
УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«25» мая 2026 г.
Протокол №2

**Рабочая программа дисциплины (модуля)
«Современные методы статистики»**

Направление подготовки: 02.03.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Прикладной искусственный интеллект

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Срок освоения программы: 4 года

Год набора: 2026

**Москва
2026**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения	5
3. Тематический план	10
4. Содержание дисциплины (модуля)	10
5. Учебно-методическое обеспечение	12
6. Материально-техническое обеспечение	12
7. Методические и оценочные материалы	14

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Современные методы статистики» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – бакалавриат по направлению 02.03.01 «Математика и компьютерные науки», профиль «Прикладной искусственный интеллект», утвержденный приказом Министерства науки и высшего образования Российской Федерации № 807 от 23.08.2017 года.

Изучение дисциплины (модуля) «Современные методы статистики» позволяет освоить строгие статистические методы, необходимые для корректного анализа данных в условиях реальной неопределённости, эндогенности и сложной структуры. Полученные знания и навыки формируют фундамент для проведения научных исследований, А/В тестирования, эконометрического моделирования и причинно-следственного анализа в науке и индустрии.

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 «Прикладной искусственный интеллект» и входит в вариативную часть Блока 1 формируемую участниками образовательных отношений как дисциплина по выбору.

Дисциплина (модуль) изучается на 3 курсе в 5 семестре. Доступна к изучению после успешного освоения дисциплины (модуля): «Математическая статистика».

Цель изучения дисциплины (модуля): сформировать у обучающихся глубокое понимание современных статистических методов, включая причинно-следственный вывод, асимптотические и байесовские подходы, а также методы оценки эффектов воздействия в условиях сложных данных и эндогенности.

Задачи изучения дисциплины (модуля):

- изучить базовые и продвинутые методы математической статистики, включая асимптотику, бутстрап, методы коррекции смещения и условия применимости оценок метода наименьших квадратов;
- освоить теоретические основы линейной регрессии, включая теоремы Гаусса-Маркова и Фриша-Ву-Ловелла, и научиться интерпретировать результаты оценивания;
- научиться выявлять и устранять смещение оценок, вызванное пропущенными переменными и эндогенностью, с применением причинно-следственных моделей и метода инструментальных переменных;
- освоить современные подходы к выбору моделей и оценке эффектов воздействия, включая двойное машинное обучение и анализ на причинных графах;
- развить навыки байесовского вывода с использованием МСМС-методов, фильтров Калмана и многочастичных фильтров для работы с моделями пространства состояний.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- базовые понятия и методы статистики: асимптотика, А/В тестирование, виды бутстрапа, метод Холла;
- теоремы линейного регрессионного анализа: Фриша-Ву-Ловелла, Гаусса-Маркова, свойства оценок МНК;
- причины смещения оценок: пропущенные переменные, эндогенность;
- причинно-следственные модели: модель рубина, причинные графы, критерий обходных путей;
- подходы к выбору моделей и проблемы неравномерной инференции по Либу-Потчеру;
- основы методов двойного машинного обучения;

— основы байесовской статистики и МСМС: ЕМ-алгоритм, сэмплер Гиббса, метод Метрополиса-Гастингса, модель стохастических блоков, фильтры Калмана и многочастичный.

уметь:

- применять асимптотические методы для построения доверительных интервалов и статистических тестов;
- использовать бутстрап для коррекции смещения и оценки неопределённости;
- оценивать простую и множественную регрессию методом наименьших квадратов и проверять условия теоремы гаусса-маркера;
- интерпретировать результаты регрессионного анализа и выявлять смещение из-за пропущенных переменных;
- применять метод инструментальных переменных (2SLS) для оценки эффектов воздействия в условиях эндогенности;
- строить байесовские А/В тесты с расчётом минимального детектируемого эффекта;
- применять ЕМ -алгоритм для решения задач классификации и кластеризации.

владеть:

- методами бутстрапа различных видов, включая метод Холла, для коррекции смещения и построения доверительных интервалов в сложных моделях;
- техникой двойного машинного обучения (двойное LASSO) для оценки причинных эффектов при большом числе контрольных переменных;
- причинно-следственным выводом на графах с использованием критерия обходных путей и моделью рубина для формализации эффектов воздействия;
- методами МСМС, включая сэмплер Гиббса и алгоритм Метрополиса-Гастингса, для байесовского вывода в моделях пространства состояний и стохастических блоковых моделях;
- фильтром Калмана и многочастичным фильтром для анализа и прогнозирования в моделях пространства состояний.

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:

Код и содержание компетенции	Индикатор компетенции и перечень планируемых результатов обучения по дисциплине (модулю)
ППК-Р4. Способен использовать базы данных при создании программных модулей и компонентов	ППК-Р4.1. Пишет программный код с использованием языков определения и манипулирования данными в базах данных (базовый) ППК-Р4.1. 3-1. Знает архитектуры современных систем управления баз данных, включая SQL и noSQL ППК-Р4.1. У-1. Умеет применять выбранные языки работы с базами данных
	ППК-Р4.2. Проектирует базы данных для программных модулей и компонентов (средний) ППК-Р4.2. 3-1. Знает современные подходы к проектированию реляционных и нереляционных баз данных ППК-Р4.2. У-2. Умеет использовать выбранную среду программирования для работы с данными в базе.
	ППК-Р4.3. Оптимизирует производительность работы с базами данных (средний) ППК-Р4.3. 3-1. Знает внутреннее устройство СУБД выбранной архитектуры ППК-Р4.3. У-2. Умеет вырабатывать варианты оптимизации производительности запросов в базе данных
ППК-Р5. Способен разрабатывать требования и проектировать программное обеспечение	ППК-Р5.1. Анализирует возможности реализации требований к компьютерному программному обеспечению (средний) ППК-Р5.1. 3-1. Знает возможности существующей программно-технической архитектуры ППК-Р5.1. У-3. Умеет вырабатывать варианты реализации требований к компьютерному программному обеспечению
	ППК-Р5.2. Разрабатывает технические спецификаций на программные компоненты и их взаимодействие (средний) ППК-Р5.2. 3-2. Знает методы и средства проектирования программных интерфейсов ППК-Р5.2. У-1. Умеет выбирать средства реализации требований к компьютерному программному обеспечению
	ППК-Р5.3. Проектирует компьютерное программное обеспечение (продвинутый) ППК-Р5.3. 3-1. Знает принципы построения и виды архитектуры компьютерного программного обеспечения ППК-Р5.3. У-2. Умеет проектировать структуры данных
	ППК-Р5.4. Выполняет логическое проектирование системы (продвинутый) ППК-Р5.4. 3-3. Знает базовые технологии взаимодействия и интеграции систем и компонентов ППК-Р5.4. У-1. Умеет описывать интеграции со смежными системами на логическом уровне
ППК-Р6. Способен участвовать в промышленной разработке программного обеспечения	ППК-Р6.1. Работает в соответствии с промышленными методологиями разработки (средний) ППК-Р6.1. 3-1. Знает принципы Agile и их применение в промышленных проектах ППК-Р6.1. У-2. Умеет работать в команде с использованием инструментов управления проектами

	ППК-Р6.2. Использует инструменты промышленной разработки (средний) ППК-Р6.2. 3-1. Знает принципы Continuous Integration and Continuous Delivery (CI/CD) ППК-Р6.2. У-1. Умеет настраивать потоки работ CI/CD
	ППК-Р6.3. Разрабатывает масштабируемый и поддерживаемый код (продвинутый) ППК-Р6.3. 3-1. Знает принципы чистого кода, SOLID, DRY, KISS и др. ППК-Р6.3. У-1. Умеет разрабатывать модульный и тестируемый программный код
	ППК-Р6.4. Участвует в развертывании и поддержке программного обеспечения (продвинутый) ППК-Р6.4. 3-1. Знает стратегии развертывания промышленного программного обеспечения ППК-Р6.4. У-1. Умеет развертывать приложения в облаке
	ППК-Р7. Способен применять искусственный интеллект (ИИ) для генерации и отладки программного кода
	ППК-Р7.1. Применяет ИИ-инструменты для генерации программного кода (базовый) ППК-Р7.1. 3-1. Знает принципы работы современных генеративных ИИ-моделей для генерации кода ППК-Р7.1. У-1. Умеет формулировать корректные текстовые запросы (промты) для генерации кода
	ППК-Р7.2. Использует ИИ для анализа и отладки кода (средний) ППК-Р7.2. 3-1. Знает методы ИИ-анализа кода ППК-Р7.2. У-2. Умеет интерпретировать рекомендации ИИ по исправлению кода
	ППК-Р7.3. Оптимизирует код с помощью ИИ (продвинутый) ППК-Р7.3. 3-1. Знает методы ИИ-оптимизации ППК-Р7.3. У-2. Умеет проверять корректность оптимизаций, предложенных ИИ
	ППК-Р7.4. Оценивает этические и профессиональные аспекты применения ИИ в разработке (продвинутый) ППК-Р7.4. 3-1. Знает этические нормы использования ИИ (конфиденциальность, плагиат кода и т.п.) ППК-Р7.4. У-2. Умеет документировать использование ИИ в разработке
ППК-ДА4. Способен выявлять бизнес-проблемы или бизнес-возможности	ППК-ДА4.1. Собирает информацию о бизнес-проблемах или бизнес-возможностях (средний) ППК-ДА4.1. 3-3. Знает инструменты, техники анализа бизнес-ситуации и предметной области ППК-ДА4.1. У-1. Умеет собирать, классифицировать, систематизировать и обеспечивать хранение и актуализацию информации для бизнес-анализа
	ППК-ДА4.2. Выявление истинных бизнес-проблем или бизнес-возможностей (продвинутый) ППК-ДА4.2. 3-2. Знает инструменты, техники анализа бизнес-ситуации и предметной области, включая анализ данных ППК-ДА4.2. У-1. Умеет выявлять и классифицировать бизнес-проблемы или бизнес-возможности
ППК-ДА5. Способен применять методы статистического анализа и теорию эксперимента для планирования, проведения и	ППК-ДА5.1. Разрабатывает дизайн эксперимента (базовый) ППК-ДА5.1. 3-1. Знает основные принципы планирования экспериментов ППК-ДА5.1. У-2. Умеет выбирать метрики для оценки эффекта воздействия

интерпретации результатов экспериментов	ППК-ДА5.2. Проводит статистический анализ данных эксперимента (средний) ППК-ДА5.2. 3-1. Знает методы проверки гипотез (t-тест, χ^2 , ANOVA) ППК-ДА5.2. У-2. Умеет интерпретировать p-value и уровень значимости
	ППК-ДА5.3. Интерпретирует результаты экспериментов (продвинутый) ППК-ДА5.3. У-2. Умеет формулировать рекомендации на основе статистических выводов
МО-2. Способен проектировать и внедрять системы автоматизации жизненного цикла машинного обучения (MLOps)	МО-2.1. Проектирует и настраивает CI/CD-пайплайны для машинного обучения (базовый) МО-2.1. 3-1. Знает принципы CI/CD применительно к моделям машинного обучения МО-2.1. У-1. Умеет настраивать CI/CD-пайплайны для МО с использованием инструментов
	МО-2.2. Организует версионирование моделей, данных и экспериментов (средний) МО-2.2. 3-1. Понимает системы версионирования моделей и данных МО-2.2. У-2. Умеет использовать инструменты MLOps для воспроизводимости и масштабируемости экспериментов
	МО-2.3. Внедряет системы мониторинга моделей в продуктивную эксплуатацию (средний) МО-2.3. 3-1. Знает метрики мониторинга моделей МО МО-2.3. У-1. Умеет разворачивать системы мониторинга моделей МО
	МО-2.4. Автоматизирует процессы переобучения и развертывания моделей (продвинутый) МО-2.4. 3-1. Знает стратегии автоматизации переобучения МО-2.4. У-1. Умеет автоматизировать переобучение моделей и A/B-тестирование развертываний
РГ-1. Способен руководить процессами разработки компьютерного программного обеспечения	РГ-1.1. Руководит разработкой программного кода (средний) РГ-1.1. У-1. Способен распределять задачи на разработку программного кода между исполнителями РГ-1.1. У-3. Оценивает качество и эффективность программного кода
	РГ-1.2. Руководит проверкой работоспособности компьютерного программного обеспечения (средний) РГ-1.2. У-2. Умеет оценивать результаты проверки работоспособности компьютерного программного обеспечения
	РГ-1.3. Руководит интеграцией программных модулей и компонентов (продвинутый) РГ-1.3. 3-2. Знает методы и программные интерфейсы взаимодействия компьютерного программного обеспечения с внешними программными компонентами
РГ-2. Способен организовать процессы разработки компьютерного программного обеспечения	РГ-2.1. Управляет проектированием компьютерного программного обеспечения (средний) РГ-2.1. 3-1. Знает принципы построения архитектуры компьютерного программного обеспечения и виды архитектуры программного обеспечения
	РГ-2.2. Управляет процессом разработки компьютерного программного обеспечения (продвинутый) РГ-2.2. У-1. Умеет составлять планы процесса разработки программного продукта

	РГ-2.3. Управляет запросами на изменения, дефектами и проблемами (продвинутый) РГ-2.3. 3-1. Знает методы и средства выявления дефектов, проблем и причин их возникновения в компьютерном программном обеспечении
	РГ-2.4. Управляет конфигурациями и выпусками программного продукта (продвинутый) РГ-2.5. У-1. Умеет управлять версиями отдельных компонентов и программного продукта в целом
	РГ-4. Способен организовать промышленную разработку программного обеспечения РГ-4.1. Внедряет промышленные методологии разработки (средний) РГ-4.1. 3-1. Знает принципы Agile и их применение в промышленных проектах
	РГ-4.2. Внедряет инструменты промышленной разработки (средний) РГ-4.2. 3-1. Знает принципы Continuous Integration and Continuous Delivery (CI/CD) РГ-4.3. Организует разработку масштабируемого и поддерживаемого кода (продвинутый) РГ-4.3. У-1. Умеет управлять разработкой модульного и тестируемого программного кода РГ-4.4. Организует развертывание и поддержку программного обеспечения (продвинутый) РГ-4.4. У-1. Умеет организовать развертывание приложения в облаке
РГ-5. Способен организовать процессы применения искусственного интеллекта (ИИ) для генерации и отладки программного кода	РГ-5.1. Организует применение ИИ-инструментов для генерации программного кода (средний) РГ-5.1. 3-1. Знает принципы работы современных генеративных ИИ-моделей для генерации кода
	РГ-5.2. Организует применение ИИ для анализа и отладки кода (продвинутый) РГ-5.2. 3-1. Знает методы ИИ-анализа кода
	РГ-5.3. Организует оптимизацию кода с помощью ИИ (продвинутый) РГ-5.3. 3-1. Знает методы ИИ-оптимизации
	РГ-5.4. Оценивает этические и профессиональные аспекты применения ИИ в разработке (продвинутый) РГ-5.4. 3-1. Знает этические нормы использования ИИ (конфиденциальность, плагиат кода и т.п.)
УК-4. (УНК-01.) Способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(ых) языке(ах) <i>(соответствует УНК-01 – Способен применять коммуникационные компетенции в профессиональной деятельности)</i>	УНК-01.01. Осуществляет коммуникации в профессиональной деятельности (базовый) УНК-01.01.01. У-1 Владеет навыками презентации и публичной дискуссии УНК-01.01.01. У-3 Способен формулировать и понимать технологические и бизнес-требования
	УНК-01.02. Демонстрирует владение профессиональной культурой (базовый) УНК-01.02.01. У-3. Умеет грамотно оформлять документацию и вести коммуникацию в соответствии с профессиональными стандартами
УК-3. (УНК-02.) Способен осуществлять социальное взаимодействие и реализовывать свою роль в команде	УНК-02.01. Принимает участие в групповом взаимодействии (базовый) УНК-02.01.01. У-5. Умеет работать в команде, проявлять инициативу и поддерживать коллег
	УНК-02.02. Проявляет лидерство и осуществляет

<i>(соответствует УНК-02 – Способен применять методы коллективной работы в профессиональной деятельности)</i>	наставничество (базовый) УНК-02.02.01. У-3. Умеет делиться знаниями и опытом, оказывать поддержку и осуществлять наставничество коллегам
---	--

3. Тематический план

№п/ п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостоятельная работа	
		Лекции	Практические занятия			
1	А/В -тестирование	4	4		16	Домашнее задание
2	Линейные модели и оценка эффектов воздействия в неэкспериментальных данных	18	18	2	76	Домашнее задание Контрольная работа
3	Байесовские методы	8	8		34	Домашнее задание
	Экзамен			2		
Итого:		30	30	4	126	
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	А/В -тестирование	Базовые понятия и методы статистики. Асимптотические методы для тестов и доверительных интервалов. Эффекты воздействия, как параметры распределения: модель Рубина. Экспериментальная оценка эффектов воздействия, А/В тестирование Бутстрап для коррекции смещения оценок и инференции. Метод Холла. Применение бутстрапа при А/В тестировании
2	Линейные модели и оценка эффектов воздействия в неэкспериментальных данных	Линейные модели с стохастическими регрессорами. Отличие прогнозной и причинно-следственной задачи при исследовании данных. Оценки МНК и проекции. Теорема Фриша-Ву-Ловелла Регрессия по всей популяции. Идентификация параметров линейной модели. Наилучший линейный предиктор. L2 пространство и его свойства. Проекция в L2 Предположения линейной модели и статистические свойства МНК оценок. Теорема Гаусса-Маркова. Тесты: t , F Асимптотические свойства оценок. Тесты параметров линейной регрессии: z -тест, тест Вальда. Мисспецификация модели: смещение из-за пропущенных переменных Контрольные переменные, инструментальные переменные. Выбор модели: секвенциальное тестирование, информационные критерии Неравномерная инференция по Либу-Потчеру Использование графических моделей для выбора регрессионной модели. Критерий обходных путей LASSO, post-LASSO, двойное LASSO Двойное машинное обучение

3	Байесовские методы	Постановка задачи байесовской статистики, примеры. Условные распределения, свойства нормальных векторов Модели пространства состояний: фильтр Калмана, многочастичный фильтр Сэмплер Гиббса, алгоритм Метрополиса-Гастингса, сходимость. Модель стохастических блоков Байесовское A/B тестирование, минимальный выявляемый эффект. Байесовские алгоритмы классификации и кластеризации: EM-алгоритм
---	--------------------	---

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. Мэрфи, К. П. Вероятностное машинное обучение: введение : практическое руководство / К. П. Мэрфи ; пер. с англ. А. А. Слинкина. - Москва : ДМК Пресс, 2023. - 990 с. - ISBN 978-5-93700-119-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2109489>.

2. Курт, У. Байесовская статистика: Star Wars, LEGO, резиновые уточки и многое другое : практическое руководство / У. Курт. - Санкт-Петербург : Питер, 2021. - 304 с. - (Библиотека программиста). - ISBN 978-5-4461-1655-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1817495>.

Дополнительная литература:

1. Брюс, П. Практическая статистика для специалистов Data Science / П. Брюс, Э. Брюс. - 2-е изд. - Санкт-Петербург : БХВ-Петербург, 2021. - 352 с. - ISBN 978-5-9775-6705-3. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2142261>.

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) «Современные методы статистики» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, контрольная работа, 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 удастся, 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 и продемонстрировать навыки их практического применения.

Бонусные баллы — это оценки, которые студенты могут получить за выполнение дополнительных заданий.

Электронный документ

Формат бонусных баллов позволяет студентам улучшить общую оценку по дисциплине (модулю) и стимулирует углубленное изучение материала.

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

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) «Современные методы статистики»

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 *экзамена*, 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, если иное не указано в syllabus дисциплины:

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	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, 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.
9	Отлично	
8	Отлично	
7	Хорошо	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
6	Хорошо	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
5	Удовлетворительно	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 минимальным набором методов исследования.
4	Удовлетворительно	
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задаст дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Современные методы статистики» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	30%	7	Набор задач по темам недели
Контрольная работа	20%	1	Письменная работа с набором задач, которые нужно решить за ограниченное время
Экзамен	50%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

В рамках изучения дисциплины (модуля) возможно получение бонусных баллов.

Формула расчёта итоговой оценки по дисциплине (модулю) «Современные методы статистики»: $0,3 \times \text{среднее за домашние задания} + 0,2 \times \text{контрольная работа} + 0,5 \times \text{экзамен}$.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1.

1. Объясните, в чём заключается разница между асимптотическими и бутстрап-подходами к построению доверительных интервалов в А/В тестировании;
2. Опишите, как модель Рубина формализует понятие эффекта воздействия и какие предположения необходимы для его идентификации;
3. Реализуйте бутстрап-оценку разности средних в А/В тесте и сравните полученный доверительный интервал с асимптотическим подходом;
4. Объясните, как метод Холла улучшает точность бутстрап-интервалов по сравнению с обычным процентильным методом;
5. Проанализируйте, в каких случаях стандартное А/В тестирование может давать смещённые оценки и как бутстрап помогает скорректировать неопределённость.

Домашнее задание 2.

1. Оцените модель линейной регрессии методом МНК и проверьте выполнение предпосылок теоремы Гаусса-Маркова на примере реальных или синтетических данных;
2. Примените теорему Фриша-Ву-Ловелла для интерпретации коэффициентов множественной регрессии как частичных эффектов;
3. Объясните, как пропущенная переменная, коррелирующая с регрессором, приводит к смещению оценки МНК, и продемонстрируйте это на примере;
4. Реализуйте оценку эффекта воздействия с помощью метода инструментальных переменных (2SLS) в случае эндогенности регрессора;
5. Примените двойное LASSO для отбора контрольных переменных и оценки причинного эффекта в условиях большого числа ковариат.

Домашнее задание 3.

1. Постройте байесовский A/B тест для биномиальных данных, используя сопряжённые априорные распределения, и рассчитайте вероятность превосходства одной группы над другой;
2. Реализуйте ЕМ-алгоритм для задачи кластеризации методом смеси нормальных распределений и проанализируйте сходимость;
3. Опишите, как работает сэмплер Гиббса в байесовской линейной регрессии, и реализуйте его на простом примере;
4. Примените фильтр Калмана для оценки скрытого состояния в модели случайного блуждания с наблюдениями с шумом;
5. Объясните, в каких случаях предпочтительнее использовать многочастичный фильтр вместо фильтра Калмана, и опишите его основные шаги: предсказание, взвешивание, ресэмплирование.

Примерные задания для контрольной работы

1. Объясните, в чём заключается различие между прогнозной и причинно-следственной интерпретацией модели линейной регрессии, и приведите пример, когда высокая предсказательная способность не означает корректной оценки эффекта воздействия;
2. Опишите, как модель Рубина формализует понятие потенциальных исходов и какие предположения (SUTVA, случайная выборка, отсутствие скрытой эндогенности) необходимы для идентификации среднего эффекта обработки (ATE);
3. Реализуйте A/B тест на синтетических данных: симулируйте две группы, оцените разность средних, постройте доверительный интервал с использованием асимптотического подхода и сравните с результатом бутстрапа;
4. Объясните, как метод Холла улучшает точность бутстрап-доверительных интервалов за счёт учёта смещения и асимметрии распределения оценки;
5. Оцените модель линейной регрессии методом МНК на примере набора данных и проверьте выполнение ключевых предпосылок теоремы Гаусса-Маркова: линейность, случайность выборки, отсутствие мультиколлинеарности, гомоскедастичность, ненормальность остатков;
6. Примените теорему Фриша-Ву-Ловелла для интерпретации коэффициента в множественной регрессии как результата двухэтапной проекции: сначала зависимой переменной, затем регрессора на пространство остальных предикторов;
7. Объясните, как пропущенная переменная, коррелирующая с включённым регрессором, приводит к смещению оценки МНК, и продемонстрируйте этот эффект на численном примере;
8. Реализуйте оценку эффекта воздействия с помощью метода инструментальных переменных (2SLS) в случае, когда регрессор эндогенен из-за обратной связи или скрытой переменной;

9. Опишите, как критерий обходных путей (backdoor criterion) используется в причинных графах для выбора корректного набора контрольных переменных, и приведите пример графа, где контроль по одной переменной устраняет смещение;

10. Примените технику двойного машинного обучения (double machine learning) с двойным LASSO для оценки причинного эффекта в условиях большого числа потенциальных ковариат, используя модель предсказания результата и модель предсказания воздействия.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция	Индикатор
1.	Назовите основную цель А/В-тестирования в контексте причинно-следственного вывода.	Оценка эффекта воздействия путём сравнения контрольной и тестовой групп	ППК-ДА5	ППК-ДА5.1.3-1
2.	Укажите, как в модели Рубина определяется потенциальный исход для индивида при отсутствии воздействия.	$Y_i(0)$ — результат, который был бы наблюдаем, если бы индивид не получил treatment	ППК-ДА5	ППК-ДА5.1.3-1
3.	Объясните, что означает предпосылка SUTVA в модели потенциальных исходов.	Отсутствие взаимодействия между индивидами и однозначность воздействия	ППК-ДА5	ППК-ДА5.1.3-1
4.	Назовите условие, при котором оценка АТЕ является несмещённой в рандомизированном эксперименте.	Случайное назначение воздействия и отсутствие систематических различий между группами	ППК-ДА5	ППК-ДА5.2.У-2
5.	Назовите асимптотический тест для проверки гипотезы о разности пропорций в А/В-тесте.	Z-тест для долей	ППК-ДА5	ППК-ДА5.2.3-1
6.	Объясните, зачем используется бутстрап при построении доверительных интервалов в условиях малой выборки.	Для аппроксимации распределения статистики без предположений о нормальности	ППК-ДА5	ППК-ДА5.2.У-2
7.	Укажите, в чём заключается преимущество метода Холла при построении бутстрап-интервалов.	Корректирует смещение и асимметрию, обеспечивая более точные покрытия	ППК-ДА5	ППК-ДА5.3.У-2
8.	Назовите проблему, возникающую при анализе данных без учёта множественной проверки гипотез.	Увеличение вероятности ложноположительного результата (Family-wise Error Rate)	ППК-ДА5	ППК-ДА5.2.3-1

9.	Назовите теорему, утверждающую, что МНК-оценки являются наилучшими линейными несмещёнными.	Теорема Гаусса-Маркова	ППК-Р5	ППК-Р5.3.3-1
10.	Объясните, как теорема Фриша-Ву-Ловелла помогает интерпретировать коэффициенты множественной регрессии.	Показывает, что коэффициент — это эффект после "очистки" от влияния других переменных	ППК-Р5	ППК-Р5.4.3-3
11.	Укажите, какое свойство оценок МНК нарушается при наличии эндогенности регрессора.	Несмещённость и состоятельность	ППК-Р5	ППК-Р5.3.3-1
12.	Назовите тест, используемый для проверки значимости группы коэффициентов в линейной модели.	F-тест	ППК-ДА5	ППК-ДА5.2.3-1
13.	Назовите метод, позволяющий оценить причинный эффект в условиях эндогенности с помощью дополнительной переменной.	Метод инструментальных переменных (IV)	ППК-Р5	ППК-Р5.3.У-2
14.	Объясните, как критерий обходных путей помогает выбрать набор контрольных переменных в причинном графе.	Определяет, какие переменные нужно контролировать, чтобы закрыть все backdoor-пути	ППК-Р5	ППК-Р5.4.У-1
15.	Назовите подход, позволяющий отбирать переменные и оценивать причинный эффект при большом числе ковариат.	Двойное машинное обучение с двойным LASSO	ППК-Р5	ППК-Р5.3.У-2
16.	Назовите метод, используемый для оценки параметров в моделях с пропущенными данными или латентными переменными.	ЕМ-алгоритм	ППК-Р5	ППК-Р5.3.3-1
17.	Назовите байесовский подход к А/В-тестированию, основанный на апостериорном распределении.	Байесовское А/В-тестирование с расчётом Posterior Probability	ППК-ДА5	ППК-ДА5.3.У-2
18.	Укажите, как сэмплер Гиббса используется для байесовского вывода в моделях со скрытыми переменными.	Путём последовательной выборки из условных распределений	ППК-Р5	ППК-Р5.3.У-2
19.	Назовите рекурсивный алгоритм для оценки скрытого состояния в линейной модели с гауссовским шумом.	Фильтр Калмана	ППК-Р5	ППК-Р5.3.3-1
20.	Назовите метод, используемый для байесовского вывода в нелинейных и не гауссовских моделях пространства состояний.	Многочастичный фильтр / Particle Filter	ППК-Р5	ППК-Р5.3.У-2
21.	Какой тип базы данных подходит для хранения результатов статистических экспериментов и метрик?	PostgreSQL / Time-series DB / MLflow Tracking / Experiment Database	ППК-Р4	ППК-Р4.2.3-1

22.	Какой метод используется для оптимизации запросов к данным для статистического анализа?	Индексирование / Partitioning / Caching / Materialized Views	ППК-Р4	ППК-Р4.3. У-2
23.	Какой принцип Agile применяется в разработке статистических экспериментов?	Scrum / Kanban / Iterative / Sprint Planning	РГ-4	РГ-4.1. 3-1
24.	Что означает аббревиатура CI/CD в контексте автоматизации статистических пайплайнов?	Continuous Integration / Continuous Delivery	МО-2	МО-2.1. 3-1
25.	Какой инструмент используется для версионирования экспериментов и статистических моделей?	MLflow / DVC / Weights & Biases / Experiment Tracker	МО-2	МО-2.2. У-2
26.	Какая метрика важна для мониторинга качества статистических моделей в продакшене?	P-value / Confidence Interval Coverage / Effect Size / Power	МО-2	МО-2.3. 3-1
27.	Какой процесс позволяет автоматизировать переобучение статистических моделей при изменении данных?	Auto-retraining Pipeline / Trigger-based Retraining / Drift Detection	МО-2	МО-2.4. У-1
28.	Какой навык необходим для распределения задач в команде статистического анализа?	Лидерство / Делегирование / Планирование / Sprint Planning	РГ-1	РГ-1.1. У-1
29.	Как называется документ, описывающий дизайн статистического эксперимента для заказчика?	Техническая спецификация / Experimental Design Document / Statistical Analysis Plan	ППК-Р5	ППК-Р5.2. 3-2
30.	Какой принцип чистого кода важен при разработке статистических пайплайнов?	DRY / SOLID / KISS / Modular Design / Reproducibility	ППК-Р6	ППК-Р6.3. 3-1
31.	Что такое технический долг в контексте статистических проектов?	Накопленные проблемы кода / Устаревшие модели / Отсутствие документации / Quick Fixes	РГ-4	РГ-4.3. У-1
32.	Какой инструмент ИИ помогает генерировать код для статистического анализа и экспериментов?	GitHub Copilot / ChatGPT / Cursor / Codeium	ППК-Р7	ППК-Р7.1. У-1
33.	Что необходимо учитывать при использовании ИИ-кода с точки зрения этики в статистических проектах?	Лицензирование / Безопасность / Конфиденциальность данных / Reproducibility	ППК-Р7	ППК-Р7.4. 3-1
34.	Какой метод используется для проверки гипотез в эксперименте со статистической моделью?	T-test / Z-test / ANOVA / Chi-square / Statistical Significance	ППК-ДА5	ППК-ДА5.2. У-2

35.	Как называется процесс сбора информации о требованиях к статистическому анализу от заказчика?	Интервью / Анализ требований / Business Analysis / Stakeholder Interview	ППК-ДА4	ППК-ДА4.1. У-1
36.	Какой навык важен для презентации результатов статистического анализа заказчику?	Коммуникация / Визуализация / Презентация / Storytelling	УНК-01 (УК-4)	УНК-01.01.01. У-1
37.	Какое действие проявляет поддержку коллег в команде статистического анализа?	Наставничество / Code Review / Обратная связь / Knowledge Sharing	УНК-02 (УК-3)	УНК-02.01.01. У-5
38.	Какой инструмент используется для управления задачами в статистическом проекте?	Jira / Trello / Asana / GitHub Projects / Linear	РГ-4	РГ-4.1. 3-1
39.	Что такое стратегия развёртывания промышленной статистической системы?	Canary / Blue-Green / Rolling / Shadow Mode / A/B Testing Pipeline	ППК-Р6	ППК-Р6.4. 3-1
40.	Какой метод используется для оптимизации вычислений в статистических моделях?	Vectorization / Parallel Computing / Caching / Model Distillation	ППК-Р6	ППК-Р6.3. У-1
41.	Как называется план процесса разработки статистического продукта?	Roadmap / План проекта / Gantt Chart / Timeline / Sprint Plan	РГ-2	РГ-2.2. У-1
42.	Какой принцип важен при проектировании архитектуры статистической системы для продакшена?	Modularity / Scalability / Reproducibility / Low Latency / Observability	ППК-Р5	ППК-Р5.3. 3-1
43.	Как называется процесс адаптации к изменениям в требованиях к статистической модели?	Change Management / Adaptation / Model Retraining / Feedback Loop	РГ-2	РГ-2.3. 3-1
44.	Какой инструмент используется для поиска уязвимостей в коде статистических систем?	SAST / Security Scan / Static Analysis / Dependency Check	ППК-Р7	ППК-Р7.2. 3-1
45.	Как называется метод интерпретации рекомендаций ИИ по оптимизации статистического кода?	Code Review / Validation / Manual Verification / Security Audit	ППК-Р7	ППК-Р7.3. У-2
46.	Что такое долгосрочное хранение данных экспериментов и артефактов в статистических проектах?	Data Lake / Artifact Store / Experiment Registry / Statistical Database	ППК-Р4	ППК-Р4.2. У-2
47.	Какой навык необходим для работы в команде над проектом статистического анализа?	Коллаборация / Коммуникация / Teamwork / Cross-functional Collaboration	УНК-02 (УК-3)	УНК-02.01.01. У-5

48.	Как называется процесс оценки результатов проверки работоспособности статистической системы?	Testing / Validation / QA / Integration Testing / Performance Testing	РГ-1	РГ-1.2. У-2
49.	Какой метод используется для выявления бизнес-проблем в проекте статистического анализа?	Анализ предметной области / Business Analysis / Problem Framing / ROI Analysis	ППК-ДА4	ППК-ДА4.2. У-1
50.	Как называется документация, оформленная в соответствии с профессиональными стандартами для статистических проектов?	Техническая документация / API Docs / Statistical Analysis Plan / Runbooks	УНК-01 (УК-4)	УНК-01.02.01. У-3