

УТВЕРЖДАЮ
Проректор по УМР

_____ Е.В. Коновалова

15 июня 2023 г., протокол УМС №5

МОДУЛЬ ДИСЦИПЛИН ПРОФИЛЬНОЙ НАПРАВЛЕННОСТИ

Анализ данных

рабочая программа дисциплины (модуля)

Закреплена за кафедрой	Прикладной математики
Учебный план	b010302-ТехнолПрог-23-1.plx Направление 01.03.02 ПРИКЛАДНАЯ МАТЕМАТИКА И ИНФОРМАТИКА Направленность (профиль): Технологии программирования и анализ данных
Квалификация	бакалавр
Форма обучения	очная
Общая трудоемкость	4 ЗЕТ

Часов по учебному плану	144
в том числе:	
аудиторные занятия	64
самостоятельная работа	44
часов на контроль	36

Виды контроля в семестрах:
экзамены 7

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	7 (4.1)		Итого	
	17 2/6			
Неделя	уп	рп	уп	рп
Лекции	32	32	32	32
Лабораторные	32	32	32	32
Итого ауд.	64	64	64	64
Контактная работа	64	64	64	64
Сам. работа	44	44	44	44
Часы на контроль	36	36	36	36
Итого	144	144	144	144

Программу составил(и):

к.ф.-м.н., доцент, Гореликов А.В.

Рабочая программа дисциплины

Анализ данных

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 01.03.02 Прикладная математика и информатика (приказ Минобрнауки России от 10.01.2018 г. № 9)

составлена на основании учебного плана:

Направление 01.03.02 ПРИКЛАДНАЯ МАТЕМАТИКА И ИНФОРМАТИКА

Направленность (профиль): Технологии программирования и анализ данных

утвержденного учебно-методическим советом вуза от 15.06.2023 протокол № 5.

Рабочая программа одобрена на заседании кафедры

Прикладной математики

Зав. кафедрой *к.ф.-м.н., доцент Гореликов А.В.*

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	
1.1	Формирование у обучающихся знаний теоретических и прикладных основ анализа данных, видов аналитики, методов и инструментальных средств анализа больших данных, технологий анализа больших данных.
1.2	Формирование у обучающихся навыка программирования на языках высокого уровня, ориентированных на работу с большими данными.
1.3	Формирование у обучающихся навыка аналитической работы с использованием методов, инструментальных средств и технологий анализа больших данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП	
Цикл (раздел) ООП:	Б1.В.01
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Методы машинного обучения
2.1.2	Методы оптимизации
2.1.3	Объектно-ориентированное программирование
2.1.4	Численные методы
2.1.5	Разработка программного обеспечения в ОС Linux
2.1.6	Базы данных
2.1.7	Математический анализ
2.1.8	Дискретная математика
2.1.9	Теория вероятностей и математическая статистика
2.1.10	Алгоритмы и структуры данных
2.1.11	Экономика и управление предприятием
2.1.12	Дифференциальные уравнения
2.1.13	Математическая логика и теория алгоритмов
2.1.14	Комбинаторика и теория графов
2.1.15	Введение в профессиональную деятельность
2.1.16	Информатика
2.1.17	Алгебра и геометрия
2.1.18	Программирование
2.1.19	Искусственный интеллект
2.1.20	Информационные технологии
2.1.21	Операционные системы
2.2	Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Выполнение и защита выпускной квалификационной работы
2.2.2	Производственная практика, преддипломная практика
2.2.3	Производственная практика, научно-исследовательская работа
2.2.4	Визуализация данных
2.2.5	Подготовка к сдаче и сдача государственного экзамена
2.2.6	Финансовая математика
2.2.7	Теория игр и исследование операций

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)
ПК-4.1: Понимает теоретические и прикладные основы анализа данных, видов аналитики, методов и инструментальных средств анализа больших данных, технологий анализа больших данных, методов интерпретации и визуализации больших данных
ПК-4.2: Проводит аналитические работы с использованием методов и технологий больших данных
ПК-4.3: Программирует на языках высокого уровня, ориентированных на работу с большими данными

ПК-2.1: Понимает и анализирует цели и задачи научно-исследовательских разработок

ПК-2.2: Проводит научно-исследовательские разработки по отдельным разделам темы проекта

ПК-2.3: Составляет отчеты по результатам исследований и разработок и оценивает полученные результаты

ПК-1.1: Собирает и обрабатывает научно-техническую информацию с использованием существующих информационно-коммуникационных технологий

ПК-1.2: Анализирует и обобщает результаты и опыт передовых исследований в соответствующей области знаний

В результате освоения дисциплины обучающийся должен

3.1 Знать:	
3.1.1	теоретические и прикладные основы анализа данных, видов аналитики, методов и инструментальных средств анализа больших данных, технологий анализа больших данных.
3.2 Уметь:	
3.2.1	проводить аналитические работы с использованием методов и технологий больших данных;
3.2.2	программировать на языках высокого уровня, ориентированных на работу с большими данными;
3.2.3	понимать и анализировать цели и задачи научно-исследовательских разработок в области анализа больших данных;
3.2.4	собирать и обрабатывать научно-техническую информацию с использованием существующих информационно-коммуникационных технологий;
3.2.5	анализировать и обобщать результаты и опыт передовых исследований в области анализа данных;
3.2.6	проводить научно-исследовательские разработки по отдельным разделам темы проекта в области анализа данных и оформлять результаты исследований и разработок.
3.3 Владеть:	
3.3.1	навыком аналитической работы с использованием методов и технологий больших данных.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература	Примечание
	Раздел 1. Линейная регрессия, свойства метода наименьших квадратов					
1.1	Коэффициент детерминации (R^2), информационные критерии (AIC, BIC), метрики (MSE, MAE, MAPE). Гауссовская линейная модель - доверительные интервалы для коэффициентов модели и для отклика, гипотезы о незначимости коэффициента и группы коэффициентов, общая линейная гипотеза, сравнение моделей. /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	

1.2	Коэффициент детерминации (R^2), информационные критерии (AIC, BIC), метрики (MSE, MAE, MAPE). Гауссовская линейная модель - доверительные интервалы для коэффициентов модели и для отклика, гипотезы о незначимости коэффициента и группы коэффициентов, общая линейная гипотеза, сравнение моделей. /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
1.3	Коэффициент детерминации (R^2), информационные критерии (AIC, BIC), метрики (MSE, MAE, MAPE). Гауссовская линейная модель - доверительные интервалы для коэффициентов модели и для отклика, гипотезы о незначимости коэффициента и группы коэффициентов, общая линейная гипотеза, сравнение моделей. /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 2. Анализ остатков						
2.1	Дисперсия остатков линейной модели в условиях гетероскедастичности, визуальный анализ. Критерии проверки на гомоскедастичность: Бройша-Пагана и Голдфелда-Квандта. Преобразование Бокса-Кокса. Устойчивые оценки дисперсии Уайта, асимптотическая нормальность /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
2.2	Дисперсия остатков линейной модели в условиях гетероскедастичности, визуальный анализ. Критерии проверки на гомоскедастичность: Бройша-Пагана и Голдфелда-Квандта. Преобразование Бокса-Кокса. Устойчивые оценки дисперсии Уайта, асимптотическая нормальность /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
2.3	Дисперсия остатков линейной модели в условиях гетероскедастичности, визуальный анализ. Критерии проверки на гомоскедастичность: Бройша-Пагана и Голдфелда-Квандта. Преобразование Бокса-Кокса. Устойчивые оценки дисперсии Уайта, асимптотическая нормальность /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 3. Обобщенная линейная модель, статистические свойства оценки коэффициентов, построение доверительных интервалов						
3.1	Частные случаи - пуассоновская регрессии. (лог. регрессия в МЛ). Байесовский классификатор. Построение оптимального байесовского классификатора с док-вом теоремы об оптимальности. Квадратичный и линейный дискриминантный анализ. Оценки параметров, вид разделяющей поверхности. Наивный байесовский классификатор. /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	

3.2	Частные случаи - пуассоновская регрессии. (лог. регрессия в МЛ). Байесовский классификатор. Построение оптимального байесовского классификатора с док-вом теоремы об оптимальности. Квадратичный и линейный дискриминантный анализ. Оценки параметров, вид разделяющей поверхности. Наивный байесовский классификатор /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
3.3	Частные случаи - пуассоновская регрессии. (лог. регрессия в МЛ). Байесовский классификатор. Построение оптимального байесовского классификатора с док-вом теоремы об оптимальности. Квадратичный и линейный дискриминантный анализ. Оценки параметров, вид разделяющей поверхности. Наивный байесовский классификатор /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 4. Пропуски в данных - типы пропусков, методы работы						
4.1	Робастная регрессия. Методы на основе ближайшего соседа - kNN, взвешенный kNN, их свойства. Непараметрическая регрессия, локальное усреднение, оценка Надарая-Ватсона. Условия сходимости оценки Надарая-Ватсона, выбор ширины ядра, доверительная лента. Локальная линейная регрессионная модель. Регрессионное дерево, метод построения, свойства. Случайный лес и его свойства. /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
4.2	Робастная регрессия. Методы на основе ближайшего соседа - kNN, взвешенный kNN, их свойства. Непараметрическая регрессия, локальное усреднение, оценка Надарая-Ватсона. Условия сходимости оценки Надарая-Ватсона, выбор ширины ядра, доверительная лента. Локальная линейная регрессионная модель. Регрессионное дерево, метод построения, свойства. Случайный лес и его свойства. /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
4.3	Робастная регрессия. Методы на основе ближайшего соседа - kNN, взвешенный kNN, их свойства. Непараметрическая регрессия, локальное усреднение, оценка Надарая-Ватсона. Условия сходимости оценки Надарая-Ватсона, выбор ширины ядра, доверительная лента. Локальная линейная регрессионная модель. Регрессионное дерево, метод построения, свойства. Случайный лес и его свойства. /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 5. Причины избыточности информации в данных, типы методов снижения размерности						

5.1	Метод главных компонент (РСА) как выбор направлений с максимальной дисперсией, формулы перехода в сжатое пространство и обратно. Дисперсии образа, выбор размерности сжатого пространства на основе доли необъясненной дисперсии /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
5.2	Метод главных компонент (РСА) как выбор направлений с максимальной дисперсией, формулы перехода в сжатое пространство и обратно. Дисперсии образа, выбор размерности сжатого пространства на основе доли необъясненной дисперсии /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
5.3	Метод главных компонент (РСА) как выбор направлений с максимальной дисперсией, формулы перехода в сжатое пространство и обратно. Дисперсии образа, выбор размерности сжатого пространства на основе доли необъясненной дисперсии /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
	Раздел 6. Теорема об SVD-разложении. Док-во существования SVD-разложения					
6.1	Методы SNE и t-SNE: первоначальный вариант SNE, симметричный SNE, проблема скученности, метод t-SNE как решение проблемы. Метод UMAP. Постановка задачи: графы, функционал качества (KL). Общие слова о том, какая “метрика” используется, и почему в этом случае нет проблемы проклятия размерности /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
6.2	Методы SNE и t-SNE: первоначальный вариант SNE, симметричный SNE, проблема скученности, метод t-SNE как решение проблемы. Метод UMAP. Постановка задачи: графы, функционал качества (KL). Общие слова о том, какая “метрика” используется, и почему в этом случае нет проблемы проклятия размерности /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
6.3	Методы SNE и t-SNE: первоначальный вариант SNE, симметричный SNE, проблема скученности, метод t-SNE как решение проблемы. Метод UMAP. Постановка задачи: графы, функционал качества (KL). Общие слова о том, какая “метрика” используется, и почему в этом случае нет проблемы проклятия размерности /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
	Раздел 7. Коэффициенты корреляции Пирсона, Спирмена и Кендалла, их свойства					
7.1	Таблицы сопряженности 2x2, точный тест Фишера, меры взаимосвязи, определение количества наблюдений. Общий случай таблиц сопряженности, типы вероятностных моделей, критерий хи-квадрат. Влияние признаков на целевую переменную: корреляции, подход с помощью решающих деревьев – важность признаков на основе Mean Decrease Impurity, Permutation feature importance, Drop Column feature importance. /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	

7.2	Таблицы сопряженности 2x2, точный тест Фишера, меры взаимосвязи, определение количества наблюдений. Общий случай таблиц сопряженности, типы вероятностных моделей, критерий хи-квадрат. Влияние признаков на целевую переменную: корреляции, подход с помощью решающих деревьев – важность признаков на основе Mean Decrease Impurity, Permutation feature importance, Drop Column feature importance. /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
7.3	Таблицы сопряженности 2x2, точный тест Фишера, меры взаимосвязи, определение количества наблюдений. Общий случай таблиц сопряженности, типы вероятностных моделей, критерий хи-квадрат. Влияние признаков на целевую переменную: корреляции, подход с помощью решающих деревьев – важность признаков на основе Mean Decrease Impurity, Permutation feature importance, Drop Column feature importance. /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 8. Виды задач дисперсионного анализа, примеры						
8.1	Критерии проверки однородности для бернуллиевских выборок, доверительные интервалы для разности (простые и Уилсона). Проверка на равенство средних нормальных выборок (t-test, 3 сл.), проверка равенства дисперсий, проверка однородности нормальных выборок. АВ -тестирование. Принципы разбиения, особенности. АА-тесты. Разбиение на тестовые группы, сроки теста, проверка нескольких гипотез. Пример, в котором события, соответствующие одному пользователю, зависимы. Бакетное семплирование как способ решения проблемы /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
8.2	Критерии проверки однородности для бернуллиевских выборок, доверительные интервалы для разности (простые и Уилсона). Проверка на равенство средних нормальных выборок (t-test, 3 сл.), проверка равенства дисперсий, проверка однородности нормальных выборок. АВ -тестирование. Принципы разбиения, особенности. АА-тесты. Разбиение на тестовые группы, сроки теста, проверка нескольких гипотез. Пример, в котором события, соответствующие одному пользователю, зависимы. Бакетное семплирование как способ решения проблемы /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	



8.3	Критерии проверки однородности для бернуллиевских выборок, доверительные интервалы для разности (простые и Уилсона). Проверка на равенство средних нормальных выборок (t-test, 3 сл.), проверка равенства дисперсий, проверка однородности нормальных выборок. АВ -тестирование. Принципы разбиения, особенности. АА-тесты. Разбиение на тестовые группы, сроки теста, проверка нескольких гипотез. Пример, в котором события, соответствующие одному пользователю, зависимы. Бакетное семплирование как способ решения проблемы /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 9. Виды альтернатив в непараметрическом случае						
9.1	Критерии Смирнова и Розенблатта. Критерий Уилкоксона-Манна-Уитни, его свойства, связанная с ним оценка параметра сдвига. Связные выборки, предположения модели, пример, когда предположения не выполняются. Критерий знаков, его свойства, связанная с ним оценка параметра сдвига. Критерий ранговых сумм Уилкоксона, его свойства, связанная с ним оценка параметра сдвига. Проверка симметрии /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
9.2	Критерии Смирнова и Розенблатта. Критерий Уилкоксона-Манна-Уитни, его свойства, связанная с ним оценка параметра сдвига. Связные выборки, предположения модели, пример, когда предположения не выполняются. Критерий знаков, его свойства, связанная с ним оценка параметра сдвига. Критерий ранговых сумм Уилкоксона, его свойства, связанная с ним оценка параметра сдвига. Проверка симметрии /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
9.3	Критерии Смирнова и Розенблатта. Критерий Уилкоксона-Манна-Уитни, его свойства, связанная с ним оценка параметра сдвига. Связные выборки, предположения модели, пример, когда предположения не выполняются. Критерий знаков, его свойства, связанная с ним оценка параметра сдвига. Критерий ранговых сумм Уилкоксона, его свойства, связанная с ним оценка параметра сдвига. Проверка симметрии /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 10. Комбинирование критериев для построения более мощных процедур на примере одновременной проверки на нормальность и однородность двух выборок с условием на контроль FWER						

10.1	Сравнение интенсивностей двух экспоненциальных выборок. Сравнение интенсивностей пуассоновских процессов. Перестановочные критерии - идея, примеры для гипотез о среднем, а также для гипотез о равенстве средних двух выборок. Множественная проверка гипотез с помощью перестановок: версия max-T, обобщенный вариант /Лек/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
10.2	Сравнение интенсивностей двух экспоненциальных выборок. Сравнение интенсивностей пуассоновских процессов. Перестановочные критерии - идея, примеры для гипотез о среднем, а также для гипотез о равенстве средних двух выборок. Множественная проверка гипотез с помощью перестановок: версия max-T, обобщенный вариант /Лаб/	7	2	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
10.3	Сравнение интенсивностей двух экспоненциальных выборок. Сравнение интенсивностей пуассоновских процессов. Перестановочные критерии - идея, примеры для гипотез о среднем, а также для гипотез о равенстве средних двух выборок. Множественная проверка гипотез с помощью перестановок: версия max-T, обобщенный вариант /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
	Раздел 11. Однофакторный дисперсионный анализ для случая независимых выборок					
11.1	F-критерий и критерий Бартлетта, их применимость. Критерий Краскела-Уоллиса и Джонкхиера. Post-hoc анализ: LSD Фишера, HSD Тьюки, критерии Неменья и Данна, оценка контраста. Однофакторный дисперсионный анализ для случая связанных выборок. F-критерий, критерии Фридмана и Пейджа. Post-hoc анализ. Двухфакторный дисперсионный анализ, случай дополнительной контрольной группы. /Лек/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
11.2	F-критерий и критерий Бартлетта, их применимость. Критерий Краскела-Уоллиса и Джонкхиера. Post-hoc анализ: LSD Фишера, HSD Тьюки, критерии Неменья и Данна, оценка контраста. Однофакторный дисперсионный анализ для случая связанных выборок. F-критерий, критерии Фридмана и Пейджа. Post-hoc анализ. Двухфакторный дисперсионный анализ, случай дополнительной контрольной группы. /Лаб/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	

11.3	F-критерий и критерий Бартлетта, их применимость. Критерий Краскела-Уоллиса и Джонкхиера. Post-hoc анализ: LSD Фишера, HSD Тьюки, критерии Неменья и Данна, оценка контраста. Однофакторный дисперсионный анализ для случая связанных выборок. F-критерий, критерии Фридмана и Пейджа. Post-hoc анализ. Двухфакторный дисперсионный анализ, случай дополнительной контрольной группы. /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 12. Практическая аналитика						
12.1	Какие особенности в данных могут присутствовать? Воронка. Парадокс Симпсона, примеры и выводы. Контрафактивная модель, причинно-следственный эффект, статистическая связь, утверждение о том, что связь не есть причинность. Равенство величины причинно-следственного эффекта и статистической связи при случайном назначении воздействия. Контрафактивная модель на примере парадокса Симпсона /Лек/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
12.2	Какие особенности в данных могут присутствовать? Воронка. Парадокс Симпсона, примеры и выводы. Контрафактивная модель, причинно-следственный эффект, статистическая связь, утверждение о том, что связь не есть причинность. Равенство величины причинно-следственного эффекта и статистической связи при случайном назначении воздействия. Контрафактивная модель на примере парадокса Симпсона /Лаб/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
12.3	Какие особенности в данных могут присутствовать? Воронка. Парадокс Симпсона, примеры и выводы. Контрафактивная модель, причинно-следственный эффект, статистическая связь, утверждение о том, что связь не есть причинность. Равенство величины причинно-следственного эффекта и статистической связи при случайном назначении воздействия. Контрафактивная модель на примере парадокса Симпсона /Ср/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 13. Ориентированные ациклические графы, терминология						
13.1	Марковское распределение на графе, примеры. Условная независимость и ее свойства. Оценка распределений в графе методом максимального правдоподобия. Интервенция, средний условный эффект как способ оценки причинно-следственного эффекта по графу. Примеры. Связь оценки причинно-следственного эффекта методом интервенции с контрафактивной моделью. /Лек/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	

13.2	Марковское распределение на графе, примеры. Условная независимость и ее свойства. Оценка распределений в графе методом максимального правдоподобия. Интервенция, средний условный эффект как способ оценки причинно-следственного эффекта по графу. Примеры. Связь оценки причинно-следственного эффекта методом интервенции с контрафактивной моделью. /Лаб/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
13.3	Марковское распределение на графе, примеры. Условная независимость и ее свойства. Оценка распределений в графе методом максимального правдоподобия. Интервенция, средний условный эффект как способ оценки причинно-следственного эффекта по графу. Примеры. Связь оценки причинно-следственного эффекта методом интервенции с контрафактивной моделью. /Ср/	7	4	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
Раздел 14. Терминология в ориентированных ациклических графах						
14.1	Марковское свойство, примеры. Свойства d-разделимости и d-связности, теорема об условной независимости на множестве вершин. Построение причинно-следственных графов по данным: метод индуктивной причинности. Оценка условной независимости: частная корреляция, причинность по Грейнджеру /Лек/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
14.2	Марковское свойство, примеры. Свойства d-разделимости и d-связности, теорема об условной независимости на множестве вершин. Построение причинно-следственных графов по данным: метод индуктивной причинности. Оценка условной независимости: частная корреляция, причинность по Грейнджеру /Лаб/	7	3	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
14.3	Марковское свойство, примеры. Свойства d-разделимости и d-связности, теорема об условной независимости на множестве вершин. Построение причинно-следственных графов по данным: метод индуктивной причинности. Оценка условной независимости: частная корреляция, причинность по Грейнджеру /Ср/	7	4	ПК-4.1 ПК-4.2 ПК-4.3 ПК-1.1 ПК-1.2 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
14.4	/Контр.раб./	7	0	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	
14.5	/Экзамен/	7	36	ПК-4.1 ПК-4.2 ПК-4.3 ПК-2.1 ПК-2.2 ПК-2.3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Л2.4Л3.1 Л3.2 Э1 Э2 Э3 Э4 Э5	

5. ОЦЕНОЧНЫЕ СРЕДСТВА
5.1. Оценочные материалы для текущего контроля и промежуточной аттестации
Представлены отдельным документом
5.2. Оценочные материалы для диагностического тестирования
Представлены отдельным документом

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)				
6.1. Рекомендуемая литература				
6.1.1. Основная литература				
	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л1.1	Маккинли, Уэс, Слинкина, А.	Python и анализ данных	Саратов: Профобразование, 2019, электронный ресурс	1
Л1.2	Волкова, В. М., Семёнова, М. А., Четвертакова, Е. С., Вожов, С. С.	Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python: учебное пособие	Новосибирск: Новосибирский государственный технический университет, 2017, электронный ресурс	1
6.1.2. Дополнительная литература				
	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Лемешко Б. Ю., Постовалов С. Н., Лемешко С. Б., Чимитова Е. В.	Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход	Москва: ООО "Научно- издательский центр ИНФРА-М", 2015, электронный ресурс	1
Л2.2	Мельниченко А. С.	Математическая статистика и анализ данных: Учебное пособие	Москва:  издательский Дом М  СиС, 2018, электронный ресурс	1
Л2.3	Сапожников П.Н., Макаров А.А.	Теория вероятностей, математическая статистика в примерах, задачах и тестах: Учебное пособие	Москва: ООО "КУРС", 2022, электронный ресурс	1
Л2.4	Зенков, А. В.	Математическая статистика в задачах и упражнениях: учебное пособие	Москва, Вологда: Инфра- Инженерия, 2022, электронный ресурс	1
6.1.3. Методические разработки				
	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л3.1	Гефан Г. Д., Ширяева Н. К.	Математическая статистика : практикум	Иркутск: ИрГУПС, 2018, электронный ресурс	1

	Авторы, составители	Заглавие	Издательство, год	Колич-во
ЛЗ.2	Фендель Т. В.	Математическая статистика в научных исследованиях: учебно-методическое пособия	Чайковский: ЧГИФК, 2017, электронный ресурс	1

6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Сайт разработчиков на Питоне - http://diveinto.python.ru/toc.html			
Э2	R — свободная программная среда для статистических вычислений и графики			
Э3	Научная электронная библиотека eLibrary			
Э4	курс МФТИ на НПОО "Анализ данных на практике"			
Э5	видеозаписи лекций и семинарских занятий Физтех-школы ФПМИ			

6.3.1 Перечень программного обеспечения

6.3.1.1	Операционная система семейства Linux.Интерпретатор языка Python 3.7 и выше, компилятор GCC 10 и выше, среда разработки PyCharm Community, текстовый редактор VSCodium (свободно-распространяемое программное обеспечение).			
6.3.1.2	Libre office.			

6.3.2 Перечень информационных справочных систем

6.3.2.1	http://www.garant.ru Информационно-правовой портал Гарант.ру			
6.3.2.2	http://www.consultant.ru/ Справочно-правовая система Консультант Плюс			

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1	Учебные аудитории для проведения занятий лекционного типа, групповых и индивидуальных консультаций, укомплектованные специализированной мебелью и техническими средствами обучения (доска, экран (стационарный или переносной), проектор (стационарный или переносной)). Учебные аудитории			
7.2	для проведения лабораторных занятий - компьютерный класс, оборудованный техникой из расчета один компьютер на одного обучающегося, с обустроенным рабочим местом преподавателя.			
7.3	Требуются персональные компьютеры, локальная вычислительная сеть с выходом в глобальную сеть Internet.			
7.4	Помещения для самостоятельной работы обучающихся, оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечения доступа в электронную информационно-образовательную среду организации.			