Федеральное государственное бюджетное образовательное учреждение высшего образования «Новосибирский государственный аграрный университет»

Рег. № МиСМОНИ.БР.18-5
«30» сентября 2022 г.

Утверждаю
Ректор
_____ Е.В. Рудой
«29» _____ 2022г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«2.1.5 *Математические и статистические методы обработки научных исследований»*

Научная специальность: 1.5.20. Биологические ресурсы

Новосибирск 2022

Рабочая программа дисциплины «Математические и статистические методы обработки научных исследований» составлена в соответствии с Федеральными государственными требованиями к структуре программ подготовки научных и научно-педагогических кадров в аспирантуре, условиям их реализации, срокам освоения этих программ с учетом различных форм обучения, образовательных технологий и особенностей отдельных категорий аспирантов, утвержденными приказом Министерства науки и высшего образования Российской Федерации от 20 октября 2021 г. № 951.

РАЗРАБОТЧИК(И):

проректор по научной и международной деятельности, доктор биологических наук, доцент _____ Е.В. Камалдинов

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ РАССМОТРЕНА И ОДОБРЕНА:

- на заседании кафедры ветеринарной генетики и биотехнологии

Протокол заседания № 1 от «20» 09 2022 г.

Зав. кафедрой _____ Н.Н. Кочнев
подпись

- на заседании Учебно-методического совета биолого-технологического факультета

Протокол заседания № 7 от «29» 09 2022 г.

Председатель комиссии
Профессор каф., д-р биол. наук, доцент _____ М.Л. Кочнева
подпись

1. Перечень сокращений, используемых в тексте рабочей программы дисциплины

- з.е. – зачетная единица
- ФГТ– Федеральные государственные требования
- ФОС – фонд оценочных средств
- Пр – практическое занятие
- Лаб – лабораторное занятие
- Лек – лекции
- СР – самостоятельная работа

2. Цель изучения дисциплины

Дать обучающимся углубленные знания по дисциплине «Математические и статистические методы научных исследований», выработать умения первичной обработки исходных данных необходимые для обоснования выводов и принятия решений.

3. Требования к результатам освоения дисциплины

Знать:

-методику современных статистических методов обработки методов среде программирования R

-актуальные статистические методы решения задач в соответствующей профессиональной области

Уметь:

-использовать готовые библиотеки и писать собственный код для решения исследовательских и практических задач, в том числе в междисциплинарных областях

-использовать готовые библиотеки и писать собственный код для решения исследовательских и практических задач, в том числе в междисциплинарных областях

Владеть:

-приёмами и методами обработки данных с использованием языка программирования R

-навыками прикладного программирования для решения аналитических задач

4. Объем дисциплины

4.1 Общая трудоемкость дисциплины составляет: 3 з.е. /108 ч.
(из них 27 ч. – самостоятельная работа обучающихся).

4.2. Контактная работа включает:
- лекции: 26 ч.
- практические занятия: 28 ч.

- мероприятия промежуточной аттестации 27 ч.

4.3. Форма контроля – зачет, экзамен:
- экзамен проводится - в 6 семестре 3 курса.

5. Содержание дисциплины

5.1. Тематические разделы курса

Таблица 2 – Тематический план курса «Математические и статистические методы обработки научных исследований»

№ п/п	Темы	Количество часов				Всего по темам
		Контактная работа обучающихся с преподавателем			СР	
		Лек	Пр	КР		
1	2	3	4	5	6	7
1	Разнообразие признака. Описательная статистика	3	2	2	2	9
2	Виды распределений	3	2	2	2	9
3	Оценка достоверности различий между выборками	3	2	2	2	9
4	Оценка взаимосвязи между признаками	3	2	2	2	9
5	Методы оценки качественных признаков	3	2	2	2	9
6	Значимость группирующего фактора	3	2	2	2	9
7	Понятие о регрессионных моделях	4	2	2	2	10
8	Линейные смешанные модели	4	4	4	4	16
9	Введение в язык программирования R	-	2	2	2	6
10	Работа с matrix и data.frame	-	4	3	3	10
11	Объекты класса list	-	2	2	2	6
12	Визуализация данных	-	2	2	2	6
	ИТОГО	26	28	27	27	108
Форма контроля		Экзамен				

5.2. Содержание лекционного курса, практических занятий

5.2.1. Лекционный курс

Тема 1. Разнообразие признака. Описательная статистика.

Группировка данных. Показатели описательной статистики. Иерархия средних значений при характеристике выборочной совокупности. Усечение данных, выбросы, метод Тьюки. Оценка параметров генеральной совокупности по параметрам выборочной совокупности.

Тема 2. Виды распределений.

Нормальное (гауссовское) распределение. Биноминальное распределение и распределение Пуассона. Экспоненциальное распределение. Критерии оценки нормальности распределения (хи-квадрат, Колмогорова-Смирнова, Шапиро-

Уилка, Андерсона-Дарлинга).

Тема 3. Оценка достоверности различий между выборками.

Нулевая гипотеза. Параметрические и непараметрические критерии. Критерии сравнения (Стьюдента, Манна-Уитни, Колмогорова-Смирнова, хи-квадрат Пирсона). Поправка Бонферони. Сравнение двух выборочных совокупностей. Сравнение ожидаемых и эмпирических распределений и двух эмпирических распределений.

Тема 4. Методы оценки качественных признаков.

Признаки с альтернативной изменчивостью. Вариация качественного признака (p, Sp). Критерий хи-квадрат для признаков с альтернативной изменчивостью. Поправка Йетса.

Тема 5. Оценка взаимосвязи между признаками.

Сопряжённость частот анализируемых признаков, коэффициент корреляции Пирсона. Непараметрические методы, коэффициент корреляции Фехнера. Корреляция рангов, коэффициенты Спирмена и Кендала.

Тема 6. Значимость группирующего фактора.

Дисперсионный анализ. Внутригрупповая, межгрупповая и общая дисперсии. Оценка разности дисперсий, критерии Бартлетта и Кохрена. Критерий Краскелла-Уоллиса.

Тема 7. Понятие о регрессионных моделях.

Виды регрессионных моделей. Построение уравнения прямолинейной регрессии. Методы отбора переменных в регрессионные модели: метод прямого отбора (Forward selection), метод обратного исключения (Backward Elimination) и метод последовательного отбора (Stepwise). Сравнение моделей, критерии Акаике и Шварца.

Тема 8. Линейные смешанные модели.

Определение и структура модели. Факторы и ковариантны. Подбор предикторов. BLUP, наследственная ковариационная матрица. Операции с матрицами. Методы решения системы линейный уравнений: метод Гаусса, Крамера и матричный метод. Пример решения системы уравнений с использованием биологических данных. Решение систем линейных уравнений с помощью специальных компьютерных программ.

5.2.2. Перечень практических занятий

Тема 1. Введение в язык программирования R.

Классы объектов. Типы данных. Функции. Пакеты. Условные операторы. Сортировка, рандомизация, поиск дубликатов записей в объектах класса vector.

Тема 2. Работа с matrix и data.frame.

Особенности сортировки и адресации в *matrix и data.frame.* Импорт и экспорт данных.

Тема 3. Объекты класса list

Особенности адресации. Преобразование в объекты класса vector и data.frame.

Тема 4. Визуализация данных.

Базовые инструменты визуализации: hist, plot, barplot, qqplot, boxplot. Продвинутая визуализация, графические библиотеки *ggplot2 и car*

Тема 5. Расчёт описательной статистики в языке программирования R. Достоверность различий между выборками.

Функции расчета показателей описательной статистики, summary. Пакет psych, функция describe. Критерии достоверности. Пакет dplyr.

Тема 6. Анализ распределения признака.

Функции распределения. Сравнение фактического распределение с нормальным. Пакет nortest.

Тема 7. Оценка связи между признаками. Влияние группирующего фактора.

Функция cor, cor.test, corrsign. Функции aov, kruskal-test.

Тема 8. Линейные и нелинейные модели в R.

Структура модели. Оценка качества модели. Поиск оптимальной модели, критерии Акаике и Шварца, метод anova.

Тема 9. Решение систем линейных уравнений.

Операции с матрицами. Методы решения системы линейный уравнений: метод Гаусса, Крамера и матричный метод. Пример решения системы уравнений с использованием биологических данных. Решение систем линейных уравнений с помощью специальных компьютерных программ.

6. Ресурсное обеспечение.

Кафедра ветеринарной генетики и биотехнологии располагает кадровыми ресурсами, гарантирующими качество подготовки аспиранта по дисциплине «Математические и статистические методы обработки научных исследований» в соответствии с ФГТ.

7. Материально-техническое обеспечение дисциплины

Для введения занятий по дисциплине «Математические и статистические методы обработки научных исследований» используются различные виды образовательных технологий, которые предусматривают использование материально-технического оборудования. При этом материально-техническое обеспечение дисциплины включает:

Таблица 3

№ п/п	Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
1.	Занятия лекционного типа проводятся в лекционной аудитории № НК 302	Компьютерный класс: аудитория для лабораторных, практических занятий, самостоятельной работы, дипломного и курсового проектирования ПО (выполнения курсовых работ)
2.	Практические занятия проводятся в аудитории № НК 302	Компьютерный класс: аудитория для лабораторных, практических занятий, самостоятельной работы, дипломного и курсового проектирования ПО (выполнения курсовых работ)

3.	Помещение для самостоятельной работы № НК 302	Компьютерный класс: аудитория для лабораторных, практических занятий, самостоятельной работы, дипломного и курсового проектирования ПО (выполнения курсовых работ)

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины: «2.1.5 Математические и статистические методы обработки научных исследований»

8.1. Список основной литературы

1. Федотова, Е. Л. Информационные технологии в науке и образовании : Учебное пособие / Е.Л. Федотова, А.А. Федотов. — Москва : ФОРУМ : ИНФРА-М, 2021. – 335 с. – (Высшее образование). - ISBN 978-5-8199-0884-6. Текст : электронный. - URL: https://znanium.com/catalog/product/1588599

2. Камалдинов, Е.В. Методы обработки экспериментальных данных и математического моделирования процессов: учебное пособие, 3-е изд., доп./ сост.: Е.В. Камалдинов, С.Г. Куликова, М.Л. Кочнева; Новосиб. гос. аграр. ун-т. – Новосибирск, 2021. – 158 с. [Электронный ресурс библиотеки ФГБОУ ВО «НГАУ»]

8.2. Список дополнительной литературы

1. Абрамян, М.Э. Инструменты и методы разработки электронных образовательных ресурсов по компьютерным наукам : монография / М.Э. Абрамян ; Южный федеральный университет. — Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2018. — 260 с.- ISBN 978-5-9275-2785-4. - Текст : электронный. - URL: https://znanium.com/catalog/product/1039686

2. Мешалкин, В. П. Основы информатизации и математического моделирования экологических систем : учебное пособие / В. П. Мешалкин, О. Б. Бутусов, А. Г. Гнаук. — Москва : ИНФРА-М, 2020. — 357 с. — (Высшее образование). - ISBN 978-5-16-009747-3. - Текст : электронный. - URL: https://znanium.com/catalog/product/1111403

3. Дадян, Э. Г. Методы, модели, средства хранения и обработки данных : учебник / Э.Г. Дадян, Ю.А. Зеленков. — Москва : Вузовский учебник : ИНФРА-М, 2018. — 168 с. - ISBN 978-5-9558-0490-3. - Текст : электронный. - URL: https://znanium.com/catalog/product/1834412

9. Перечень ресурсов информационно-телекоммуникационной сети «интернет», необходимых для освоения дисциплины

Таблица 4 – Перечень информационных ресурсов

№ п/п	Наименование	Адрес
-	Камалдинов Е.В. Использование R	http://www.rbiostats.blogspot.ru

	для обработки экспериментальных данных	
-	The R Development Core Team. The R Journal	http://www.r-project.org/doc/Rnews/bib/Rnewsbib.html
-	Графический метод решения задач линейного программирования	http://ru.wikipedia.org/wiki/Графический_метод_решения_задачи_линейного_программирования
-	*Вялый М.Н. Линейные неравенства и комбинаторика.*	*http://www.mccme.ru/dubna/2001/material/vyalyi.pdf*
-	Метод Симплекс.	http://ru.wikipedia.org/wiki/Симплекс-метод

10. Перечень лицензионного программного обеспечения:

Таблица 5 – Перечень лицензионного программного обеспечения

№ п/п	Наименование	Тип лицензии или номер лицензии
1	Браузер MozillaFireFox	Свободно-распространяемая
2	Rstudio	Свободно-распространяемая
3	R forWindows	Свободно-распространяемая
4	Zotero	Свободно-распространяемая
5	Mendeley	Свободно-распространяемая
9	LibreOffice	Свободно-распространяемая
10	OpenOffice	Свободно-распространяемая

11. Оценочные средства

Оценочные средства представлены в Приложение 1.