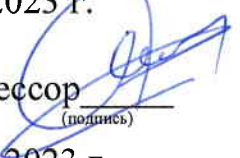


(ФИО, ученая степень, ученое звание)

(подпись)

«21» августа 2023 г.

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



(подпись)

«21» августа 2023 г.

Зав. Отделом комплектования ЦНБ



(подпись)

**Бумажный экземпляр ПП, электронные варианты ПП и оценочных материалов получены:**

Методический отдел УМУ \_\_\_\_\_ «\_\_» \_\_\_\_\_ 20\_\_ г

## СОДЕРЖАНИЕ

<b>1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....</b>	<b>4</b>
<b>2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ.....</b>	<b>4</b>
<b>3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ «ПРОГРАММИРОВАНИЕ И АНАЛИЗ БИОДАНЫХ R», СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ.....</b>	<b>5</b>
<b>4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ .....</b>	<b>5</b>
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ .....	5
ПО СЕМЕСТРАМ .....	5
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ .....	8
4.3 ПРАКТИЧЕСКИЕ ЗАНЯТИЯ .....	10
<b>5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ .....</b>	<b>15</b>
<b>6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....</b>	<b>15</b>
6.1. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ.....	15
<b>7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ .....</b>	<b>20</b>
7.1 ОСНОВНАЯ ЛИТЕРАТУРА.....	20
7.2 ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА .....	20
<b>8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ) .....</b>	<b>20</b>
<b>9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ .....</b>	<b>20</b>
<b>10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ .....</b>	<b>21</b>
<b>11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ СТУДЕНТАМ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ.....</b>	<b>21</b>
Виды и формы отработки пропущенных занятий.....	22
<b>12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ .....</b>	<b>22</b>

**АННОТАЦИЯ**  
**рабочей программы учебной дисциплины**  
**Б1.В.ДВ.02.02 «Программирование и анализ биоданных R»**  
для подготовки магистра по направлению 35.04.05 Садоводство  
направленности «Биотехнология и селекция растений»

**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 35.04.05 Садоводство

**Требования к результатам освоения дисциплины:** в результате освоения дисциплины формируются следующие компетенции: 1 профессиональную компетенцию ПКос-1 (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-1.4).

**Краткое содержание дисциплины:** Дисциплина "Программирование и анализ биоданных R" 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:** 180/5 (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 «Программирование и анализ биоданных R» включена в часть профессионального цикла, формируемую участниками образовательных отношений. Реализация в дисциплине «Программирование и анализ биоданных R» требований ФГОС ВО, ОПОП и Учебного плана по направлению 35.04.05 Садоводство для подготовки магистров направленности «Биотехнология и селекция растений».

Предшествующими курсами, на которых непосредственно базируется дисциплина «Программирование и анализ биоданных R», являются «Статистика», «Биология», «Молекулярная биология».

Дисциплина «Программирование и анализ биоданных R» является основополагающей для изучения следующих дисциплин: «Концепции ускоренной селекции растений», «Генетические основы селекции растений», «Геномика растений».

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 «Программирование и анализ биоданных R» 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 «Программирование и анализ биоданных R», 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 трудоёмкости дисциплины по видам работ по семестрам**

Общая трудоёмкость дисциплины составляет 5 зач.ед. (180 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-1	Способен проводить полевые и лабораторные опыты с использованием традиционных и современных методов, в т.ч. цифровых	<p>ПКос-1.1 проводить поиск и анализ данных (в том числе с использованием методов биоинформатики), научной литературы для достижения поставленной цели научного исследования</p> <p>ПКос-1.2 организация закладки полевых и лабораторных опытов в рамках испытаний растений и влияния условий на проявление их признаков и свойств</p>	<p>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</p>	<p>Использовать R для статистической обработки экспериментальных и клинических данных. Уметь интерпретировать полученные результаты.</p>	<p>Написания простейших программ, построения графиков и диаграмм в R. Владеть основными методами статистики, методами машинного обучения, навыками их использованию с помощью среды R.</p>
				<p>Основные понятия теории вероятностей и математической статистики, основные методы описания данных</p>	<p>Ориентироваться в основных алгоритмах статистической обработки данных для решения практических задач и графического представления результатов</p>	<p>Соответствующими приемами программирования, компьютерными методами обработки данных биологических исследований</p>

		<p>ПКос-1.3 производить учеты и наблюдения в опытах для испытания растений с оценкой влияния условий на проявление признаков и свойств, в т.ч. с использованием цифровых методов</p>	<p>Принципы применения теории вероятностей и математической статистики для описания данных биологических исследований</p>	<p>Применять методы описания, статистического анализа и визуализации данных биологических исследований</p>	<p>Компьютерными средствами для подготовки, хранения и анализа данных биологических исследований</p>
		<p>ПКос-1.4 определять комплекс традиционных и современных (полевых и лабораторных) методов исследования для решения научных задач</p>	<p>основные термины и понятия биоинформатики, современные концепции биоинформатики; объекты изучения биоинформатики: последовательности нуклеиновых кислот и аминокислот</p>	<p>находить, анализировать, обобщать и систематизировать научные данные, полученные в ходе лабораторных экспериментов, для постановки целей исследований и выбора оптимальных методов их достижения;</p>	<p>правилами расчетов оптимальных параметров проведения анализа, систематизации и интерпретации данных биологических объектов, и их корректирования</p>

## ОЧНАЯ ФОРМА ОБУЧЕНИЯ

Таблица 2а

### Распределение трудоёмкости дисциплины по видам работ по семестрам

Вид учебной работы	Трудоёмкость
	час.
<b>Общая трудоёмкость</b> дисциплины по учебному плану	<b>180</b>
<b>1. Контактная работа:</b>	<b>48.4</b>
<b>Аудиторная работа</b>	<b>48.4</b>
<i>в том числе:</i>	
лекции (Л)	6
практические занятия (ПЗ)	40/4
консультации перед экзаменом	2
контактная работа на промежуточном контроле (КРА)	0,4
<b>2. Самостоятельная работа (СРС)</b>	<b>107</b>
<i>самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиумам и т.д.)</i>	82,4
Подготовка к экзамену (контроль)	24,6
Вид промежуточного контроля:	Экзамен

## 4.2 Содержание дисциплины

Таблица 3

### Тематический план учебной дисциплины

Наименование разделов и тем дисциплин (укрупнённо)	Всего	Аудиторная работа			Внеаудиторная работа СР
		Л	ПЗ/С	ПКР	
<b>Раздел 1 Программирование и анализ биоданных R</b>	180	6	40	2,4	107
Тема 1. Установка R и RStudio. Особенности интерфейса RStudio. Установка библиотек.	7,4	1	5	-	13
Тема 2. Форматы данных. Загрузка данных в R. Работа с текстовыми файлами в R.	8	1	5	-	13
Тема 3. Основы работы с базами данных.	8,5	1	5	-	13
Тема 4. Базовые графики в R. Визуализация количественных и качественных данных в R.	10,5	-	5	-	14
Тема 5. Корреляционный и регрессионный анализ в R. Множественная регрессия в R.	8,5	-	5	-	14
Тема 6. Управляющие конструкции в R. Циклы for и while. Функции в R.	10,5	1	5	-	14
Тема 7. Парсинг веб-страниц средствами R. Библиотека rvest.	10,5	1	5	-	13
Тема 8. Работа с API средствами R.	10,5	1	5	-	13



Наименование разделов и тем дисциплин (укрупненно)	Всего	Аудиторная работа			Внеаудиторная работа СР
		Л	ПЗ/С	ПКР	
Консультация перед экзаменом	2	-	-	2	-
Контактная работа на промежуточном контроле (КРА)	0,4	-	-	0,4	-
Подготовка к экзамену	24,6	-	-	-	24,6
<b>Итого по дисциплине</b>	<b>180</b>	<b>6</b>	<b>40</b>	<b>2,4</b>	<b>107</b>

## Раздел 1 Программирование и анализ биоданных R

### 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 процентиля. Разведывательный анализ данных: типы и распределения данных. Проверка данных на нормальность в R: нормальная вероятностная бумага, критерий Колмогорова-Смирнова 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: heatmaps. 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
	<b>Раздел 1. Программирование и анализ биоданных R</b>		ПКос-1	устный опрос контрольная работа 1,2	<b>46</b>
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-1		1
		Практическое занятие №1. Базовые объекты R: переменные, векторы, списки, матрицы. Установка R и	ПКос-1	устный опрос	2

№ п/п	Название раздела, темы	№ и название лекций/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во часов
		RStudio. Консоль R. Знакомство с интерфейсом RStudio			
		Практическое занятие №2. Переменные в R. Типы данных: числовой, целочисленный, логический, текстовый. Преобразование типов.	ПКос-1	устный опрос	3
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-1		1
		Практическое занятие №3. Разнообразие форматов данных: таблицы Excel (xls, xlsx), таблицы Stata и SPSS (dta, sav), текстовые файлы (txt, csv), json-файлы	ПКос-1	устный опрос	4
		Рубежная контрольная работа по темам 1, 2	ПКос-1	контрольная работа 1	1
3	Тема 3. Основы работы с базами данных	Лекционное занятие №3. Основы работы с базами данных	ПКос-1		2
		Практическое занятие №4. Создание базы данных.	ПКос-1	устный опрос	3
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-1	устный опрос	5
5	Тема 5. Корреляционный и регрессионный анализ	Практическое занятие №6. Работа с пространственно-временными данными в R	ПКос-1	устный опрос	2
		Практическое занятие №7	ПКос-1	устный опрос	2

№ п/п	Название раздела, темы	№ и название лекций/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во часов
	в R. Множественная регрессия в R.	Визуализация корреляций между переменными в R			
		Рубежная контрольная работа по темам 3, 4, 5	ПКос-1	контрольная работа 2	1
6	Тема 6. Управляющие конструкции в R. Циклы for и while. Функции в R	Лекционное занятие №6. Управляющие конструкции в R. Циклы for и while. Функции в R	ПКос-1		1
		Практическое занятие №8. Операторы в R. Условные операторы if и else. Множественные и разветвленные условия.	ПКос-1	устный опрос	2
		Практическое занятие №9. Циклы в R. Циклы for и while: достоинства и недостатки. Циклы vs векторные операции в R. Устройство функций в R	ПКос-1	устный опрос	3
7	Тема 7. Парсинг веб-страниц средствами R. Библиотека rvest.	Лекционное занятие №7. Парсинг веб-страниц средствами R. Библиотека rvest.	ПКос-1		1
		Практическое занятие №10. Структура html-файлов. Извлечение информации из html-файлов.	ПКос-1	устный опрос	2
		Практическое занятие №11. Выгрузка текста из html-файлов. Функционал библиотеки rvest.	ПКос-1	устный опрос	3
8	Тема 8. Работа с API средствами R.	Лекционное занятие №8. Работа с API средствами R.	ПКос-1		1
		Практическое занятие №12. Знакомство с API.	ПКос-1	устный опрос	2
		Практическое занятие №13.	ПКос-1	устный опрос	3

№ п/п	Название раздела, темы	№ и название лекций/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во часов
		Работа с API, библиотеки для работы с API			

Таблица 5

**Перечень вопросов для самостоятельного изучения дисциплины**

№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения
<b>Раздел 1. Программирование и анализ биоданных R</b>		
1.	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. ПКос-1
2.	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. ПКос-1
3.	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.
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 стати-

№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения
		стики: среднее арифметическое, среднее квадратичное отклонение, медиана, квантили, квартили и процентиля. Разведывательный анализ данных: типы и распределения данных. Проверка данных на нормальность в R: нормальная вероятностная бумага, критерий Колмогорова-Смирнова и критерий Шапиро-Уилка. ПКос-1
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 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. ПКос-1
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-1
7.	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. ПКос-1
8	Тема 8. Работа с API средствами R.	Знакомство с API. Работа с API. ПКос-1

## 5. Образовательные технологии

Таблица 6

### Применение активных и интерактивных образовательных технологий

№ п/п	Тема и форма занятия		Наименование используемых активных и интерактивных образовательных технологий (форм обучения)
1.	Тема 1. Установка R и RStudio. Особенности интерфейса RStudio. Установка библиотек	ПЗ	Интерактивная форма: мастер-класс
1.	Тема 3. Основы работы с базами 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 rvest	ПЗ	Интерактивная форма: мастер-класс

### 6. Текущий контроль успеваемости и промежуточная аттестация по итогам 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 контроль)

##### Устный опрос

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. Перцентили 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 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) average();
- б) sqrt();
- в) abs();
- г) ceiling().

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

- а)  $a = b = 5$ ;
- б)  $7 \rightarrow d \rightarrow f$ ;
- в)  $6 = r = s$ ;
- г)  $w \leftarrow 5 \rightarrow z$ .

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 создают

**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)  $H_0: = 12$ ;
- в)  $H_0: > 12$ ;
- 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. Перцентили 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 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.

## 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			
	Устный опрос	0	2	4
Контрольная работа	0-4	5-6	7-8	9-10
Экзамен	0-8	9-13	14-17	18-20
<b>Оценка</b>	<b>Неуд.</b>	<b>Удовл.</b>	<b>Хорошо</b>	<b>Отлично</b>
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 ( $R_{\text{факт.сем}} > 50\%R_{\text{норм семестр}}$ ), 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)	Оценка по традиционной шкале
85,1-100%	Отлично
65,1 – 85 %	Хорошо
60,1 – 65 %	Удовлетворительно
Менее 60 %	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 : ГГУ 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. Парилина, А. А. Седаков. — 4-е изд., стер. — Санкт-Петербург : Лань, 2023. — 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. Гадзаов. — 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, 2023. — 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.bioconductor.org/> (Bioconductor – 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	Год разработки
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 корпуса, № аудитории)	Оснащенность специальных помещений и помещений для самостоятельной работы**
--	---

1	2
Центральная научная библиотека имени Н.И. Железнова, Читальные залы библиотеки	Столы, стулья, учебная литература
Общежитие №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, представлять 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 рекомендуемой литературы и информационных ресурсов, 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 повысить интерес к изучению дисциплины. Задачей преподавателя является приведение максимального количества позитивных приме-

ров 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

Монахос С.Г., д.с.-х.н., профессор



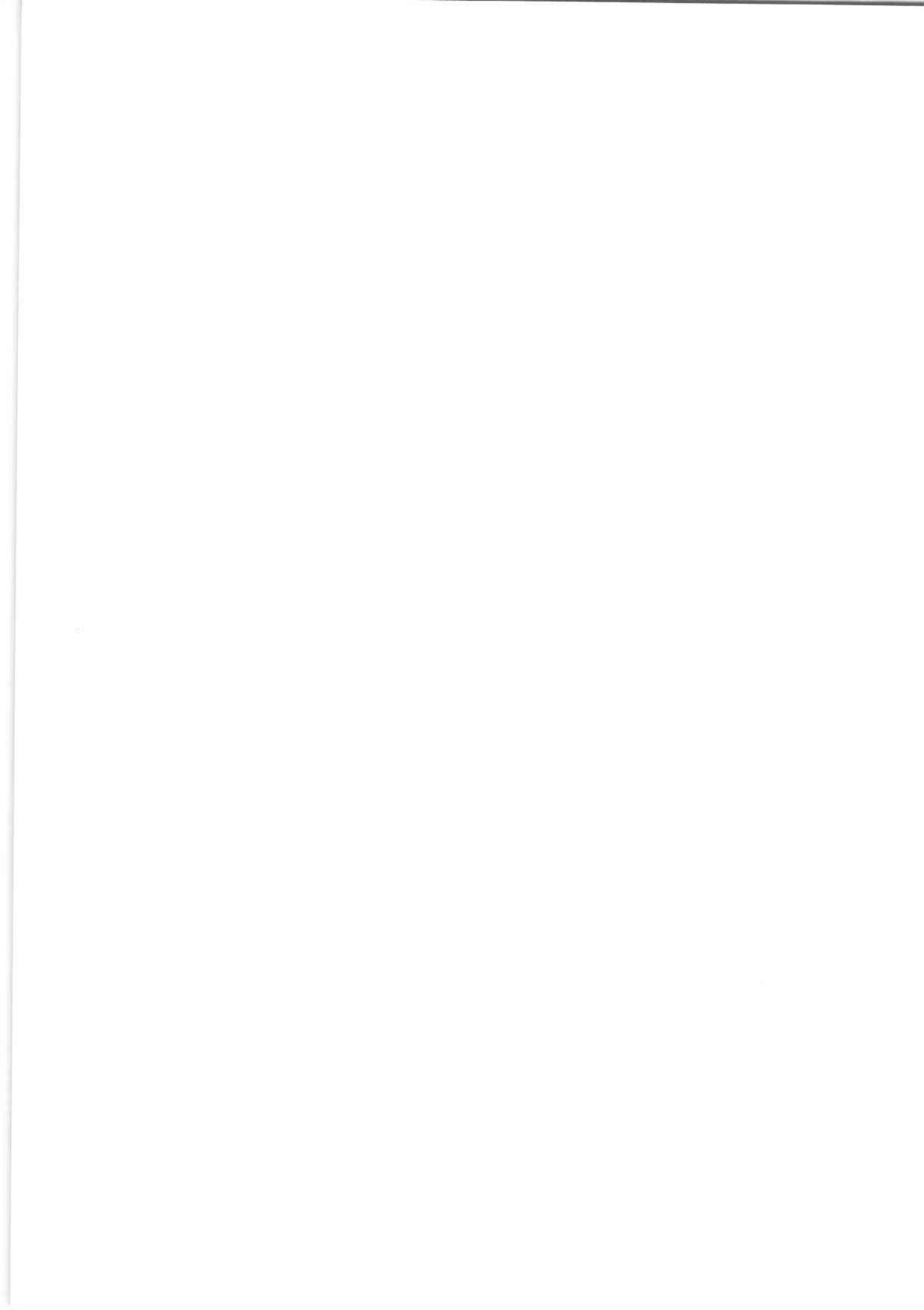
---

(подпись)



---

(подпись)





## РЕЦЕНЗИЯ

на рабочую программу дисциплины «Программирование и анализ биоданных R»  
ОПОП ВО по направлению 35.04.05 Садоводство, направленность «Биотехнология и  
селекция растений» (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 «Программирование и анализ биоданных R» ОПОП ВО по направлению 35.04.05 Садоводство, направленность «Биотехнология и селекция растений» (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 «Программирование и анализ биоданных R» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 35.04.05 Садоводство. 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 35.04.05 Садоводство.

4. В соответствии с Программой за дисциплиной «Программирование и анализ биоданных R» закреплена **1 компетенция**. Дисциплина «Программирование и анализ биоданных R» 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 «Программирование и анализ биоданных R» составляет **5 зачётных единиц (180 часов)**.

6. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Программирование и анализ биоданных R» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 35.04.05 Садоводство 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 «Программирование и анализ биоданных R» предполагает **20 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 35.04.05 Садоводство.

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 35.04.05 Садоводство.

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 35.04.05 Садоводство.

13. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины «Программирование и анализ биоданных R» 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 «Программирование и анализ биоданных R».

#### **ОБЩИЕ ВЫВОДЫ**

На основании проведенной рецензии можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины «Программирование и анализ биоданных R» ОПОП ВО по направлению 35.04.05 Садоводство, направленность «Программирование и анализ биоданных R» (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 2023 г.

(подпись)