

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение
высшего образования
«ТЮМЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
АГРАРНЫЙ ИНСТИТУТ



УТВЕРЖДАЮ

Исполнительный директор
программы развития,
руководитель научно-
исследовательской политики

Г.Н. Суворов

«25» июня 2026 г.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
ПО ДИСЦИПЛИНЕ
**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКИЕ МЕТОДЫ
ОБРАБОТКИ ИНФОРМАЦИИ В БИОЛОГИИ**
Научная специальность 4.2.5. Разведение, селекция, генетика и биотехнология животных

1. Паспорт оценочных материалов по дисциплине

№ п/п	Темы дисциплины (модуля) / Разделы (этапы) практики* в ходе текущего контроля, вид промежуточной аттестации (с указанием семестра)	Код и содержание компетенции	Оценочные материалы (виды и количество)
1	2	3	4
1.	Цифровая биология: от кода к жизни.	ПК-13 - способность к выведению новых и совершенствованию существующих пород, типов, линий, семейств сельскохозяйственных животных,	Проработка лекционного материала, подготовка к лабораторно-практическим занятиям.
2.	Описательная статистика, сравнение средних по критерию Стьюдента при равенстве и неравенстве дисперсий.	оценке результативности племенной работы при разработке селекционных программ; ПК-14 - способность к совершенствованию и разработке новых методов оценки племенных и продуктивных качеств	Проработка лекционного материала, подготовка к лабораторно-практическим занятиям.
3.	Корреляционный, регрессионный и дисперсионный анализ.	сельскохозяйственных животных с использованием молекулярно-генетических маркеров и селекционно-генетических параметров;	Проработка лекционного материала, подготовка к лабораторно-практическим занятиям. Индивидуальное задание.
4.	Многомерные методы анализа и их визуализация.	ПК-15 - готовность к совершенствованию существующих и разработке новых биотехнологических методов воспроизводства и селекции сельскохозяйственных животных.	Проработка лекционного материала, подготовка к лабораторно-практическим занятиям. Индивидуальное задание.
5.	Зачёт (4 семестр)		Подготовка к зачёту в форме устного собеседования.

2. Виды и характеристика оценочных средств

Проработка лекционного материала включает:

- прочтение конспекта лекций и рекомендованной литературы;
- анализ теоретического материала, предложенного в лекциях;
- ознакомление с материалами по соответствующей тематике из рекомендуемых источников;
- составление опорного конспекта или ментальной карты для систематизации материала.

Форма контроля: устный опрос.

Подготовка к лабораторно-практическим занятиям включает:

- обращение к конспектам лекций по соответствующему вопросу и рекомендуемым источникам для уточнения терминологии;
- анализ методов решения задач, рассмотренных в лекциях.

Форма контроля: оценка выполненных лабораторно-практических заданий.

Индивидуальное задание напрямую вытекает из исследовательской проблемы, над которой работает аспирант.

Форма контроля: отчет по результатам самостоятельной работы.

Зачет проводится в форме устного собеседования. Каждый билет содержит два вопроса. Преподаватель вправе задать уточняющий вопрос по каждому вопросу билета.

Оценка «зачтено» выставляется, если аспирант:

- демонстрирует знание основных понятий, методов и инструментов дисциплины;
- понимает логику подготовки и обработки данных в общественных науках;
- умеет выбирать базовые математические и статистические методы в соответствии с задачей исследования;
- допускает отдельные неточности, но они не искажают общий смысл ответа и не препятствуют решению практической задачи.

Оценка «не зачтено» выставляется, если аспирант:

- не знает или слабо знает основные понятия и методы дисциплины;
- не может объяснить принципы обработки данных и выбор методов анализа;
- показывает фрагментарные, несистемные или неверные представления по ключевым вопросам дисциплины.

3. Оценочные средства

Вопросы для устного опроса по лекционному материалу

Тема 1. Цифровая биология: от кода к жизни.

1. Как развитие информационных технологий повлияло на прогресс в биологии?
2. Почему биологические системы (например, ДНК) можно считать «цифровыми»? Приведите примеры дискретных биологических кодов.
3. В чём состоит аналогия между генетическим кодом и компьютерным кодом?
4. Что подразумевается под термином «цифровая биология»?
5. Как используются методы машинного обучения и искусственного интеллекта в анализе биологических данных (например, при расшифровке геномов)?
6. Что такое синтетическая геномика и как она связана с цифровой биологией?
7. Какие перспективы открывает цифровая биология для сельского хозяйства и производства продуктов питания?
8. Может ли цифровая биология привести к появлению новых форм биологического неравенства или биотерроризма? Какие меры могут это предотвратить?
9. Как изменится роль биолога в эпоху цифровой биологии? Станет ли биолог больше похож на инженера или программиста?
10. Какие прорывные открытия в области цифровой биологии можно ожидать в ближайшие 10–20 лет?

Тема 2. Описательная статистика, сравнение средних по критерию Стьюдента при равенстве и неравенстве дисперсий

1. Что такое описательная статистика? Какова её роль в первичном анализе данных?
2. Перечислите и кратко охарактеризуйте основные цели первичного анализа данных.
3. Сравните среднее арифметическое, медиану и моду. В каких случаях предпочтительнее использовать каждый из этих показателей?
4. Что показывают показатели вариации (размах, дисперсия, стандартное отклонение, стандартная ошибка)? Как они помогают в интерпретации данных?
5. Что такое асимметрия и эксцесс? Как эти показатели помогают анализировать форму распределения?
6. Как выявить выбросы в данных? Какие методы обработки выбросов существуют?

7. Приведите пример использования описательной статистики для первичной интерпретации эмпирических данных в биологии.
8. Как проверить, соответствует ли распределение данных нормальному? Какие тесты и графические методы для этого существуют?
9. Для чего применяется t-критерий Стьюдента? Сформулируйте нулевую (H_0) и альтернативную (H_1) гипотезы при сравнении двух средних.
10. Почему важно проверять равенство дисперсий перед сравнением средних?

Тема 3. Корреляционный, регрессионный и дисперсионный анализ

1. В чём отличие статистической связи от причинно-следственной связи? Приведите пример корреляции без причинности.
2. Какова задача корреляционного анализа? Какие коэффициенты корреляции вы знаете? В каких случаях применяется каждый из них?
3. Как интерпретировать силу и направление корреляционной связи? Приведите шкалу оценки силы связи и пример её применения.
4. Какие условия необходимы для корректного применения корреляционного анализа? Что делать, если эти условия не выполняются?
5. В чём назначение регрессионного анализа? Опишите процесс построения уравнения регрессии на конкретном примере.
6. Как выбрать зависимую и независимую переменные в регрессионной модели? Приведите пример с обоснованием выбора.
7. Что показывает коэффициент детерминации (R^2)? Как его интерпретировать?
8. Как оценить значимость регрессионной модели и адекватность аппроксимации?
9. В чём суть дисперсионного анализа (ANOVA)? Для каких задач он применяется?
10. Приведите пример интерпретации результатов корреляционного, регрессионного и дисперсионного анализа в контексте исследования в биологии.

Тема 4. Многомерные методы анализа и их визуализация

1. Что такое многомерный анализ данных? Почему он особенно важен для биологических исследований?
2. Как применяется анализ главных компонент в биологии? Приведите пример.
3. Что такое иерархический кластерный анализ? В чём его ключевое отличие от неиерархических методов (например, k-means)?
4. Что такое дендрограмма? Как её читать и интерпретировать? Как определить оптимальное число кластеров по дендрограмме?
5. Как влияет нормализация данных на результаты иерархической кластеризации? Приведите пример, когда без нормализации результат будет некорректен.
6. В чём принципиальное отличие факторного анализа от анализа главных компонент? Когда предпочтительнее использовать факторный анализ?
7. Какова основная цель дискриминантного анализа? Для чего биологи применяют дискриминантный анализ? Приведите пример (идентификация видов и т.д.).
8. Как выбрать оптимальный тип визуализации для представления результатов вашего исследования? Опишите критерии выбора на конкретном примере.
9. Что такое «биplot» (biplot) и как его читают в контексте биологических данных?
10. Как многомерные методы помогают анализировать данные?

Пример лабораторно-практического занятия

Тема: «Дискриминантный анализ средствами электронной таблицы»

Продолжительность: 90 минут

Цель: сформировать практические навыки применения дискриминантного анализа для решения биологических задач

Подготовительный этап

Аспиранты подготавливают данные по результатам своих исследований (в формате Excel или CSV)

Задание перед занятием:

1. Прочитать условия задачи (дана в качестве примера).

Задача: измеряли листья двух популяций растений (популяция А – в зоне промышленного загрязнения, популяция В – в чистой зоне). Необходимо выяснить, влияет ли промышленное загрязнение на размер листьев. Если влияет, то выработать правило, по которому можно (если вообще можно!) использовать размер листьев в качестве индикатора загрязнения.

***Основная идея.** Имеется две или больше групп объектов, у каждого объекта измерен набор показателей. Ни один из показателей в отдельности не позволяет отнести объект к какой-то группе. Необходимо выяснить, можно ли по набору показателей отнести объект к той или иной группе. Если да, то выяснить, все ли показатели реально нужны. Построить математическую функцию, позволяющую отнести объект к той или другой группе по набору показателей.*

2. Аспирантом нужно ознакомиться таблицей с алгоритмом решения поставленной задачи.

Ход занятия

Часть 1. Вводная часть (15 мин)

1. Краткий опрос по теории:

- ✓ Какова основная цель дискриминантного анализа?
- ✓ Когда целесообразно применять дискриминантный анализ в биологических исследованиях? Приведите 3–4 примера из разных областей биологии.
- ✓ Какие типы данных (переменных) необходимы для проведения дискриминантного анализа? Каковы требования к зависимой и независимым переменным?
- ✓ Как визуализировать результаты дискриминантного анализа?

2. Обсуждение и корректировка данных, подготовленных для анализа.

3. Обзор инструментов: демонстрация возможностей Р7-Офис/Яндекс Таблицы.

Часть 2. Практическая работа (50 мин)

Аспиранты выполняют задания индивидуально, преподаватель консультирует.

Задание 1. По исходным данным вывести описательную статистику.

Проанализировать полученные данные, среднее значение, размах варьирования и т.д.

Задание 2. Выполнить регрессионный анализ, проанализировать показатели:

- ✓ Значимость F – сделать вывод является ли наше уравнение регрессии в высшей степени значимо.
- ✓ Р-Значение для коэффициентов – проверить все ли коэффициенты в уравнении в высшей степени значимы. Незначимые исключить.

Задание 3. Составить уравнение множественной регрессии.

Используя полученное уравнение выполнить проверку. Повторить тот же регрессионный анализ, поставив галочку в пункте Остатки. Скопировать столбцы Предсказанное Код и Остатки в исходную таблицу.

Задание 4. Визуализация данных.

Построить точечную диаграмму по исходным данным и по полученным (Предсказанное Код). Сравнить результаты.

Задание 5. Интерпретация результатов.

Ответить на вопросы:

1. Какие выводы можно сделать на основе точечных диаграмм?
2. Было ли выполнено разделение данных (дискриминация)?

3. Почему ось «Предсказанное Код» называется осью дискриминации, а наше уравнение дискриминантной функцией?
4. Как интерпретировать полученный результат?

Часть 3. Обсуждение и рефлексия (25 мин)

1. Презентация результатов по 3 минуты.
 - уравнение множественной регрессии;
 - диаграммы;
 - выводы по исследовательским вопросам.
 - коллективное обсуждение полученных результатов.
2. Подведение итогов:
 - типичные ошибки при подготовке данных для анализа;
 - типичные ошибки при интерпретации значимых показателей и результатов исследования.

Пример индивидуального задания

Задание 1.

1. Подготовить данные ваших исследований для однофакторного дисперсионного анализа. Число вариантов – не меньше трех, число повторностей в каждом варианте – на свое усмотрение.
2. Провести однофакторный дисперсионный анализ, вывести описательную статистику, результаты оформить в текстовом редакторе в виде страницы для диссертации.
3. Добавить на отчетную страницу, сделанную в электронной таблице диаграмму с пределом погрешностей, таблицу с результатами дисперсионного анализа и вывод по полученным результатам.

Рекомендации по выполнению:

Каждое задание предполагает: сбор/использование данных, обработку с помощью ИТ-инструментов, статистический анализ, визуализацию результатов и научную интерпретацию.

Формат сдачи: отчет с описанием методики, расчётами, диаграммами и выводами + файлы с данными.

Вопросы к зачёту

1. Понятие информационных технологий и их роль в биологии. Основные направления развития современных ИТ в биологии.
2. Математические методы в биологии.
3. Высокопроизводительное секвенирование.
4. Синтетическая геномика: базы биологических данных.
5. Искусственный интеллект и машинное обучение в биологии.
6. Использование больших данных и искусственного интеллекта для обработки информации, прогнозирования и поддержки принятия решений.
7. Понятие описательной статистики и ее место в первичном анализе данных. Цели первичного анализа: выявление общей структуры данных, основных тенденций, разброса значений и возможных выбросов.
8. Основные формы представления результатов: таблицы частот, гистограммы, диаграммы, графики, диаграммы рассеяния.
9. Показатели центра распределения: среднее арифметическое, медиана, мода. Показатели вариации: размах, дисперсия, стандартное отклонение, стандартная ошибка.
10. Анализ формы распределения: асимметрия, эксцесс, наличие выбросов и отклонений от нормальности.

11. Применение описательной статистики для первичной интерпретации эмпирических данных в биологии.

12. Понятие статистической связи между переменными и отличие связи от причинности. Корреляционный анализ: задача оценки тесноты и направления связи между признаками, коэффициенты корреляции, интерпретация силы связи. Условия применения корреляционного анализа: типы данных, линейность или нелинейность связи, требования к объему выборки.

13. Регрессионный анализ: назначение, построение уравнения регрессии, выбор зависимой и независимой переменных, интерпретация коэффициентов модели. Оценка качества регрессионной модели: коэффициент детерминации, значимость модели, проверка адекватности аппроксимации.

14. Дисперсионный анализ: сравнение средних значений нескольких групп, выявление влияния фактора на результативный признак. Логика проверки статистических гипотез в дисперсионном анализе: нулевая и альтернативная гипотезы, уровень значимости, критерий Фишера. Интерпретация результатов анализа в биологии: выявление взаимосвязей, различий между группами и факторов, влияющих на исследуемый показатель.

15. Принципы визуализации данных: наглядность, точность, лаконичность, соответствие типу данных и цели анализа. Основные виды визуализации: гистограмма, линейный график, круговая диаграмма, диаграмма рассеяния, пузырьковая диаграмма, сетевые схемы.

16. Ошибки визуализации: перегруженность, искажение масштаба, неудачный выбор типа графика, отсутствие подписей и легенды.

17. Интерпретация визуально представленных данных: выявление тенденций, различий, выбросов, связей и аномалий. Правила научной интерпретации результатов.

18. Многомерные методы анализа: анализ главных компонент; кластерный, факторный и дискриминантный анализ.