

Магистратура, кафедра статистического моделирования, набор 2023+

1 семестр

Вероятностные и статистические модели (Голяндина Н.Э.). Общий курс для ПМИ, содержит базовые сведения о теории вероятностей и математической статистике, а также их использовании для создания вероятностных моделей и принятия решений в рамках этих моделей на основе реальных данных.	
Многомерный анализ данных (Голяндина Н.Э.). Классические методы многомерного анализа. Это курс лекций, где даются базовые сведения по статистическому анализу многомерных данных. Основные изучаемые методы – анализ главных компонент (АГК), многомерная регрессия, классификация, кластерный анализ. Курс служит базой для других курсов, в том числе, курсов по машинному обучению, по временным рядам. В частности, АГК часто используется в качестве метода feature extraction в методах машинного обучения.	
Анализ многомерных данных на компьютере (Голяндина Н.Э.) Практика в поддержку курса по многомерному анализу данных. Целью курса является обучение применению теоретических знаний в области многомерной статистики к реальным данным с помощью статистических пакетов, в частности, языка R. В курсе будут рассмотрены вопросы правильного применения методов анализа данных, таких как дискриминантный анализ (классификация), кластерный анализ, анализ главных компонент. Будет даваться много практических заданий, выполнять которые нужно на R (предпочтительно) или Python, а на занятиях результаты будут обсуждаться в интерактивном режиме. Фактически, занятия являются поддерживающей практикой к курсу лекций Многомерный анализ данных. Предполагается знание базовой статистики: первичная статистика, проверка гипотез, меры зависимости. Перед началом основного курса идет вводная часть, где рассматриваются примеры применения базовой статистики.	Применение статистических методов к реальным данным (Алексеева Н.П.) Практика частично в поддержку курса по многомерному анализу данных. Целью изучения дисциплины «Применение статистических методов к реальным данным» является освоение студентами специализированных методов обработки статистических данных на примерах из биометрической практики. Будут рассмотрены следующие темы: Непараметрические методы статистического анализа, Статистические модели, Дисперсионный анализ и анализ выбросов, Задачи классификации, Систематизации и прогнозирования, Анализ данных типа времени жизни. На занятиях рассказывается теория и разбираются практические примеры. Зачет ставится на основе выполненных дома практических заданий.
Модели и методы исследования систем массового обслуживания (Мелас В.Б.) В теоретическом курсе рассматриваются модель реальных явления, таких как очереди, телефонные линии, прием заявок, в виде системы массового обслуживания. В курсе рассматриваются теория систем массового обслуживания (Queueing theory) и примеры ее применения.	Программирование для анализа данных (Шпилев П.В.) В курсе даются базовые сведения по языку R и Python для того, чтобы дать студентам возможность эффективно проводить анализ данных, который требуется для других курсов. Даются принципы повторяемых исследований.
Учебная практика 1 (проектно-технологическая)	

<p>Главные компоненты временных рядов (Голяндина Н.Э.) Курс посвящен методу SSA. Целью дисциплины «Главные компоненты временных рядов» является дать обучающемуся теоретические основы и методологию применения метода анализа сингулярного спектра (метода главных компоненты для временных рядов) для анализа и прогноза временных рядов. Будут рассмотрены вопросы выделения тренда, выделения периодик, разложения ряда на интерпретируемые компоненты, прогноз компонент ряда, оценивание параметров, обнаружение разладки (change point detection). Также рассмотрим многомерные обобщения метода, в частности, для анализа цифровых изображений.</p>	<p>Задачи математической статистики (Ермаков М.С.) Цель курса – дать введение в современные методы статистики. Реальный объем того, что рассказывается, определяется уровнем студентов. Может случиться, что курс будет содержать только базовые вещи. Если по максимуму, то задача курса - охватить основной материал годового курса Advanced Statistics, читаемого магистрам американских университетов. С материалами соответствующих курсов студенты знакомятся в процессе обучения. Курс на более глубоком уровне рассказывает о задачах параметрической статистики, изучает задачи семипараметрической статистики, излагает основные методы непараметрического оценивания с их сравнением, скоростью сходимости, эффективностью, адаптивностью, разреженностью, оракульными неравенствами. В курсе также дается общая математическая модель задач классификации и рассказывается о методах ее изучения. Все это излагается на достаточно популярном уровне.</p>
<p>Статистический анализ и прогноз временных рядов (Голяндина Н.Э.) Практика к лекциям по SSA и классические методы анализа временных рядов. Целью дисциплины является создание общего представления об анализе временных рядов в практических задачах. Будут рассмотрены задачи сглаживания, выделения тренда, сезонного разложения, фильтрации, прогноза методами скользящего среднего, авторегрессионным подходом, анализом сингулярного спектра, методами LOESS, STL и другими. Обучающийся должен получить представление, как правильно выбрать метод для конкретного временного ряда. В процессе изучения дисциплины каждый получит практическое задание анализа и прогноза реального временного ряда, которое выполняется на языке R или Python. Результаты выполнения обсуждаются в интерактивном режиме на занятиях.</p>	<p>Страховые и финансовые вероятностные модели (Товстик Т.М.) Большая часть курса посвящена страховой (актуарной) математике. Рассматриваются вопросы оценки вероятностей разорения для некоторых моделей страховых компаний, перестрахование, расчет удержаний, принципы расчета премий.</p>
<p>Интеллектуальный анализ данных (Гученко Р.А.) Целью дисциплины является формирование навыков исследования больших данных, направленных на решение типичных задач, возникающих во время статистического исследования с использованием пакетов языка R. Будут рассмотрены задачи классификации, регрессии, снижения размерности, кластеризации, resampling, линейные модели и регуляризация, нелинейные модели в задачах регрессии и классификации. На каждую тему обучающийся получает индивидуальное задание, которое должен выполнить на языке R и уметь объяснить и обосновать результаты.</p>	<p>Случайные поля и динамические методы Монте-Карло (Товстик Т.М.) Целью теоретического курса является освоение методов моделирования динамических случайных полей, изучение методов моделирования Гиббса и Метрополиса с использованием элементов отжига, использование полученных навыков для построения алгоритмов, приводящих к созданию эффективных программ для решения прикладных задач. Будут рассмотрены предельные теоремы для цепей Маркова, методы Монте-Карло при моделировании динамических полей, алгоритмы Метрополиса и приложения.</p>

Метод Монте-Карло и параллельные стохастические вычисления (Товстик Т.М.) Программа теоретического курса предусматривает ознакомление студентов магистратуры с основными особенностями современной вычислительной техники и специальными классами алгоритмов (Монте-Карло), позволяющих эту технику эффективно использовать.	Введение в имитационное моделирование: моделирование случайности (Некруткин В.В.) Принципы и характеристики моделирования. Табличные методы моделирования дискретных распределений. Дискретные распределения, зависящие от параметра (нетабличные методы). Общие методы моделирования (метод обратных функций, метод дискретной декомпозиции, метод отбора), моделирование наиболее часто встречающихся распределений. Курс для тех, у кого не было статистического моделирования в бакалавриате.
Производственная практика 1 (научно-исследовательская работа)	Производственная практика 1 (проектно-технологическая)

3 семестр

Статистическое и машинное обучение (Голяндина Н.Э.) Семинар проходит в форме докладов студентов с рассказом о теории и практике по след. темам: Регрессия. Регуляризация в регрессии. Классификация. Функция риска. Логистическая регрессия. Feature selection и extraction. Метод опорных векторов. Выбор модели с помощью кросс-валидации. Обучение без учителя. Разделение смеси распределений. Кластеризация. Вычислительные аспекты оптимизации. Гладкие функционалы и пр. Метод стохастического градиента как метод оптимизации. Примеры на основе одного из предыдущих методов. Решающие деревья. Random Forest. Композиция методов. Бустинг. Нейронные сети (особый класс функций для оптимизации), back propagation как вычислительный подход. Нейронные сети для изображений. Активное обучение. Тематическое моделирование. Глубокое обучение.	
История и методология прикладной математики и информатики (Ермаков С.М., Бурнаева Э.Г.) Доклады по теме курса.	
Цепи Маркова и их приложения (Некруткин В.В.)	Программные средства для многомерной классификации и регрессии (Шпилев П.В.)
Стохастические вычислительные методы (Мелас В.Б.)	Методы моделирования в задачах разработки мобильных приложений (Шпилев П.В.)
Учебная практика 2 (проектно-технологическая)	
Производственная практика (педагогическая)	

4 семестр

Научная и компьютерная коммуникация в современных условиях (Некруткин В.В., Мелас В.Б.) Доклады по теме ВКР	
Статистическое обучение для моделирования и прогнозирования временных рядов (Голяндина Н.Э.) На основной части занятий рассматривается модель поведения акций в виде броуновского движения плюс кусочно-линейный тренд. На занятиях обсуждается, как предсказывать такие ряды – зная модель, не зная модель, как применять методы машинного обучения к временным рядам. Также даются задания на соревнованиях в kaggle, где нужно набрать определенное число очков за преодоление различных benchmarks.	
Биостатистика (Алексеева Н.П.)	Модели и методы тропической математики (Кривулин Н.К.)
Производственная практика 2 (научно-исследовательская работа)	Производственная практика 2 (проектно-технологическая)