ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«07» марта 2024 г.
Протокол №1

Рабочая программа дисциплины (модуля)
«Продвинутая статистика»

Направление подготовки: *02.03.01 Математика и компьютерные науки*

Направленность (профиль) подготовки: *Математика и компьютерные науки*

Квалификация (степень) выпускника: *бакалавр*

Форма обучения: *очная*

Срок освоения программы: *4 года*

Год набора: *2024*

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат: 0215B0C000D1B3FEB44A37EE631E0C27D4
Владелец: Ивашкевич Евгений Васильевич
Действителен: с 13.01.2026 до 13.04.2027

Электронный документ

Москва
2024

Содержание

1. **Краткая характеристика дисциплины (модуля)** 3
2. **Перечень планируемых результатов обучения** 4
3. **Тематический план** 7
4. **Содержание дисциплины (модуля)** 7
5. **Учебно-методическое обеспечение** 9
6. **Материально-техническое обеспечение** 9
7. **Методические и оценочные материалы** 11

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Продвинутая статистика» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – бакалавриат по специальности 02.03.01 Математика и компьютерные науки, профиль Математика и компьютерные науки, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 807 от 23.08.2017 года.

Изучение дисциплины (модуля) «Продвинутая статистика» позволяет студентам эффективно анализировать сложные данные и выявлять закономерности, что критически важно для научных, математических исследований и бизнес-анализа. Кроме того, навыки, полученные в рамках курса, помогают специалистам принимать более информированные решения на основе количественных данных.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки бакалавриата по направлению 02.03.01 Математика и компьютерные науки, профиль Математика и компьютерные науки и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений.

Дисциплина (модуль) изучается на 4 курсе в 7 или 8 семестре на выбор.

Цель изучения дисциплины (модуля): заключается в углубленном освоении студентами методов статистического анализа и интерпретации данных для принятия обоснованных решений в различных областях науки и практики.

Задачи изучения дисциплины (модуля):

— изучить основные концепции и методы продвинутой статистики, необходимые для работы с данными в искусственном интеллекте;

— научиться интегрировать статистические методы в процесс разработки и оптимизации алгоритмов;

— освоить техники статистической валидации и оценки, а также интерпретации результатов анализа данных для принятия решений в проектах ИИ.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

— формирование знаний основ параметрической статистики и методы оценивания, методы проверки статистических гипотез, основные концепции непараметрической статистики, продвинутые статистические методы;

уметь:

— применять методы точечного и интервального оценивания;

— проводить и интерпретировать проверки гипотез;

— использовать линейные модели для анализа данных;

— планировать и реализовывать бутстрап-оценок и A/B тестов;

владеть:

— эффективным использованием методов параметрического оценивания;

— проведением статистических проверок и анализ результата;

— практическими навыками работы с линейными моделями;

— применением сложных статистических методов в практических исследованиях.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-1.	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1.	Знает методы поиска и анализа информации в области искусственного интеллекта, основные принципы критической оценки источников информации и их релевантности
		УК-1.2.	Умеет критически оценивать источники информации и синтезировать данные из различных источников для решения задач, применять системный подход к анализу и решению комплексных проблем
		УК-1.3.	Имеет практический опыт работы с современными инструментами и технологиями для обработки информации, формулировании и структурировании задач на основе полученной информации
УК-2.	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	УК-2.1.	Знает действующие правовые нормы, регулирующие деятельность в области решения задач, основные методы и подходы к определению круга задач
		УК-2.2.	Умеет определять круг задач в рамках поставленной цели, выбирать оптимальные способы решения задач, учитывая имеющиеся ресурсы и ограничения
		УК-2.3.	Имеет практический опыт применения знаний о правовых нормах и ресурсах в реальных ситуациях, разработки и реализации решений в соответствии с установленными ограничениями
ОПК-1.	Способен консультировать и использовать фундаментальные знания в области математического анализа, комплексного и функционального анализа алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений, дискретной математики и	ОПК-1.1.	Знает основные концепции и теории в области математического анализа и смежных дисциплин; методы и подходы, используемые в различных областях математики
		ОПК-1.2.	Умеет применять математические методы для решения профессиональных задач
		ОПК-1.3.	Имеет практический опыт разработки и реализация математических моделей в профессиональной деятельности

	математической логики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики в профессиональной деятельности		
ОПК-4.	Способен находить, анализировать, реализовывать программно и использовать на практике математические алгоритмы, в том числе с применением современных вычислительных систем	ОПК-4.1.	Знает базовые основы современного математического аппарата, связанного с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности
		ОПК-4.2.	Умеет использовать этот математический аппарат в профессиональной деятельности
		ОПК-4.3.	Имеет практический опыт применения современного математического аппарата, связанного с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности
ПК-1.	Способен формулировать задачи с математической точностью, обосновывать утверждения строго и анализировать полученные результаты в области математики и компьютерных наук	ПК-1.1.	Знает методы и подходы к формулированию задач, а также основные принципы математического доказательства и анализа результатов
		ПК-1.2.	Умеет корректно ставить и формулировать математические задачи, применять строгие методы доказательства и анализировать полученные результаты
		ПК-1.3.	Имеет опыт работы с задачами в области математики и компьютерных наук, включая применение математических методов для решения практических задач
ПК-2.	Способен решать типовые задачи профессиональной деятельности в области искусственного интеллекта, опираясь на информационную и библиографическую культуру, используя информационно-коммуникационные технологии и учитывая основные требования информационной	ПК-2.1.	Знает основы информационной и библиографической культуры, а также принципы информационной безопасности и применения информационно-коммуникационных технологий в профессиональной деятельности
		ПК-2.2.	Умеет эффективно использовать информационно-коммуникационные технологии для решения стандартных задач профессиональной деятельности, учитывая требования

	безопасности		информационной безопасности
		ПК-2.3.	Имеет опыт работы с информационными ресурсами и технологиями в области искусственного интеллекта, включая соблюдение норм информационной безопасности
ПК-3.	Способен применять методы математического и алгоритмического моделирования для решения как теоретических, так и практических задач в рамках профессиональной деятельности	ПК-3.1.	Знает основные методы математического и алгоритмического моделирования, а также их применение для решения теоретических и прикладных задач
		ПК-3.2.	Умеет применять методы математического и алгоритмического моделирования для анализа и решения различных задач в области математики и компьютерных наук
		ПК-3.3.	Имеет опыт использования методов математического и алгоритмического моделирования при решении теоретических и прикладных задач в профессиональной деятельности

3. Тематический план

№п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Контактная работа		Контроль	Самостоятельная работа	
		Лекции	Семинарские (практические занятия)			
1	Параметрическая статистика: точечное и интервальное оценивание	6	12		20	Домашние задания Подготовка к семинару
2	Параметрическая статистика: интервальное оценивание. Параметрическая статистика: проверка гипотез	6	12		20	Домашние задания, Контрольная работа
3	Непараметрическая статистика	6	12		18	Домашние задания Подготовка к семинару
4	Линейные модели	6	12		18	Домашние задания, контрольная работа
5	Современные методы статистики	6	12		18	Домашние задания Подготовка к семинару
	Зачет с оценкой			6		
	Итого:	***30***	***60***	***6***	***94***	
Объем дисциплины (модуля) (в ак. ч.)		***190***				
Объем дисциплины (модуля) (в зач. ед.)		***5***				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Параметрическая статистика: точечное и интервальное оценивание	Основные определения и классификация статистических данных. Понятия смещения, эффективности и состоятельности оценок. Метод максимального правдоподобия (MLE): Теоретические основы и алгоритм нахождения MLE. Примеры расчетов для различных распределений (нормальное, биномиальное). Метод моментов: Определение и процедура построения оценок. Сравнительный анализ с методом максимального правдоподобия. Оценка смещенности и дисперсии оценок. Асимптотические свойства оценок: Центральная предельная теорема. Закон больших чисел. Эффективность и байесовская оценка. Эффективность оценок: Информационная матрица Фишера. Нижняя граница Рао-Крамера. Байесовский

		подход к оценке параметров: Апостериорные плотности и выбор априорных распределений. Эмпирический байесовский подход.
2	Параметрическая статистика: интервальное оценивание. Параметрическая статистика: проверка гипотез	Концепция доверительных интервалов. Точные и асимптотические доверительные интервалы. Расчёт доверительных интервалов для среднего значения и доли: Нормальное распределение. Биномиальное распределение. Доверительные интервалы для дисперсии: Хи-квадрат распределение. F-распределение для отношения двух дисперсий. Доверительные границы для различий и отношений параметров. Устойчивость и применение доверительных интервалов. Принцип устойчивости доверительных границ. Практики построения и применения доверительных интервалов: Реальные кейсы и примеры. Проблемы и особенности применения. Основы проверки статистических гипотез. Основные понятия: нулевая и альтернативная гипотеза. Ошибки I и II рода, уровень значимости (α), мощность критерия (β). Простейшие тесты для средних значений и долей: Z-тест, T-тест Стьюдента. Сложные случаи проверки гипотез. Анализ мощности критериев. Проблема множественного тестирования: Поправка Бонферрони, Последовательные тесты Вальда. Одно- и двусторонние критерии. Решение комплексных задач на проверку гипотез. Особенности проведения и интерпретации результатов: Кейсы из практики. Возможные ошибки и способы их избегания.
3	Непараметрическая статистика	Основы непараметрической статистики. Общие принципы непараметрических методов. Ранговая статистика: Коэффициенты ранговой корреляции (Спирмена, Кендалла). Ранговые суммы Вилкоксона. Критерий знаков и критерий Манна-Уитни. Критерий Колмогорова-Смирнова. Хи-квадрат тест Пирсона. Критерий Андерсона-Дарлинга. Непараметрические аналоги ANOVA: Критерий Крускал-Уоллиса. H-статистика Фридмана.
4	Линейные модели	Непараметрическая регрессия: Локально-взвешенная регрессия LOESS. Сплайновая интерполяция. Ковариационные методы: Частичная корреляция. Каноническая корреляция.
5	Современные методы статистики	Классическая линейная регрессия: Модель $Y=X\beta+\varepsilon$, Метод наименьших квадратов. Предпосылки классической линейной модели: Нормальность остатков, Гомоскедастичность. Статистический анализ коэффициентов регрессии: T-тесты и F-тесты, Прогнозирование и доверительные интервалы прогноза. Идея бутстрап-метода. Процедуры бутстрап-выборки. Построение бутстрап-доверительных интервалов. Применение бутстрапа для оценки точности статистических выводов. Преимущества и недостатки бутстрап-подхода. Основы планирования экспериментов. Полностью рандомизированные планы. Балансировка факторов и контроль смешанных эффектов. А/Б тестирование: Постановка эксперимента. Анализ результатов теста. Вычисление размера выборки и мощности теста.

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Энатская, Н. Ю. Математическая статистика и случайные процессы : учебник для вузов / Н. Ю. Энатская. — Москва : Издательство Юрайт, 2025. — 191 с. — (Высшее образование). — ISBN 978-5-9916-9808-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/561150.

2. Кремер, Н. Ш. Математическая статистика : учебник и практикум для вузов / Н. Ш. Кремер. — 5-е изд. — Москва : Издательство Юрайт, 2025. — 259 с. — (Высшее образование). — ISBN 978-5-534-01654-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/561039.

3. Малугин, В. А. Математическая статистика : учебник для вузов / В. А. Малугин. — Москва : Издательство Юрайт, 2025. — 218 с. — (Высшее образование). — ISBN 978-5-534-06965-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/563990.

4. Трофимов, А. Г. Математическая статистика : учебное пособие для вузов / А. Г. Трофимов. — 2-е изд. — Москва : Издательство Юрайт, 2025. — 182 с. — (Высшее образование). — ISBN 978-5-534-21177-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/559506.

Дополнительная литература:

1. Трофимов, А. Г. Математическая статистика : учебное пособие для вузов / А. Г. Трофимов. — 2-е изд. — Москва : Издательство Юрайт, 2025. — 182 с. — (Высшее образование). — ISBN 978-5-534-21177-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/559506.

2. Малугин, В. А. Математическая статистика : учебник для вузов / В. А. Малугин. — Москва : Издательство Юрайт, 2025. — 218 с. — (Высшее образование). — ISBN 978-5-534-06965-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/563990.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной

мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

— столами и стульями;
— компьютерной техникой;
— механическими калькуляторами;
— специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование"	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное

Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Продвинутая статистика» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания и контрольные работы, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удается, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар — это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Контрольная работа – письменная работа с набором задач, которые нужно решить за ограниченное время.

Цель контрольной работы - получить специальные знания по одной или нескольким темам дисциплины (модуля) и продемонстрировать навыки их практического применения.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины.

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Продвинутая статистика»

Оценивание уровня учебных достижений, обучающихся по дисциплине (модулю), осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме ***зачета с оценкой***, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

Десятибалльная оценка	**Пятибалльная оценка**	**Оценка за зачет**	**Общая характеристика результата обучения по дисциплине (модулю)**
10	Отлично	Зачтено	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину. Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы,
9	Отлично	Зачтено	
8	Отлично	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
7	Хорошо	Зачтено	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	Зачтено	
5	Удовлетворительно	Зачтено	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь
4	Удовлетворительно	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			минимальным набором методов исследования.
3	Не сдан	Не зачтено	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

Дисциплина (модуль) «Продвинутая статистика» оценивается следующим образом:

Активность	Вес	Описание
Домашние задания	20%	Набор задач по темам недели
Контрольные работы	50%	Письменная работа с набором задач, которые нужно решить за ограниченное время
Активность на семинаре	15%	Активная работа студента на семинаре, его ответы на вопросы преподавателя и участие в дискуссии
Зачет с оценкой	35%	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по курсу

Формула расчёта итоговой оценки по дисциплине (модулю) «Продвинутая статистика»: «0,2 × средне за домашние задания + 0,15 × среднее за активность на семинаре + 0,5 × среднее за контрольные работы + 0,35 × зачет с оценкой».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные задания для подготовки к семинарам

Параметрическая статистика: точечное и интервальное оценивание

1. Что такое параметрическая статистика и как она отличается от непараметрической?
2. Какие основные типы статистических данных вы знаете и как их классифицируют?
3. Объясните понятие смещения оценки и приведите пример смещенной и несмещенной оценки.
4. Что такое эффективность и состоятельность оценки? Как эти свойства влияют на выбор оценок?
5. В чем заключается метод максимального правдоподобия (MLE)? Опишите его теоретические основы.
6. Какой алгоритм используется для нахождения оценки методом максимального правдоподобия?
7. Приведите пример расчета MLE для нормального распределения.
8. Как рассчитать MLE для биномиального распределения? Покажите на примере.
9. Что такое метод моментов и как он применяется для построения оценок?
10. Сравните метод максимального правдоподобия и метод моментов: в чем их преимущества и недостатки?
11. Как оценить смещенность и дисперсию статистической оценки?
12. Что такое Центральная предельная теорема и как она применяется в интервальном оценивании?
13. Объясните Закон больших чисел и его роль в асимптотических свойствах оценок.

14. Что такое информационная матрица Фишера и как она связана с эффективностью оценок?
15. В чем заключается нижняя граница Рао-Крамера и как она используется для оценки эффективности?

Параметрическая статистика: интервальное оценивание. Параметрическая статистика: проверка гипотез

1. Что такое доверительный интервал и как он отличается от точечной оценки?
2. Какие виды доверительных интервалов существуют (точные и асимптотические)? Приведите примеры.
3. Как рассчитать доверительный интервал для среднего значения нормального распределения?
4. Объясните расчет доверительного интервала для доли на основе биномиального распределения.
5. Как построить доверительный интервал для дисперсии с использованием хи-квадрат распределения?
6. Что такое F-распределение и как оно применяется для доверительных интервалов отношения двух дисперсий?
7. Как рассчитать доверительные границы для различий и отношений параметров?
8. Что такое принцип устойчивости доверительных границ и как он влияет на их применение?
9. Приведите реальный кейс построения и применения доверительного интервала. Какие проблемы могут возникнуть?
10. Что такое нулевая и альтернативная гипотеза в проверке статистических гипотез?
11. Объясните ошибки I и II рода, уровень значимости (α) и мощность критерия (β).
12. Как работает Z-тест для проверки среднего значения? Приведите пример.
13. Опишите T-тест Стьюдента и его применение для малых выборок.
14. Что такое проблема множественного тестирования и как ее решает поправка Бонферрони?
15. Как интерпретировать результаты проверки гипотез в практических кейсах? Какие ошибки следует избегать?

Непараметрическая статистика

1. В чем заключаются основы непараметрической статистики и ее принципы?
2. Что такое ранговая статистика и как она применяется в анализе данных?
3. Объясните коэффициент ранговой корреляции Спирмена и его расчет.
4. Как рассчитывается коэффициент ранговой корреляции Кендалла?
5. Что такое ранговые суммы Вилкоксона и как они используются для сравнения двух групп?
6. Опишите критерий знаков и его применение в непараметрической статистике.
7. Как работает критерий Манна-Уитни для сравнения двух независимых выборок?
8. Объясните критерий Колмогорова-Смирнова для проверки распределений.
9. Что такое хи-квадрат тест Пирсона и как он применяется для категориальных данных?
10. Опишите критерий Андерсона-Дарлинга и его преимущества над другими тестами.
11. Как работает непараметрический аналог ANOVA — критерий Крускал-Уоллиса?
12. Что такое H-статистика Фридмана и когда она используется?
13. Приведите пример применения непараметрических методов в реальном анализе данных.
14. Какие преимущества и недостатки имеют непараметрические методы по сравнению с параметрическими?

15. Как выбрать подходящий непараметрический тест в зависимости от типа данных и гипотезы?

Линейные модели

1. Что такое непараметрическая регрессия и как она отличается от параметрической?
2. Объясните локально-взвешенную регрессию LOESS и ее алгоритм.
3. Как работает сплайновая интерполяция в непараметрической регрессии?
4. Что такое частичная корреляция и как она рассчитывается?
5. Опишите каноническую корреляцию и ее применение для анализа взаимосвязей между переменными.
6. Как строится классическая линейная регрессия по модели $Y = X\beta + \varepsilon$?
7. Что такое метод наименьших квадратов и как он применяется для оценки параметров регрессии?
8. Какие предпосылки классической линейной модели (нормальность остатков, гомоскедастичность) и почему они важны?
9. Как проводится статистический анализ коэффициентов регрессии с использованием t-тестов?
10. Что такое F-тест в линейной регрессии и как он проверяет значимость модели?
11. Как строить прогнозы и доверительные интервалы прогноза в линейной регрессии?
12. Приведите пример применения линейной регрессии в анализе данных.
13. Какие проблемы могут возникнуть в линейной регрессии (мультиколлинеарность, гетероскедастичность) и как их решать?
14. Как оценить качество линейной модели с помощью коэффициента детерминации R^2?
15. Сравните параметрическую и непараметрическую регрессию: когда использовать каждую?

Современные методы статистики

1. Как работает идея бутстрап-метода для оценки статистических параметров?
2. Опишите процедуры бутстрап-выборки и их шаги.
3. Как построить бутстрап-доверительные интервалы для оценки?
4. Приведите пример применения бутстрапа для оценки точности статистических выводов.
5. Какие преимущества и недостатки имеет бутстрап-подход по сравнению с традиционными методами?
6. Что такое планирование экспериментов и зачем оно нужно?
7. Объясните полностью рандомизированные планы в планировании экспериментов.
8. Как обеспечивается балансировка факторов и контроль смешанных эффектов в экспериментах?
9. Что такое А/Б тестирование и как его правильно поставить?
10. Как анализировать результаты А/Б теста: расчет эффекта и статистическая значимость?
11. Как вычислить размер выборки и мощность теста в А/Б экспериментах?
12. Приведите реальный кейс применения бутстрапа или А/Б тестирования в статистическом анализе.
13. Как бутстрап помогает в оценке дисперсии и других характеристик моделей?
14. Какие этические аспекты следует учитывать при проведении А/Б тестирования?
15. Сравните бутстрап и традиционные доверительные интервалы: когда бутстрап предпочтительнее?

Примерные домашние задания

Домашнее задание: Параметрическая статистика: точечное и интервальное оценивание

1. Определите, что такое точечное оценивание и приведите два примера точечных оценок.
2. Объясните, что такое доверительный интервал и как он рассчитывается. Приведите пример.
3. Рассчитайте 95% доверительный интервал для среднего значения выборки с данными: 5, 7, 8, 6, 9. Предположите, что стандартное отклонение известно.
4. Опишите методы построения доверительных интервалов для пропорций. Приведите пример расчета.
5. Обсудите, как размер выборки влияет на ширину доверительного интервала.

Домашнее задание: Параметрическая статистика: интервальное оценивание и проверка гипотез

1. Объясните, что такое нулевая и альтернативная гипотезы. Приведите примеры.
2. Опишите процесс проверки гипотез, включая выбор уровня значимости. Как он влияет на результаты?
3. Рассчитайте t-критерий для следующих данных: 12, 15, 14, 10, 13. Используйте уровень значимости 0.05.
4. Приведите пример, когда вы бы использовали Z-критерий вместо t-критерия. Объясните, почему.
5. Объясните ошибки первого и второго рода. Приведите примеры ситуаций, в которых они могут возникнуть.

Домашнее задание: Непараметрическая статистика

1. Определите, что такое непараметрическая статистика и когда она используется. Приведите примеры.
2. Объясните, как тест Манна-Уитни применяется для сравнения двух групп. Приведите пример.
3. Рассчитайте и интерпретируйте результаты знакового теста для следующего набора данных: 1, 2, 3, 4, 5. Используйте уровень значимости 0.05.
4. Опишите критерий Краскала-Уоллиса и в каких случаях его следует использовать.
5. Обсудите преимущества и недостатки непараметрических методов по сравнению с параметрическими методами.

Домашнее задание: Обобщение и применение

1. Сравните точечное и интервальное оценивание. В каких ситуациях лучше использовать каждое из них?
2. Объясните, как интервал может помочь в принятии решений в бизнесе или науке.

3. Приведите пример исследования, где использовались методы проверки гипотез, и обсудите результаты.

4. Рассмотрите ситуацию, в которой данные не соответствуют нормальному распределению. Как это повлияет на выбор методов анализа?

5. Обсудите, как современные методы статистики могут дополнить традиционные методы, включая как параметрические, так и непараметрические подходы.

Примерные задания по контрольной работе

Контрольная работа №1

1. Определите непараметрическую статистику и объясните, в каких ситуациях она предпочтительнее параметрической статистики.

2. Приведите пример непараметрического теста и опишите, как он применяется для анализа данных.

3. Объясните, как тест Манна-Уитни используется для сравнения двух независимых выборок. Приведите пример применения.

4. Рассчитайте и интерпретируйте результаты знакового теста для следующего набора данных: 3, 5, 6, 4, 7. Используйте уровень значимости 0.05.

5. Опишите критерий Краскала-Уоллиса и в каких случаях его следует использовать. Как он отличается от ANOVA?

6. Объясните, как тест Вилкоксона применяется для сравнения двух связанных выборок. Приведите пример.

7. Обсудите преимущества и недостатки непараметрических методов по сравнению с параметрическими методами.

8. Приведите пример реального исследования, в котором использовались непараметрические методы, и обсудите его результаты.

Контрольная работа №2

1. Определите линейную регрессию и объясните, как она используется для моделирования зависимостей между переменными.

2. Что такое простая линейная регрессия? Приведите формулу и объясните значения каждого параметра.

3. Рассчитайте коэффициенты линейной регрессии для следующего набора данных (x, y): (1, 2), (2, 3), (3, 5), (4, 4). Определите интерсепт и наклон.

4. Объясните, что такое множественная линейная регрессия и как она отличается от простой линейной регрессии.

5. Приведите пример, когда использование множественной линейной регрессии будет более подходящим, чем простая линейная регрессия.

6. Объясните, как интерпретировать коэффициенты множественной линейной регрессии. Что означает положительный и отрицательный коэффициент?

7. Обсудите, что такое мультиколлинеарность и как она может повлиять на результаты множественной линейной регрессии.

8. Приведите пример реального исследования, в котором использовалась линейная регрессия, и обсудите его результаты.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1	Укажите метод оценки параметров распределений, основанный на максимизации функции правдоподобия.	Метод максимального правдоподобия	УК-1
2	Укажите теорему, которая описывает поведение выборочных средних при увеличении размера выборки.	Центральная предельная теорема	УК-1
3	Укажите название матрицы, используемой для оценки эффективности оценок параметров.	Информационная матрица Фишера	УК-1
4	Укажите подход к оценке параметров, который использует априорные распределения и апостериорные плотности.	Байесовский подход	УК-2
5	Укажите распределение, используемое для построения доверительных интервалов для среднего при известной дисперсии.	Нормальное распределение	УК-2
6	Укажите тест, применяемый для проверки гипотез о средних значениях при неизвестной дисперсии.	Т-тест Стьюдента	УК-2
7	Укажите критерий, используемый для сравнения двух независимых выборок в непараметрической статистике.	Критерий Манна-Уитни	ОПК-1
8	Укажите метод, применяемый для оценки размера выборки в A/B тестировании.	Расчёт мощности теста	ОПК-1
9	Укажите алгоритм, используемый для локально-взвешенной регрессии в непараметрических моделях.	LOESS	ОПК-1
10	Укажите статистику, используемую в хи-квадрат тесте Пирсона для проверки согласия распределений.	Сумма квадратов отклонений	ОПК-4
11	Укажите метод, применяемый для оценки коэффициентов в классической линейной регрессии.	Метод наименьших квадратов	ОПК-4
12	Укажите процедуру, используемую в бутстрап-методе для построения доверительных интервалов.	Перевыборка с возвращением	ОПК-4
13	Укажите коэффициент, измеряющий ранговую корреляцию между двумя переменными.	Коэффициент Спирмена	ПК-1
14	Укажите непараметрический аналог ANOVA для сравнения нескольких групп.	Критерий Крускалла-Уоллиса	ПК-1
15	Укажите метод, используемый для выявления общих закономерностей в данных через частичную корреляцию.	Ковариационные методы	ПК-1
16	Укажите подход, применяемый для планирования экспериментов с балансировкой факторов.	Полностью рандомизированные планы	ПК-2
17	Укажите тест, используемый для проверки гипотез о долях в биномиальном распределении.	Z-тест	ПК-2
18	Укажите критерий, применяемый для проверки равенства дисперсий двух выборок.	F-распределение	ПК-2
19	Укажите метод, используемый для интерполяции данных в непараметрической регрессии.	Сплайновая интерполяция	ПК-3
20	Укажите поправку, применяемую для коррекции уровня значимости при множественном тестировании.	Поправка Бонферрони	ПК-3