
УТВЕРЖДЕНА

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

**Рабочая программа дисциплины (модуля)
«Основы статистики. Продвинутый уровень»**

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Управление продуктами

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2026

**Москва
2026**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	8
4. Содержание дисциплины (модуля).....	8
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) «Основы статистики. Продвинутый уровень» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по направлению 02.04.01 Математика и компьютерные науки, профиль Управление продуктами, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Основы статистики. Продвинутый уровень» играет ключевую роль в формировании у студентов навыков анализа, интерпретации и обобщения данных, необходимых для принятия обоснованных решений на основе количественной информации. Освоение продвинутых статистических методов позволяет специалистам эффективно работать с данными, строить надёжные выводы и обосновывать стратегические решения в условиях сложной и динамичной реальности.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Управление продуктами и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений как дисциплина по выбору.

Дисциплина (модуль) доступна к изучению на 2 курсе в 3 или 4 семестре на выбор.

Цель изучения дисциплины (модуля): сформировать у студентов теоретические знания и практические навыки в области математической статистики, необходимые для корректного анализа данных, построения оценок, проверки гипотез и проведения A/B тестов в условиях реальных задач.

Задачи изучения дисциплины (модуля):

- освоить ключевые понятия теории вероятностей и математической статистики: распределения, оценки, доверительные интервалы, гипотезы;
- научиться применять методы точечного и интервального оценивания, включая метод моментов и максимального правдоподобия;
- освоить техники проверки статистических гипотез и интерпретации их результатов;
- разработать навыки проектирования и анализа A/B тестов с учётом мощности, размера выборки и корректности выводов;
- приобрести опыт работы с корреляцией, ковариацией и линейной регрессией, а также применения метода бутстрапа и базовых понятий теории информации.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- определения, параметры и характеристики основных распределений;
- точечные оценки и их свойства;
- метод моментов и метод максимального правдоподобия;
- основные предельные теоремы математической статистики;
- методика построения и виды доверительных интервалов;
- методика бутстрапа для различных задач;
- понятия и алгоритмы, связанные с тестированием гипотез;
- понятие ковариации и различные виды корреляции, а также значимость корреляций;
- основные понятия линейной регрессии;
- этапы и различные техники A/B тестирования;
- базовые понятия теории информации.

уметь:

- визуализировать плотность и рассчитывать характеристики различных распределений;
- рассчитывать и интерпретировать точечные оценки на выборке;
- строить доверительные интервалы;
- тестировать гипотезы с помощью параметрических и непараметрических критериев;
- оцифровывать зависимость между случайными величинами и определять значимость зависимости;
- проектировать, запускать и анализировать A/B тесты;
- применять расчёт энтропии в статистических задачах.

владеть:

- навыком строить распределения и рассчитывать их характеристики в python;
- способами рассчитывать и интерпретировать точечные и интервальные оценки;
- опытом делать выводы с помощью тестирования гипотез;
- практикой использовать бутстрап для построения оценок и тестирования гипотез;
- опытом делать выводы о корреляции и каузации случайных величин;
- практикой запускать, мониторить и оцифровывать результаты проведения A/B тестов.

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 критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1.	Знает методы декомпозиции сложных продуктовых задач, принципы работы с неопределенностью в ИТ-среде и основы стратегического планирования в условиях быстро меняющегося рынка
		УК-1.2.	Умеет применять фреймворки системного анализа для выявления ключевых причин проблем в продукте, формулировать гипотезы и выбирать оптимальные стратегии на основе данных
		УК-1.3.	Имеет опыт проведения анализа реальных кейсов, способен обосновывать стратегические решения с учетом технических ограничений, бизнес-рисков и этических норм работы с данными
ОПК-1.	Способен находить, формулировать и решать актуальные и значимые проблемы прикладной и компьютерной математики	ОПК-1.1.	Знает основные методы и подходы к решению задач прикладной и компьютерной математики, включая алгоритмы, математическое моделирование и теорию оптимизации, а также современные инструменты и технологии, применяемые для прогнозирования и решения задач управления продуктом
		ОПК-1.2.	Умеет анализировать и формулировать математические задачи на основе проблем управления продуктом, выбирать и применять соответствующие методы и алгоритмы для их решения, а также интерпретировать и представлять результаты в понятной и доступной форме для принятия обоснованных решений
		ОПК-1.3.	Имеет практический опыт

			работы над проектами или исследованиями в области прикладной и компьютерной математики, включая участие в конкурсах, хакатонах или научных публикациях, где были решены актуальные и значимые задачи управления продуктом, подтвержденные публикацией результатов или внедрением разработанных алгоритмов
ПК-5.	Способен оценивать техническую реализуемость продуктовых решений в ИТ-среде и интегрировать модели машинного обучения	ПК-5.1.	Знает возможности и ограничения применения моделей машинного обучения, основы архитектуры программных систем и принципы внедрения данных решений в ИТ-среде
		ПК-5.2.	Умеет оценивать трудозатраты на реализацию функций, формулировать задачи для команд аналитики данных и машинного обучения, а также составлять технические требования
		ПК-5.3.	Имеет практический опыт участия в проектировании решений с использованием моделей машинного обучения, обеспечивая баланс между бизнес-ценностью и технической сложностью внедрения
ПК-6.	Способен разрабатывать программное обеспечение для решения прикладных задач в сфере управления продуктами в ИТ-среде	ПК-6.1.	Знает основные языки программирования и инструменты анализа данных, основы статистики и принципы проектирования программных систем, применяемых для решения задач управления продуктами в ИТ-среде
		ПК-6.2.	Умеет разрабатывать программные модули и скрипты для анализа метрик и валидации гипотез, используя современные технологии, а также проводить тестирование и отладку созданных решений
		ПК-6.3.	Имеет практический опыт разработки программного обеспечения для решения задач управления продуктами в ИТ-среде, включая автоматизацию рабочих процессов и презентацию

			результатов заинтересованным сторонам
--	--	--	--

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостояте льная работа	
		Лекции	Семинары (практичес кие занятия)			
1	Введение. Аксиоматика теории вероятностей	3	6		10	Домашние задания
2	Распределения и оценки	8	15		25	Домашние задания
3	Гипотезы и А/В тесты	8	15	2	25	Домашние задания, Контрольная работа
4	Зависимость случайных величин	8	15		25	Домашние задания
5	Теория информации	3	7		11	Домашние задания
	Экзамен			4		
Итого:		30	58	6	96	
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	Введение. Аксиоматика теории вероятностей	Аксиоматика теории вероятностей
2	Распределения и оценки	Дискретные и непрерывные распределения Свойства точечных оценок - 1 Свойства точечных оценок - 2 Методы построения оценок Доверительные интервалы и численные методы
3	Гипотезы и А/В тесты	Введение в гипотезы и продуктивные эксперименты р-значение, минимальный детектируемый эффект и мощность Критерий Стьюдента и тесты на дисперсию Множественное тестирование и Ratio-метрики Непараметрические критерии
4	Зависимость случайных величин	Зависимость случайных величин Корреляция Линейная регрессия - 1 Линейная регрессия - 2
5	Теория информации	Основы теории информации

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. Кремер, Н. Ш. Теория вероятностей и математическая статистика : учебник и практикум для вузов / Н. Ш. Кремер. — 5-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 538 с. — (Высшее образование). — ISBN 978-5-534-10004-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/565694>.

2. Гмурман, В. Е. Теория вероятностей и математическая статистика : учебник для вузов / В. Е. Гмурман. — 12-е изд. — Москва : Издательство Юрайт, 2025. — 479 с. — (Высшее образование). — ISBN 978-5-534-00211-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/559584>.

Дополнительная литература:

1. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике : учебник для вузов / В. Е. Гмурман. — 11-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 395 с. — (Высшее образование). — ISBN 978-5-534-21643-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/581860>.

2. Статистика : учебник для вузов / под редакцией В. С. Мхитаряна. — 3-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 503 с. — (Высшее образование). — ISBN 978-5-534-18687-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/568978>.

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) «Основы статистики. Продвинутый уровень» 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 (модуля) и продемонстрировать навыки их практического применения.

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

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) «Основы статистики. Продвинутый уровень»

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 силлабусе 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 (модуль). 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,
6	Хорошо	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		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.
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 задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Основы статистики. Продвинутый уровень» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	60%	14	Набор задач по темам недели
Контрольная работа	20%	1	Письменная работа с набором задач, которые нужно решить за ограниченное время
Экзамен	20%	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) «Основы статистики. Продвинутый уровень»: « $0,6 \times$ среднее за домашние задания + $0,2 \times$ контрольная работа + $0,2 \times$ экзамен».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1. Аксиомы теории вероятностей и свойства случайных величин

Задание:

- Сформулируйте три аксиомы Колмогорова теории вероятностей. Приведите по одному примеру, иллюстрирующему каждую аксиому (например, бросок кубика, выбор карты и т.п.).
- Дано вероятностное пространство:
 - Множество элементарных исходов: $\Omega = \{1, 2, 3, 4, 5, 6\}$ (бросок игрального кубика)
 - События: $A = \{1, 3, 5\}$ — выпало нечётное число, $B = \{2, 4, 6\}$ — чётное, $C =$

- $\{6\}$ — выпала шестёрка
 - Вычислите:
 - $P(A)$, $P(B)$, $P(C)$
 - $P(A \cup B)$, $P(A \cap C)$
 - Проверьте, являются ли события A и B несовместными и полной группой событий
3. Пусть X — дискретная случайная величина с распределением:
- $P(X \leq 3)$
 - $M(X)$ — математическое ожидание
 - $D(X)$ — дисперсию
 - Постройте функцию распределения $F(x)$ и нарисуйте её график (схематично)
4. Приведите пример непрерывной случайной величины. Запишите, какие условия должна удовлетворять её плотность распределения $f(x)$, согласно аксиоматике.
5. Докажите, что $P(\emptyset) = 0$, используя аксиомы Колмогорова.

Требования к оформлению:

- Решения должны быть обоснованы;
- Все формулы — с пояснениями;
- График функции распределения — от руки или в электронном виде;
- Формат: PDF или рукописный вариант (отсканированный).

Домашнее задание 2. Работа с распределениями и точечными оценками

Задание:

1. Для заданной выборки (например, объём продаж за 30 дней) рассчитайте точечные оценки: выборочное среднее, дисперсию, медиану, моду.
2. Постройте гистограмму и эмпирическую функцию плотности (с помощью **seaborn** или **matplotlib**).
3. Определите, какое теоретическое распределение (нормальное, экспоненциальное, Пуассона) лучше всего описывает данные. Обоснуйте выбор.
4. Сравните выборочные оценки с теоретическими параметрами распределения.
5. Оцените смещённость и состоятельность выборочного среднего на основе моделирования (сгенерируйте выборки разного объёма).

Требования:

- Код на Python (Jupyter Notebook или .py)
- Пояснения к каждому шагу
- Визуализации
- Выводы по результатам

Домашнее задание 3. Построение доверительных интервалов и применение бутстрапа

Задание:

1. На основе выборки (например, время загрузки страницы) постройте 95% доверительный интервал для среднего с использованием:
 - Аналитического метода (при условии нормальности)
 - Метода бутстрапа (с 1000 повторений)
2. Сравните результаты.
3. Проверьте, выполняется ли центральная предельная теорема при малом и большом объёме выборки (смоделируйте).
4. Оцените, как меняется ширина доверительного интервала при увеличении объёма выборки.

Требования:

- Реализация бутстрапа на Python

- Визуализация распределения средних
- Интерпретация результатов
- Отчёт с выводами

Домашнее задание 4. А/В тестирование и проверка гипотез

Задание:

1. Даны две выборки: время пребывания пользователей на старой и новой версии сайта.
2. Сформулируйте нулевую и альтернативную гипотезы.
3. Проверьте гипотезу о равенстве средних с помощью t-критерия Стьюдента.
4. Проверьте нормальность распределения (с помощью критерия Шапиро-Уилка) и равенство дисперсий (критерий Левене).
5. Если условия t-критерия не выполняются — примените непараметрический критерий (Манна-Уитни).
6. Рассчитайте p-значение, минимальный детектируемый эффект и мощность теста.
7. Сделайте вывод: есть ли статистически значимое различие?

Требования:

- Код с проверкой всех условий
- Интерпретация p-значения и мощности
- Отчёт с пошаговым обоснованием выбора критерия
- Рекомендация по внедрению новой версии

Примерные задания для контрольной работы

Задание 1

Сформулируйте аксиомы теории вероятностей по Колмогорову. Проверьте, может ли заданная функция P быть вероятностью на $\Omega = \{a, b, c\}$:
 $P(a) = 0.4$, $P(b) = 0.5$, $P(c) = 0.2$. Обоснуйте ответ.

Задание 2

Даны события A и B , такие что $P(A) = 0.6$, $P(B) = 0.5$, $P(A \cap B) = 0.3$.

Найдите:

- $P(A \cup B)$
- $P(\bar{A})$ — вероятность противоположного события
- Являются ли A и B независимыми? Обоснуйте.

Задание 3

Дана выборка: [12][15][14][10][13][16][11][14].

Рассчитайте выборочное среднее и исправленную выборочную дисперсию.

Постройте 90% доверительный интервал для среднего, если известно, что данные распределены нормально.

Задание 4

Сформулируйте нулевую и альтернативную гипотезы для проверки, увеличивает ли новый баннер конверсию.

Объясните, что такое p-значение и как его интерпретировать.

Что произойдёт с вероятностью ошибки I рода при множественном тестировании? Как её можно контролировать?

Задание 5

Даны два признака: возраст и время просмотра видео.

Рассчитайте коэффициент Пирсона и Спирмена.

Объясните разницу между ними.

Можно ли утверждать, что между признаками есть причинно-следственная связь, если корреляция высокая?

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Где целесообразно проверять математическую строгость для выработки стратегии действий?	Учебники / Научные статьи	УК-1
2.	Каким способом обеспечивается воспроизводимость при критическом анализе экспериментов?	Фиксация random seed	УК-1
3.	Какие этические аспекты важны при системном анализе пользователей?	Информированное согласие	УК-1
4.	Каким образом планируется развитие компетенций для работы с неопределенностью?	Индивидуальный план обучения	УК-1
5.	Запишите формулу аксиомы аддитивности для решения математической задачи.	$P(A \cup B) = P(A) + P(B)$	ОПК-1
6.	Как выглядит формула математического ожидания в прикладной математике?	$E[X] = \sum x_i * p_i$	ОПК-1
7.	Что выражает формула дисперсии в математическом моделировании?	$Var(X) = E[(X - \mu)^2]$	ОПК-1
8.	Как записывается плотность нормального распределения в математических моделях?	$f(x) = (1/\sigma\sqrt{2\pi}) * e^{-(x-\mu)^2/2\sigma^2}$	ОПК-1
9.	Какой критерий применяется для оценки технической реализуемости сравнения выборок?	t-критерий Стьюдента	ПК-5
10.	Как вычисляется p-значение для валидации гипотез в ML-моделях?	$P(T \geq t)$	Н0)
11.	Что означает уровень значимости $\alpha = 0.05$ для оценки технических рисков?	5% риск ошибки I рода	ПК-5
12.	Как рассчитывается MDE для формирования технических требований?	$MDE \propto \sqrt{(Var/n)}$	ПК-5
13.	Запишите формулу корреляции Пирсона для разработки ПО анализа данных.	$r = cov(X,Y) / (\sigma_x * \sigma_y)$	ПК-6
14.	Как выглядит уравнение линейной регрессии для программной реализации?	$y = \beta_0 + \beta_1 * x + \varepsilon$	ПК-6

15.	Что минимизирует метод наименьших квадратов в алгоритме ПО?	$\Sigma(y_i - \hat{y}_i)^2$	ПК-6
16.	Как записывается формула мощности теста для интеграции моделей ML?	Power = 1 - β	ПК-5
17.	Какая поправка применяется при множественном тестировании в скриптах?	Бонферрони (α/k)	ПК-6
18.	Как вычисляется энтропия в математической теории информации?	$H(X) = -\Sigma p_i * \log(p_i)$	ОПК-1
19.	Что измеряет KL-дивергенция в математических распределениях?	Расхождение P от Q	ОПК-1
20.	Какой критерий используется в программных пайплайнах при нарушении нормальности?	Манна-Уитни / Вилкоксона	ПК-6