

**КАЗАНСКИЙ КООПЕРАТИВНЫЙ ИНСТИТУТ (ФИЛИАЛ)
АВТОНОМНОЙ НЕКОММЕРЧЕСКОЙ ОБРАЗОВАТЕЛЬНОЙ
ОРГАНИЗАЦИИ ВЫСШЕГО ОБРАЗОВАНИЯ
ЦЕНТРОСОЮЗА РОССИЙСКОЙ ФЕДЕРАЦИИ
«РОССИЙСКИЙ УНИВЕРСИТЕТ КООПЕРАЦИИ»**

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ

Специальность: 10.05.01 Компьютерная безопасность

Специализация: «Разработка защищенного программного обеспечения»

Формы обучения: очная

Объем дисциплины:

в зачетных единицах: 5 з.е.

в академических часах: 180 ак.ч.

Форма промежуточной аттестации: экзамен

Рабочая программа по дисциплине «Многомерный статистический анализ» по специальности 10.05.01 Компьютерная безопасность, специализации «Разработка защищенного программного обеспечения», составлена в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования – специалитет по специальности 10.05.01 Компьютерная безопасность, утвержденного приказом Министерства науки и высшего образования Российской Федерации от «26» ноября 2020 г. № 1459, профессионального стандарта 06.032 «Специалист по безопасности компьютерных систем и сетей», утвержденного приказом Министерства труда и социальной защиты Российской Федерации от 14.09.2022 г. № 533н, профессионального стандарта 06.033 «Специалист по защите информации в автоматизированных системах», утвержденного приказом Министерства труда и социальной защиты Российской Федерации от 14.09.2022 г. №525н.

Рабочая программа:

обсуждена и рекомендована к утверждению решением кафедры естественных дисциплин, сервиса и туризма от 25 марта 2025 г., протокол № 6

одобрена Научно-методическим советом Казанского кооперативного института (филиала) от 2 апреля 2025 г., протокол № 3

СОДЕРЖАНИЕ

1. Цель и задачи освоения дисциплины	4
2. Место дисциплины в структуре образовательной программы	4
3. Перечень планируемых результатов обучения по дисциплине	5
4. Объем дисциплины и виды учебной работы.....	5
5. Содержание дисциплины, структурированное по темам (разделам) с указанием количества академических часов и видов учебных занятий.....	6
5.1. Содержание дисциплины	6
5.2. Разделы, темы дисциплины и виды занятий	9
6. Практические занятия	11
7. Лабораторные работы	11
8. Примерная тематика курсовых работ (проектов)	16
9. Самостоятельная работа студента	16
10. Оценивание результатов обучения и уровня сформированности компетенций	18
11. Учебно-методическое обеспечение дисциплины	18
12. Перечень программного обеспечения, профессиональных баз данных и информационных справочных систем, необходимых для освоения дисциплины.....	19
13. Материально–техническое обеспечение дисциплины	20
Обновление рабочей программы дисциплины (модуля)	22

1. Цель и задачи освоения дисциплины

Цель дисциплины - является изучение основных методов многомерного статистического анализа данных с точки зрения их практического применения; привитие навыков работы с соответствующими разделами ППП STATISTICA и правильной интерпретации результатов исследований.

Задачи:

- помочь студентам понять и, освоить методологию многомерного статистического анализа данных;
- привить теоретические и практические знания в области прикладного многомерного анализа данных;
- познакомить студентов и обучить максимально широкому инструментарию многомерного анализа данных в среде ППП STATISTICA;
- выработать в процессе обучения у студентов навыки грамотного использования аппарата вероятностно-статистического моделирования посредством применения передовых информационных технологий.

2. Место дисциплины в структуре образовательной программы

Дисциплина «Многомерный статистический анализ» относится к Дисциплинам специализации основной образовательной программы – программы специалитета по специальности 10.05.01 Компьютерная безопасность специализация «Разработка защищенного программного обеспечения».

Код компетенции	Предшествующие дисциплины (модули), практики	Изучаемые в текущем семестре дисциплины (модули), практики	Последующие дисциплины (модули), практики
ОПК-3	Дифференциальные уравнения Численные методы Практикум по численным методам Уравнения математической физики Теория игр Комбинаторный анализ Теория графов и ее приложения	Многомерный статистический анализ	Математический анализ Алгебра и аналитическая геометрия

3. Перечень планируемых результатов обучения по дисциплине

Изучение дисциплины направлено на формирование у обучающихся компетенций.

Формируемые компетенции (код и наименование компетенции)	Индикаторы достижения компетенций	Планируемые результаты обучения
ОПК-3 Способен на основании совокупности математических методов разрабатывать, обосновывать и реализовывать процедуры решения задач профессиональной деятельности	ОПК-3.4 Демонстрирует умение обосновывать и реализовывать процедуры решения стандартных задач профессиональной деятельности с применением методов анализа данных	Знать: основных математических методов и их модификаций для обоснования процедур решения задач профессиональной деятельности Уметь: реализовывать процедуры решения многомерных статистических задач с применением основных математических методов Владеть: навыками теоретического и экспериментального исследования совокупности математических методов для разработки, обоснования и реализации процедур решения задач профессиональной деятельности
ОПК-3.1 Способен проводить анализ программного кода с целью поиска потенциальных уязвимостей и недокументированных возможностей	ОПК-3.1.1 Знает методы и средства защиты информации в компьютерных сетях, операционных системах и системах управления базами данных	Знать: основные понятия и современные принципы многомерного анализа. Уметь: строить прогностические модели для прогнозирования новых угроз, строить закономерности на данных о ранее совершенных атаках Владеть: математическим аппаратом для построения моделей
	ОПК-3.1.2 Демонстрирует умение применять методы анализа защищенности информации от атак на компьютерные системы и сети	Знать: основные понятия и современные принципы многомерного анализа. Уметь: логически верно, аргументировано и ясно строить устную и письменную речь Владеть: культурой мышления, навыками обобщенного анализа и восприятия информации.

4. Объем дисциплины и виды учебной работы

Объем дисциплины и виды учебной работы в академических часах с выделением объема контактной работы обучающихся с преподавателем и самостоятельной работы обучающихся

очная форма обучения

Вид учебной деятельности	ак. часов	
	Всего	По семестрам
		7
1. Контактная работа обучающихся с преподавателем:	69	69

Аудиторные занятия, часов всего, в том числе:	68	68
• занятия лекционного типа	34	34
• занятия семинарского типа:	34	34
практические занятия	-	-
лабораторные занятия	34	34
в том числе занятия в форме практической подготовки	-	-
Консультации	0,5	0,5
Контактные часы на аттестацию в период экзаменационных сессий	0,5	0,5
2. Самостоятельная работа студентов всего, в том числе	75	75
- подготовка курсовой работы	-	-
- проработка учебного (теоретического) материала	25	25
- выполнение домашних заданий по практическим занятиям	25	25
- подготовка к экзамену	25	25
Промежуточная аттестация: экзамен	36	36
ИТОГО:		
общая трудоемкость	часов	144
	зач. ед.	4

5. Содержание дисциплины, структурированное по темам (разделам) с указанием количества академических часов и видов учебных занятий

5.1. Содержание дисциплины

Раздел 1. Основные приемы работы с ППП STATISTICA

Тема 1. Элементарные понятия статистики.

Статистика — область науки, занимающаяся сбором, обработкой, анализом и интерпретацией данных для выявления закономерностей и закономерностей, присущих большим объемам информации. Характеристики или свойства, присущие объектам выборки. Количественное выражение признака, принимающее конкретные значения. Методы сбора данных: наблюдение, анкетирование, эксперимент.

Тема 2. Измерительные шкалы. Работа с данными. Основные операции над случаями и переменными

Измерительные шкалы играют ключевую роль в процессе научного познания, позволяя объективно фиксировать и анализировать информацию. Важно уметь правильно подбирать и применять подходящие шкалы, а также владеть операциями над случаями и переменными, необходимыми для дальнейшей обработки данных.

Раздел 2. Основные статистики и методы сравнения средних величин

Тема 3. Основные статистики, корреляционный анализ.

Основная задача статистики заключается в сборе, обработке и анализе

данных с целью обнаружения закономерностей и взаимосвязей. Один из важнейших аспектов статистического анализа — изучение связей между различными переменными, чему посвящен корреляционный анализ.

Тема 4. Сравнение средних величин критерием Стьюдента.

Критерий Стьюдента (t-критерий) широко применяется в статистике для проверки гипотез о равенстве средних значений двух выборок. Этот метод особенно полезен при сравнении небольших выборок с примерно нормально распределёнными данными.

Тема 5. Непараметрическая статистика. Сравнение величин методами непараметрической статистики.

Параметрические методы предполагают знание точного закона распределения данных, тогда как непараметрические методы не требуют таких жестких предположений и подходят даже для ненормально распределённых данных. Одним из преимуществ непараметрических методов является устойчивость к выбросам и нечувствительность к отклонениям от нормальной формы распределения.

Тема 6. Группировка и однофакторная ANOVA.

Однофакторный дисперсионный анализ (ANOVA) представляет собой статистический метод, предназначенный для проверки гипотез о равенстве средних значений нескольких групп. Он применяется, когда исследователь хочет сравнить три или более групп, относящихся к одной независимой переменной (фактора), с целью выяснить, влияют ли различия между этими группами на зависимую переменную.

Тема 7. Дисперсионный анализ.

Дисперсионный анализ (Analysis of Variance, ANOVA) — это статистический метод, используемый для проверки гипотез о равенстве средних значений нескольких групп. Основной задачей ANOVA является выявление существенности различий между средним значением зависимых переменных в разных группах, определяемых одним или несколькими факторами.

Раздел 3. Методы многомерного анализа

Тема 8. Таблицы частот, сопряженности, флагов и заголовков.

Одним из первых шагов в анализе данных является просмотр и первичная обработка информации. Именно для этого созданы специальные средства: таблицы частот, таблицы сопряженности, системы маркировки данных с помощью флагов и удобочитаемые заголовки.

Эти инструменты важны как для исследователей, так и для пользователей данных, так как они облегчают восприятие и интерпретацию информации.

Тема 9. Канонический анализ.

Канонический анализ — это метод многомерной статистики, позволяющий находить скрытые взаимосвязи между двумя множествами переменных, образуя новые комбинации переменных, называемые каноническими переменными. Суть метода заключается в поиске таких линейных комбинаций переменных, которые максимально сильно связаны друг с другом.

Цель канонического анализа — снижение размерности данных и установление внутренних взаимосвязей между двумя наборами переменных.

Тема 10. Линейное и нелинейное моделирование взаимосвязей.

Одной из важнейших задач статистического анализа является моделирование взаимосвязей между переменными. Существуют два основных класса моделей: линейные и нелинейные. В данном материале мы рассмотрим, как строятся и интерпретируются такие модели, а также сравним их возможности и ограничения.

Тема 11. Дискриминантный анализ.

Дискриминантный анализ — это метод многомерной статистики, направленный на классификацию объектов по известным признакам и изучение различий между классами. Этот метод позволяет решать задачи отнесения нового объекта к одному из предварительно установленных классов, а также выявления наиболее значимых признаков, влияющих на принадлежность к классу.

Тема 12. Кластерный анализ.

Кластерный анализ — это раздел многомерной статистики, посвященный поиску естественных групп (кластеров) в наборе данных. Цель кластерного анализа — разделить множество объектов на подмножества (кластеры), внутри которых объекты схожи между собой больше, чем с объектами из других кластеров.

Тема 13. Деревья классификации.

Деревья классификации — это интуитивно понятный и универсальный метод анализа данных, позволяющий эффективно решать задачи классификации и прогнозирования. Их главная особенность — наглядность и доступность интерпретации, что делает их популярным выбором для многих областей: медицины, маркетинга, экономики и других.

Тема 14. Факторный анализ.

Факторный анализ — это метод многомерной статистики, используемый для выявления скрытых общих факторов, лежащих в основе набора переменных. Главная идея факторного анализа заключается в снижении размерности данных путём объединения тесно связанных

переменных в небольшие группы факторов, сохраняя при этом максимальное количество информации.

Тема 15. Многомерное шкалирование.

Многомерное шкалирование (multidimensional scaling, MDS) — это метод многомерной статистики, применяемый для реконструкции геометрии высокого пространства признаков в низкоространственном представлении. Основная цель — сохранить соотношения расстояний между объектами, сохранив при этом главное свойство исходных данных.

Тема 16. Обзор пройденного материала.

Данный документ предназначен для повторения и систематизации основных понятий и методов, рассмотренных в предыдущем учебном периоде. Материал охватывает фундаментальные темы статистики и анализа данных, что позволит освежить знания и подготовиться к следующей стадии обучения.

5.2. Разделы, темы дисциплины и виды занятий

очная форма обучения

№ п/п	Наименование раздела, темы дисциплины	Виды занятий, включая самостоятельную работу студентов (в часах)			Индикаторы достижения компетенций
		занятия лекционного типа	занятия семинарского типа / из них в форме практической подготовки (лабораторные работы)	самос- тоятельная работа	
Раздел 1. Основные приемы работы с ППП STATISTICA					
1	Тема 1. Элементарные понятия статистики. Из- мерительные шкалы	2	2/2	5	ОПК-3, ОПК- 3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
2	Тема 2. Работа с данными. Основные операции над случаями и переменными	2	2/2	5	ОПК-3, ОПК- 3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
Раздел 2. Основные статистики и методы сравнения сред них величин					
3	Тема 3. Основные статистики, корреляционный анализ	2	2/2	5	ОПК-3, ОПК- 3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
4	Тема 4. Сравнение средних величин критерием Стьюдента	2	2/2	5	ОПК-3, ОПК- 3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2

№ п/п	Наименование раздела, темы дисциплины	Виды занятий, включая самостоятельную работу студентов (в часах)			Индикаторы достижения компетенций
		занятия лекционного типа	занятия семинарского типа / из них в форме практической подготовки (лабораторные работы)	самос- тоятельная работа	
5	Тема 5. Непараметрическая статистика. Сравнение средних величин методами непараметрической статистики	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
6	Тема 6. Группировка и однофакторная ANOVA	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
7	Тема 7. Дисперсионный анализ	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
Раздел 3. Методы многомерного анализа					
8	Тема 8. Таблицы частот, сопряженности, флагов и заголовков	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
9	Тема 9. Канонический анализ	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
10	Тема 10. Линейное и нелинейное моделирование взаимосвязей	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
11	Тема 11. Дискриминантный анализ	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
12	Тема 12. Кластерный анализ	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
13	Тема 13. Деревья классификации	2	2/2	5	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
14	Тема 14. Факторный анализ	2	2/2	4	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
15	Тема 15. Многомерное шкалирование	2	2/2	3	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2
16	Тема 16. Обзор пройденного материала.	4	4/4	3	ОПК-3, ОПК-3.4, ОПК-3.1, ОПК-3.1.1, ОПК-3.1.2

№ п/п	Наименование раздела, темы дисциплины	Виды занятий, включая самостоятельную работу студентов (в часах)			Индикаторы достижения компетенций
		занятия лекционного типа	занятия семинарского типа / из них в форме практической подготовки (лабораторные работы)	самос- тоятельная работа	
	Всего	34	34/34	75	

6. Практические занятия

Практические занятия не предусмотрены.

7. Лабораторные работы

№ п/п	Наименование раздела, темы дисциплины	Содержание лабораторных работ	Объем (час.)
			Очная форма обучения
Раздел 1. Основные приемы работы с ППП STATISTICA			
1.	Тема 1.Элементарные понятия статистики. Измерительные шкалы.	Освоение базовых понятий статистики и классификация видов измерительных шкал. Формирование умения анализировать эмпирические данные в зависимости от типа шкалы измерения. Основные задачи лабораторной работы: ознакомиться с базовыми элементами статистики (варианта, признак, генеральная совокупность, выборочная совокупность), научиться различать типы измерительных шкал (номинальная, порядковая, интервальная, относительная), применять полученные знания на практике путем анализа реальных примеров. Понятие статистики и её роль в научных исследованиях. Базовые элементы статистики: вариант, признак, совокупностью и выборка, типы измерительных шкал: номинальная шкала (примеры: пол, национальность, категория товара); порядковая шкала (примеры: степень удовлетворенности качеством услуги, уровень квалификации сотрудника); интервальная шкала (температура по Цельсию, возраст); относительная шкала (вес, расстояние, денежные показатели).	2
2.	Тема 2. Работа с данными. Основные операции над случаями и переменными	Закрепление навыков работы с данными и освоение основных операций над случаями и переменными в среде статистического анализа. Развитие способности эффективно обрабатывать и интерпретировать данные с использованием специализированных инструментов. Основные задачи лабораторной работы: освоить базовые приемы подготовки данных перед проведением статистического анализа, овладеть методами фильтрации и сортировки данных, уметь преобразовывать данные (создание новых переменных, группировка значений), понимать важность контроля качества исходных данных.	2

№ п/п	Наименование раздела, темы дисциплины	Содержание лабораторных работ	Объем (час.)
			Очная форма обучения
		Основные принципы организации и хранения данных в электронных таблицах. Понятия случая (наблюдения) и переменной (признака). Возможности программного обеспечения для управления данными (фильтрация, сортировка, агрегирование). Операции над переменными: создание производных признаков, расчет простых функций (сумма, среднее значение, процентиля).	
Раздел 2. Методы сравнения средних величин			
3.	Тема 3. Основные статистики, корреляционный анализ	Изучение методов расчета и интерпретации основных статистик и проведение корреляционного анализа для оценки взаимосвязи между признаками исследуемого множества данных. Основные задачи лабораторной работы: рассчитать и проанализировать основные статистические характеристики выборочных данных (среднее, дисперсия, стандартное отклонение, асимметрия, эксцесс), осуществить проверку гипотез о параметрах распределения выборки, освоить технику построения корреляционной матрицы и диаграмм рассеяния, проинтерпретировать коэффициенты корреляции Пирсона и Спирмена, оценить силу связи между признаками.	2
4.	Тема 4. Сравнение средних величин критерием Стьюдента	Изучить методику сравнения двух выборочных средних значений с применением t-критерия Стьюдента и освоить практические навыки принятия решений на основе статистических тестов. Основные задачи лабораторной работы: научиться проверять гипотезы о равенстве средних значений двух независимых выборок, использовать критерий Стьюдента для подтверждения или отрицания предположений о равенстве средних, усвоить порядок интерпретации p-значений и доверительных интервалов, получить опыт анализа реального массива данных и принятие обоснованного решения.	2
5.	Тема 5. Непараметрическая статистика. Сравнение средних величин методами непараметрической статистики	Изучение основ непараметрических методов статистики и применение их для сравнения средних величин при неизвестных или нарушенных условиях нормального распределения данных. Основные задачи лабораторной работы: ознакомиться с принципами непараметрических методов, используемых для сравнительного анализа данных, применить тесты Манна-Уитни и Вилкоксона для проверки гипотез о различиях средних величин двух независимых выборок, узнать правила принятия решений на основе рангового подхода и интерпретации полученных результатов, научиться сравнивать выборочные центральные моменты, используя непараметрические критерии.	2
6.	Тема 6. Группировка и однофакторная ANOVA	Изучение методики группировки данных и использование однофакторного дисперсионного анализа (ANOVA) для проверки гипотез о равенстве средних в трех и более группах. Основные задачи лабораторной работы: освоить процедуру группировки данных по одному факторному	2

№ п/п	Наименование раздела, темы дисциплины	Содержание лабораторных работ	Объем (час.)
			Очная форма обучения
		признаку, применить однофакторный дисперсионный анализ (ANOVA) для оценки влияния фактора на зависимую переменную, интерпретировать результаты ANOVA и принять обоснованное решение о равенстве средних, создать навыки анализа вариации внутри и между группами.	
7.	Тема 7. Дисперсионный анализ	Изучение методов дисперсионного анализа (ANOVA) для проверки гипотез о равенстве средних значений в различных группах. Основные задачи лабораторной работы: освоить методологию проведения дисперсионного анализа, научиться применять однофакторный и многофакторный дисперсионный анализ, освоить расчёт суммы квадратов, среднего квадрата и оценку эффектов факторов, освоить приём проверки статистических гипотез с использованием F-критерия. Понятие дисперсионного анализа и его назначение. Однофакторный и многофакторный дисперсионный анализ. Критерии разделения общей дисперсии на компоненты (внутригрупповая и межгрупповая дисперсия). Правила принятия решений на основе р-значения и F-критерия.	2
Раздел 3. Методы многомерного анализа			
8.	Тема 8. Таблицы частот, сопряженности, флагов и заголовков	Научиться создавать и интерпретировать таблицы частот, сопряженности, а также правильно оформлять столбцы и строки таблиц (флаги и заголовки) для эффективной визуализации данных. Основные задачи лабораторной работы: освоить построение таблиц частот и понимание частоты встречаемости отдельных категорий, научиться составлять таблицы сопряженности (кросс-таблицы) для изучения взаимозависимости признаков, понять принципы выделения важных элементов таблиц с помощью флагов и заголовков, интерпретировать результаты таблиц и формировать полезные выводы. Понятие таблицы частот и ее структура. Назначение таблиц сопряженности (контингенции) и способы их составления. Значение флагов и заголовков в оформлении таблиц. Способы визуального выделения важных данных в таблице.	2
9	Тема 9. Канонический анализ	Изучение принципов канонического анализа и получение навыков его применения для установления связей между двумя множествами переменных. Основные задачи лабораторной работы: освоить концепцию канонических корреляций и их роли в установлении связей между множеством переменных, научиться выбирать оптимальные канонические комбинации и оценивать их значимость, овладеть процедурой канонического анализа с использованием специального программного обеспечения, правильно интерпретировать результаты канонического анализа и делать обоснованные выводы. Основные понятия канонического анализа: каноническая корреляция, канонические веса, канонические нагрузки. Алгоритм расчета канонической	2

№ п/п	Наименование раздела, темы дисциплины	Содержание лабораторных работ	Объем (час.)
			Очная форма обучения
		корреляции. Особенности применения канонического анализа в социальных науках и экономике.	
10	Тема 10. Линейное и нелинейное моделирование взаимосвязей	Изучение методов линейного и нелинейного моделирования взаимосвязей между переменными, обучение выбору адекватных моделей регрессии и проведению анализа остатков. Основные задачи лабораторной работы: изучить особенности построения линейных и нелинейных моделей регрессии, научиться строить модели на основе данных и оценивать качество подобранных уравнений, овладеть методом подбора наилучшей формы модели и диагностики ошибок регрессионного анализа, научиться делать прогнозы на основе полученных моделей. Основные понятия линейной и нелинейной регрессий. Методы наименьших квадратов и максимизации правдоподобия. Проблемы спецификации модели, коллинеарности и гетероскедастичности. Диагностика ошибок регрессионного анализа.	2
11	Тема 11. Дискриминантный анализ	Изучение метода дискриминантного анализа и овладение техникой классификации объектов на основе набора признаков. Основные задачи лабораторной работы: освоить методику проведения дискриминантного анализа, научиться рассчитывать дискриминантные функции и интерпретировать их коэффициенты, научиться классифицировать объекты на основе дискриминантных функций, оценивать точность классификации и интерпретировать результаты. Понятие дискриминантного анализа и его основное предназначение. Формула дискриминантной функции и правило классификации. Процесс расчета дискриминантных функций и оценки их значимости. Методы оценки точности классификации (матрицы ошибок, индекс точности).	2
12	Тема 12. Кластерный анализ	Изучение методов кластерного анализа и развитие навыков классификации объектов на основе их признаков. Основные задачи лабораторной работы: освоить методы иерархического и неиерархического кластерного анализа, научиться определять оптимальное количество кластеров, овладеть приемами интерпретации результатов кластерного анализа, исследовать структуру данных с помощью кластерного анализа. Понятие кластера и задачи кластерного анализа. Основные подходы к формированию кластеров (иерархические и неиерархические методы). Методы оценки оптимального количества кластеров (метод локтя, silhouette score). Этапы проведения кластерного анализа и их последовательность.	2
13	Тема 13. Деревья классификации	Изучение метода деревьев классификации и практика его применения для построения моделей предсказательной аналитики. Основные задачи лабораторной работы: освоить процесс построения дерева классификации, научиться выбирать оптимальный способ	2

№ п/п	Наименование раздела, темы дисциплины	Содержание лабораторных работ	Объем (час.)
			Очная форма обучения
		ветвления узлов и останова роста дерева, научиться интерпретировать дерево классификации и понимать смысл каждого узла, оценить эффективность построенной модели с помощью перекрестной проверки. Основные концепции деревьев классификации: узлы, ветви, листья. Алгоритмы ветвления (энтропия, прирост информации, индекс Джини). Методы останова роста дерева и ограничения глубины. Способы оценки эффективности модели (ошибка классификации, кривая ROC).	
14	Тема 14. Факторный анализ	Изучение метода факторного анализа и умение извлекать скрытые факторы, объясняющие вариацию большого числа переменных. Основные задачи лабораторной работы: освоить процедуру извлечения главных компонент и факторного вращения, научиться интерпретировать факторные нагрузки и выделять наиболее информативные факторы, овладеть приемом сокращения размерности пространства признаков, оценить влияние изменения числа факторов на объясненную долю дисперсии. Основные понятия факторного анализа: латентные переменные, общие и специфичные факторы. Цели и области применения факторного анализа. Методы извлечения факторов (главных компонент, максимального правдоподобия). Стратегии вращений факторов (варимакс, кватримакс, прокрустово вращение).	2
15	Тема 15. Многомерное шкалирование	Изучение метода многомерного шкалирования (Multidimensional Scaling, MDS) и овладение навыком снижения размерности данных путём сохранения попарных расстояний между объектами. Основные задачи лабораторной работы: освоить основные идеи и методы многомерного шкалирования, научиться применять метрическое и неметрическое многомерное шкалирование, овладеть техникой интерпретации результатов MDS, изучить возможности визуализации сложных структур данных с помощью MDS. Концепция многомерного шкалирования и его роль в анализе данных. Отличие метрического и неметрического многомерного шкалирования. Алгоритмы реализации MDS (классический, метрический, неметрический). Инструменты визуализации и интерпретации результатов.	2
16	Тема 16. Обзор пройденного материала.	Основные задачи лабораторной работы: повторение и углубленное изучение ключевых понятий и методов статистики, осмысление пройденного материала и устранение пробелов в знаниях, совершенствование навыков самостоятельной работы с литературой и источниками информации, самостоятельное решение задач и выполнение практических заданий по ключевым вопросам курса.	4
	Всего		34

8. Примерная тематика курсовых работ (проектов)

Курсовые работы (проекты) не предусмотрены.

9. Самостоятельная работа студента

Самостоятельная работа студента при изучении дисциплины «Многомерный статистический анализ» направлена на:

- освоение нормативных правовых актов и учебной литературы, необходимых для освоения дисциплины;
- самостоятельный поиск информации в Интернете и других источниках;
- выполнение домашних заданий по практическим занятиям;
- подготовку к экзамену.

Краткие рекомендации по выполнению самостоятельной работы:

Успешное усвоение дисциплины предполагает большой, упорный, серьезный, систематический труд студентов. Важнейшая его составная часть – выполнение разных видов самостоятельной работы.

1. Составление тематического конспекта на основе изученной основной и дополнительной учебной литературы. В тематическом конспекте за основу берется содержание темы, вопросы для обсуждения.

Этапы работы.

1.1 Конспектирование делается только после того, как прочитан или усвоен материал для конспектирования.

1.2. Необходимо мысленно или письменно составить план конспекта. По этому плану и будет строиться конспект далее.

1.3. Составление самого конспекта. Можно сказать, что конспект – это расширенные тезисы, дополненные рассуждениями и доказательствами, содержащимися в материалах для конспекта, а также собственными мыслями и положениями составителя конспекта.

Писать конспект рекомендуется четко и разборчиво. В конспекте можно выделять места текста в зависимости от их значимости. Для этого применяются различного размера буквы, подчеркивания, замечания на полях.

Примерные задания на лабораторные работы

ОПК-3, ОПК-3,1

Задание 1. Произвести основные операции над переменными и случаями по представленным файлам данных

Задание 2. Построить двухмерные графики по представленным файлам данных.

Задание 3. Вычислить описательные статистики, просчитать параметрические и непараметрические коэффициенты корреляция по представленным файлам данных.

Задание 4. Произвести подгонку законов распределения по представленным файлам данных, генерацию случайных чисел. Решить

задачи.

Задание 5. Произвести сравнение средних посредством t-критерия по представленным файлам данных.

Задание 6. Реализовать метод Группировка и однофакторная ANOVA на представленных файлах данных.

Задание 7. Реализовать метод Дисперсионный анализ на представленных файлах данных.

Задание 8. Реализовать метод Непараметрическая статистика на представленных файлах данных.

Задание 9. Реализовать метод частотный анализ на представленных файлах данных. Задание 10. Реализовать метод Таблицы кросстабуляции и таблицы флагов и заголовков на представленных файлах данных.

Задание 11. Реализовать методы Множественная регрессия, фиксированная нелинейная регрессия на представленных файлах данных.

Задание 12. Реализовать методы Нелинейная регрессия, модели бинарных откликов на представленных файлах данных.

Задание 13. Реализовать метод Дискриминантный анализ на представленных файлах данных.

Задание 14. Реализовать метод Кластерный анализ на представленных файлах данных.

Задание 15. Реализовать метод Деревья классификации на представленных файлах данных.

Задание 16. Реализовать метод Факторный анализ на представленных файлах данных.

Задание 17. Реализовать метод Канонический анализ на представленных файлах данных.

Задание 18. Реализовать метод Многомерное шкалирование на представленных файлах данных.

Зачетно-экзаменационные материалы для промежуточной аттестации (экзамен/зачет)

Примерный перечень вопросов к экзамену

1. Инструменты для работы с данными
2. Структура электронной таблицы
3. Основные операции над переменными и случаями
4. Основные операции с таблицами данных
5. Обмен данными с другими приложениями WINDOWS
6. Двухмерная графика
7. Трехмерная графика
8. Формирование отчета и рабочей книги
9. Описательные статистики
10. Корреляционная матрица
11. t-критерий сравнения средних
12. Группировка и однофакторная ANOVA
13. Таблицы частот

14. Таблицы кросстабуляции
15. Таблицы флагов и заголовков
16. Подбор закона распределения
17. Линейная регрессионная модель
18. Модуль множественная регрессия
19. Линеаризующие преобразования
20. Модели бинарных откликов
21. Описание модуля Нелинейное оценивание
22. Экспоненциальная регрессия
23. Кусочно-линейная регрессия
24. Определенная пользователем регрессия
25. Дискриминантный анализ. Описание модуля Дискриминантный анализ
26. Кластерный анализ. Описание модуля Кластерный анализ.
27. Задача факторного анализа. Описание модуля Факторный анализ
28. Деревья классификации
29. Задача канонического анализа. Описание метода. Модуль Канонический анализ
30. Многомерное шкалирование
31. Анализ соответствий

10. Оценивание результатов обучения и уровня сформированности компетенций

Оценивание результатов обучения по дисциплине и уровня сформированности компетенций (части компетенции) осуществляется в рамках текущего контроля успеваемости и промежуточной аттестации в соответствии с оценочными материалами.

11. Учебно-методическое обеспечение дисциплины

Основная литература:

1. Халафян, А.А. Методы машинного обучения в Data Mining пакета STATISTICA/ А. А. Халафян. - М.: [Горячая линия-Телеком], 2022. - 260 с.
2. Чураков, Е. П. Введение в многомерные статистические методы / Е. П. Чураков. — 3-е изд., стер. — Санкт-Петербург : Лань, 2023. — 148 с. — ISBN 978-5-507-47141-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/330530> (дата обращения: 13.06.2023). — Режим доступа: для авториз. пользователей.
3. Ганичева, А. В. Прикладная статистика / А. В. Ганичева. — 4-е изд., стер. — Санкт-Петербург : Лань, 2023. — 172 с. — ISBN 978-5-507-47980-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/336800> (дата обращения: 13.06.2023). — Режим доступа: для авториз. пользователей.

Для освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья имеются издания в электронном виде в электронно-библиотечной системе «Лань».

Дополнительная литература:

1. Халафян, А.А. Статистический анализ данных. STATISTICA 6 / А. А. Халафян. - М.: [БИНОМ-Пресс], 2010. - 522 с.
2. Халафян, А.А. Статистический анализ данных. STATISTICA 6 / А. А. Халафян. - М.: [Бином-Пресс], 2009. - 522 с.
3. Халафян, А.А. Математическая статистика с элементами теории вероятностей. STATISTICA 6 / А. А. Халафян. - М. : БИНОМ, 2010. - 491 с.
4. Туганбаев, А.А. Теория вероятностей и математическая статистика / А.А. Туганбаев, В.Г. Крупин. – Санкт-Петербург: Лань, 2011. – 320 с. [Электронный ресурс]: – Режим доступа: <https://e.lanbook.com/book/652>
5. Боровиков, В. П. Популярное введение в современный анализ данных в системе STATISTICA. Методология и технология современного анализа данных : учебное пособие/ В. П. Боровиков. — Москва : Горячая линия-Телеком, 2018. — 288 с. — ISBN 978-5- 9912-0326-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/111023> (дата обращения: 13.06.2023). — Режим доступа: для авториз. пользователей.

Электронные библиотечные системы (ЭБС) и электронные образовательные ресурсы

1. Электронно-библиотечная система ZNANIUM <https://www.znanium.com>
2. Универсальная справочно-информационная полнотекстовая база данных периодических изданий «East View» (ИВИС) <https://dlib.eastview.com/>
3. Электронно-библиотечная система «BOOK.ru» www.book.ru
4. Электронно-библиотечная система «IPRbooks» <http://www.iprbookshop.ru/>
5. ЭБС «Национальный цифровой ресурс «Руконт» <https://lib.rucont.ru/search>
6. Образовательная платформа ЮРАЙТ <http://www.urait.ru>

Электронный каталог Университета – Автоматизированная информационная библиотечная система (АИБС «МегаПро»)

12. Перечень программного обеспечения, профессиональных баз данных и информационных справочных систем, необходимых для освоения дисциплины

Лицензионное программное обеспечение, в том числе отечественного производства:

– DsktpEdu ALNG LicSARk OLV F1 1Y Acdmc Ent (MS Windows)

(подписка на ПО).

Свободно распространяемое программное обеспечение, в том числе отечественного производства:

- 7-Zip (свободный файловый архиватор с высокой степенью сжатия данных) (отечественное ПО).

- FastStone Image Viewer (программа для просмотра изображений в Microsoft Windows (лицензия бесплатного программного обеспечения)).

- Foxit Reader (Бесплатное прикладное программное обеспечение для просмотра электронных документов в стандарте PDF (лицензия бесплатного программного обеспечения)).

- Indigo (система программ для создания и проведения компьютерного тестирования знаний).

- Google Chrome, ЯндексБраузер (браузер).

- Adobe Acrobat Reader DC (• ПО для просмотра, печати и комментирования документов в формате PDF)

- Visual Studio Code (интегрированная среда разработки программного обеспечения)

Профессиональные базы данных не используются.

Информационные справочные системы:

- СПС КонсультантПлюс. Компьютерная справочная правовая система, широко используется учеными, студентами и преподавателями (подписка на ПО)

Каждый обучающийся в течение всего периода обучения обеспечивается индивидуальным неограниченным доступом к электронно-библиотечной

13. Материально–техническое обеспечение дисциплины

Образовательный процесс обеспечен учебными аудиториями для проведения учебных занятий, а именно для проведения занятий лекционного типа, практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, помещениями для самостоятельной работы студентов.

Учебная аудитория для проведения занятий лекционного типа

Рабочее место преподавателя (стол, стул), рабочие места обучающихся, доска для написания мелом, кафедра.

Технические средства обучения: персональный компьютер-1, мультимедийное оборудование (проектор, проекционный экран) - 1

Учебная аудитория для проведения практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Рабочее место преподавателя (стол, стул), рабочие места обучающихся, доска для написания мелом, кафедра.

Технические средства обучения: персональный компьютер-1,

мультимедийное оборудование (проектор, проекционный экран) – 1.

Помещение для самостоятельной работы обучающихся, оснащенное компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду института.

Рабочие места обучающихся.

Технические средства обучения: персональный компьютер - 13.

Учебные аудитории соответствуют действующим противопожарным правилам и нормам, укомплектованы учебной мебелью

Учебные аудитории соответствуют действующим противопожарным правилам и нормам, укомплектованы учебной мебелью.

Обновление рабочей программы дисциплины (модуля)

Наименование раздела рабочей программы, в который внесены изменения

(измененное содержание раздела)

обсуждено и рекомендовано к утверждению решением кафедры
наименование кафедры

от ____ ____ 20__ г., протокол № ____

одобрено Научно-методическим советом ____ от ____ ____ 20__ г.,
протокол № ____