



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

федеральное государственное бюджетное образовательное учреждение
высшего образования
**«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ИГУ»)**

Институт математики и информационных технологий
Кафедра алгебраических и информационных систем

«УТВЕРЖДАЮ»
Директор ИМИТ ИГУ
М. В. Фалалеев
«25» мая 2022 г.



Рабочая программа дисциплины (модуля)

Б1.О.08 Математика для анализа данных

Направление подготовки информационные технологии	02.04.02	Фундаментальная информатика и
Направленность (профиль) подготовки машинное обучение		Анализ данных научных исследований и
Квалификация выпускника	магистр	
Форма обучения	очная	

Иркутск 2022 г.

Согласовано с УМК Института математики
и информационных технологий
Протокол № 3 от «04» апреля 2022 г.

Председатель _____
Антоник В.Г.

Рекомендовано кафедрой Алгебраических и
информационных систем ИМИТ ИГУ:
Протокол № 9 От «24» марта 2022 г.

Зав. кафедрой _____
Пантелеев В.И.

1. Цели и задачи дисциплины

Цель

Освоение базовых методов анализа статистических данных, полученных с помощью измерений в разных измерительных шкалах.

Задачи:

- Изучить наиболее распространённые методы анализа данных;
- Уметь оценивать применимость конкретных методов анализа данных для имеющейся статистики;
- Сформировать навыки применения методов анализа данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

2.1. Учебная дисциплина (модуль) относится к обязательной части программы и изучается на первом курсе.

2.2. Для изучения данной учебной дисциплины (модуля) необходимы знания, умения и навыки, сформированные математикой, теорией вероятностей и математической статистики, экономикой.

2.3. Перечень последующих учебных дисциплин, для которых необходимы знания, умения и навыки, формируемые данной учебной дисциплиной: все виды практик, машинное обучение, подготовка и защита ВКР.

3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс освоения дисциплины направлен на формирование компетенций (элементов следующих компетенций) в соответствии с ФГОС ВО по соответствующему направлению подготовки.

Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ОПК-1 Способен находить, формулировать и решать актуальные проблемы прикладной математики, фундаментальной информатики и информационных технологий	ИДК опк1.1 Умеет выделять проблемы, относящиеся к прикладной математике, фундаментальной информатике и информационным технологиям	Знает наиболее распространенные модели и методы анализа данных. Умеет оценить применимость моделей и методов анализа данных для решения проблемы. Владеет методами анализа данных.
	ИДК опк1.2 Умеет решать актуальные проблемы прикладной математики, фундаментальной информатики и информационных технологий	Знает наиболее распространенные модели и методы анализа данных. Умеет оценить применимость моделей и методов анализа данных для решения проблемы. Владеет методами анализа данных.
	ИДК опк1.3 Способен формулировать проблемы прикладной математики,	Знает наиболее распространенные модели и методы анализа данных.

	фундаментальной информатики и информационных технологий	Умеет оценить применимость моделей и методов анализа данных для решения проблемы. Владеет методами анализа данных.
ОПК-3 Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования	ИДК опк3.1 Способен строить математические модели для решения задач профессиональной деятельности	Знает модели анализа данных. Умеет строить модели процессов с помощью анализа данных. Владеет методами анализа данных для решения прикладных задач.
	ИДК опк3.2 Способен проводить анализ математических моделей, выбирать оптимальные	Знает модели анализа данных. Умеет строить модели процессов с помощью анализа данных. Владеет методами анализа данных для решения прикладных задач.
	ИДК опк3.3 Способен создавать инновационные методы решения прикладных задач профессиональной деятельности	Знает модели анализа данных. Умеет строить модели процессов с помощью анализа данных. Владеет методами анализа данных для решения прикладных задач.

4. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Объем дисциплины составляет 5 зачетных единиц, 180 часов, практическая подготовка __34__.

Форма промежуточной аттестации: 1 семестр - зачет.

4.1. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ, С УКАЗАНИЕМ ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ И ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ

№ п/п	Раздел дисциплины/темы	Се мес тр	Виды учебной работы, включая самостоятельную работу обучающихся и трудоемкость (в часах)				Формы текущего контроля успеваемости
			Контактная работа преподавателя с обучающимися			Самостоя тельная работа + контроль	
			Лекции	Семинарск ие (практичес кие занятия)	Контроль обучения		
1	Измерения, статистики и первичная обработка данных		2	2	1	10	Устный опрос
2	Меры центральной тенденции, изменчивости и связи		2	2	1	12	Устный опрос
3	Однофакторный и двухфакторный дисперсионный анализ		2	4	1	15	Устный опрос
4	Многофакторный регрессионный анализ		2	6	1	15	Устный опрос
5	Факторный анализ		2	4	1	15	Устный опрос
6	Кластерный анализ		2	6	1	20	Устный опрос
7	Дискриминантный анализ		2	4	1	15	Устный опрос
8	Многомерное шкалирование		2	6	1	20	Устный опрос
Итого часов			16	34	8	122	

4.2. План внеаудиторной самостоятельной работы обучающихся по дисциплине

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы Список источников
		Вид самостоятельной работы	Сроки выполнения	Затраты времени (час.)		
1	Измерения, статистики и первичная обработка данных	УИЛ	В соответствии с темой	10	Устный опрос	Список источников
1	Меры центральной тенденции, изменчивости и связи	УИЛ	В соответствии с темой	12	Устный опрос	Список источников
1	Однофакторный и двухфакторный дисперсионный анализ	УИЛ	В соответствии с темой	15	Устный опрос	Список источников
1	Многофакторный регрессионный анализ	УИЛ	В соответствии с темой	15	Устный опрос	Список источников
1	Факторный анализ	УИЛ	В соответствии с темой	15	Устный опрос	Список источников
1	Кластерный анализ	УИЛ	В соответствии с темой	20	Устный опрос	Список источников
1	Дискриминантный анализ	УИЛ	В соответствии с темой	15	Устный опрос	Список источников

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы Список источников
		Вид самостоятельной работы	Сроки выполнения	Затраты времени (час.)		
1	Многомерное шкалирование	УИЛ	В соответствии с темой	20	Устный опрос	
Общая трудоемкость самостоятельной работы по дисциплине (час)				122		
Из них объем самостоятельной работы с использованием электронного обучения и дистанционных образовательных технологий (час)				122		

У-выполнение упражнений, И- информационный поиск, Л- изучение литературы

4.3. СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

1. Методы анализа данных.

Тема 1. Измерения, статистики и первичная обработка данных: измерения, измерительные шкалы, статистики, первичная обработка данных, визуализация.

Тема 2. Меры центральной тенденции, изменчивости и связи: мода, медиана, среднее значение, бимодальное распределение, дисперсия, стандартное отклонение, коэффициенты вариации, корреляции, бисериальный коэффициент корреляции.

Тема 3. Однофакторный и двухфакторный дисперсионный анализ: качественные и количественные переменные, постоянные эффекты, метод Тьюки, метод Шеффе, карта различий.

Тема 4. Многофакторный регрессионный анализ: регрессия, мультиколлинеарность, средняя ошибка аппроксимации, стандартизованные коэффициенты регрессии, критерии Фишера, Стьюдента.

Тема 5. Факторный анализ: латентные характеристики, метод главных компонент, однофакторная модель Спирмена, метод главных факторов, матрица корреляций, вращение осей.

Тема 6. Кластерный анализ: меры сходства, иерархический кластерный анализ, метод к-средних, метод поиска сгущений, критерии качества классификации.

Тема 7. Дискриминантный анализ: дискриминантная функция, дискриминантные переменные, обучающие выборки, классификация объектов.

Тема 8. Многомерное шкалирование: многомерное шкалирование Торгерсона, метрическое шкалирование, неметрическое шкалирование, анализ предпочтений.

4.3.1. Перечень семинарских, практических занятий и лабораторных работ

№ п/п	№ раздела и темы	Наименование семинаров, практических и лабораторных работ	Трудоемкость (час.)		Оценочные средства	Формируемые компетенции (индикаторы)*
			Всего часов	Из них практичес- кая подготовк а		
1	2	3	4	5	6	7
1	1	Измерения, статистики и первичная обработка данных	2	2	Устный опрос	ОПК-1, ОПК-3
2	2	Меры центральной тенденции, изменчивости и связи	2	2	Устный опрос	ОПК-1, ОПК-3
3	3	Однофакторный и двухфакторный дисперсионный анализ	4	4	Устный опрос	ОПК-1, ОПК-3
4	4	Многофакторный регрессионный анализ	6	6	Устный опрос	ОПК-1, ОПК-3
5	5	Факторный анализ	4	4	Устный опрос	ОПК-1, ОПК-3
6	6	Кластерный анализ	6	6	Устный опрос	ОПК-1, ОПК-3
7	7	Дискриминантный анализ	4	4	Устный опрос	ОПК-1, ОПК-3
8	8	Многомерное шкалирование	6	6	Устный опрос	ОПК-1, ОПК-3
		Всего	34			

4.3.2. Перечень тем (вопросов), выносимых на самостоятельное изучение студентами в рамках самостоятельной работы (СР)

Измерения, статистики и первичная обработка данных
Меры центральной тенденции, изменчивости и связи
Однофакторный и двухфакторный дисперсионный анализ
Многофакторный регрессионный анализ
Факторный анализ
Кластерный анализ
Дискриминантный анализ
Многомерное шкалирование

4.4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Методические указания расположены в ИОС Educa.

4.5. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ РАБОТ (ПРОЕКТОВ)

Не предусмотрено.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

а) перечень литературы

Основная:

Анализ данных : учебник для вузов / В. С. Мхитарян [и др.]; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2022. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт].

Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2022. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт].

Толстова Ю.Н., Основы многомерного шкалирования [Текст] : учеб. пособие для студ. вузов / Ю. Н. Толстова. - М. : Университет, 2006. - 157 с.

Дополнительная:

Гласс Дж., Стэнли Д. Статистические методы в педагогике и психологии [Текст] : научное издание / Д. Гласс, Д. Стэнли ; пер. с англ. Л. И. Хайрусова, ред., авт. послесл. Ю. П. Адлер, авт. послесл. А. Н. Ковалев. - М. : Прогресс, 1976. - 495 с.

Многомерный статистический анализ в экономике [Текст] : учеб. пособие для студ. вузов / Л.А.Сошникова, В.Н.Тамашевич, Г.Уебе, М.Шефер ; Под ред. В.Н. Тамашевича. - М. : Юнити-Дана, 1999. - 598 с.

б) периодические издания

в) список авторских методических разработок:

г) базы данных, информационно-справочные и поисковые системы

1. Единое окно доступа к образовательным ресурсам. Полнотекстовая электронная библиотека учебных и учебно-методических материалов (федеральный ресурс). <http://www.window.edu.ru>.

2. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.biblio-online.ru/>

3. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>

4. ИОС ИГУ EDUca

5. Онлайн-курсы от ведущих вузов и компаний страны <https://welcome.stepik.org/ru>

6. Образовательный онлайн-проект <https://www.coursera.org/>

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. УЧЕБНО-ЛАБОРАТОРНОЕ ОБОРУДОВАНИЕ:

Для проведения лекционных занятий необходима аудитория с презентационным оборудованием, для проведения практических занятий необходима аудитория на 15-30 рабочих мест (в зависимости от численности учебной группы), оборудованная доской, презентационной техникой.

6.2. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ:

Microsoft office

6.3. ТЕХНИЧЕСКИЕ И ЭЛЕКТРОННЫЕ СРЕДСТВА:

ИОС EDUCA, DOMIC, презентационное оборудование, персональный компьютер с возможностью демонстрации презентаций в формате pdf.

7. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При реализации данного курса используются следующие образовательные технологии: технологии традиционного обучения, игровые технологии, технологии проблемного обучения, технологии обучения в сотрудничестве, технологии контекстного обучения,

интерактивные технологии, технологии дистанционного обучения, активные педагогические технологии.

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

8.1. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ВХОДНОГО КОНТРОЛЯ

8.2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ТЕКУЩЕГО КОНТРОЛЯ

Вопросы для контроля

Анализ данных.

Сравнение измерительных шкал.

Мультиколлинеарность и способы устранения.

Оценка качества регрессионной модели.

Интерпретация коэффициентов регрессии.

Для чего нужны стандартизованные коэффициенты регрессии?

Что оценивает дисперсионный анализ?

Как построить карту различий в дисперсионном анализе?

Что такое – латентная переменная?

Проблема вращений в факторном анализе.

Классификация кластерных методов.

Задачи многомерного шкалирования.

Уровень значимости.

Число степеней свободы.

Неметрическое шкалирование.

Показатели центральной тенденции.

Показатели изменчивости.

Интервальное представление данных.

8.3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПРОМЕЖУТОЧНОГО КОНТРОЛЯ

Вопросы к зачету

1. Измерения и измерительные шкалы.
2. Агрегированные характеристики набора данных.
3. Задачи анализа данных.
4. Классификация методов анализа данных.
5. Задачи дисперсионного анализа.
6. Метод Тьюки.
7. Метод Шеффе.
8. Задачи дискриминантного анализа.
9. Классификация объектов в дисперсионном анализе.
10. Подбор параметров и оценка качества результатов в регрессионном анализе.
11. Трудности в построении регрессионной модели.
12. Задачи факторного анализа.
13. Варианты факторного анализа.
14. Задачи кластерного анализа.
15. Алгоритмы кластерного анализа.
16. Задачи многомерного шкалирования.
17. Метрическое многомерное шкалирование.
18. Неметрическое многомерное шкалирование.
19. Прогнозирование процессов с помощью результатов анализа данных.
20. Подбор метода анализа данных.

Разработчики: _____ _профессор, д.э.н.____ _Амбросов Н.В.
(подпись)  (занимаемая должность) (Ф.И.О.)

Программа составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 02.04.02 «Фундаментальная информатика и информационные технологии» (уровень магистратуры), утвержденный приказом Министерства образования и науки Российской Федерации от «23» августа 2017 г. № 811, зарегистрированный в Минюсте России «13» сентября 2017 г. № 48168 с изменениями и дополнениями от 26 ноября 2020 г., 8 февраля 2021 г.

Программа рассмотрена на заседании кафедры Алгебраических и информационных систем ИМИТ ИГУ «24» марта 2022 г.

Протокол № 9 Зав. кафедрой  Пантелеев В.И.