

**Федеральное государственное автономное образовательное  
учреждение высшего образования  
«Московский физико-технический институт  
(национальный исследовательский университет)»**

**УТВЕРЖДЕНО**

**Директор физтех-школы  
прикладной математики и  
информатики**

**А.М. Райгородский**

**Рабочая программа дисциплины (модуля)**

<b>по дисциплине:</b>	Прикладная статистика
<b>по направлению:</b>	Прикладная математика и информатика
<b>профиль подготовки:</b>	Современная комбинаторика центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск" кафедра дискретной математики
<b>курс:</b>	1
<b>квалификация:</b>	магистр

Семестр, формы промежуточной аттестации: 2 (весенний) - Дифференцированный зачет

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 75 час.

Всего часов: 135, всего зач. ед.: 3

Программу составили:

А.М. Райгородский, д-р физ.-мат. наук, доцент, ассистент

Н.А. Волков

Программа обсуждена на заседании кафедры дискретной математики 30.01.2025

## Аннотация

Курс посвящен основным методам статистического анализа реальных данных, которые подкрепляются необходимой математической теорией. В курсе разбираются как параметрические методы, так и непараметрические. Особое внимание уделяется особенностям работы с реальными данными, а также их подготовке. По итогам прохождения курса студенты смогут правильно формулировать гипотезы по реальным данным, выбирать статистический метод для решения задачи и составлять правильные выводы.

### 1. Цели и задачи

#### Цель дисциплины

Изучение математических и теоретических основ современного статистического анализа, а также подготовка слушателей к дальнейшей самостоятельной работе в области анализа статистических задач прикладной математики, физики, астрономии и педагогики.

#### Задачи дисциплины

- изучение математических основ прикладной статистики;
- приобретение слушателями теоретических знаний в области современного статистического анализа;
- приобретение слушателями навыков подготовки реальных данных к обработке и выбора статистического метода для решения задачи.

### 2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-2 Способен управлять проектом на всех этапах его жизненного цикла	УК-2.1 Формулирует в рамках обозначенной проблемы цель, задачи, актуальность, значимость (научную, практическую, методическую и иную в зависимости от типа проекта), ожидаемые результаты и возможные сферы их применения
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
ПК-3 Владеет навыками участия в научных дискуссиях, выступления с сообщениями и докладами, устного, письменного и виртуального (размещение в информационных сетях) характера, представления материалов собственных исследований, проведения корректуры, редактирования, реферирования работ	ПК-3.2 Умеет вести корректную дискуссию в области информационно-коммуникационных технологий и информационных систем, задавать вопросы и отвечать на поставленные вопросы по теме научной работы

### 3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны знать:

- фундаментальные понятия, законы теории вероятностей и математической статистики;
- понятия, аксиомы, методы доказательств и доказательства основных теорем курса;
- основные свойства объектов математической статистики;
- аналитические и численные подходы и методы для решения типовых прикладных задач теории вероятностей и математической статистики;
- параметрические и непараметрические методы статистического анализа;
- базовые понятия теории проверки статистических гипотез;
- основные подходы к сравнению оценок параметров как случайных величин;
- определение общей линейной регрессионной модели и метод наименьших квадратов;
- многомерное нормальное распределение и его основные свойства;
- методы оценки корреляционной зависимости;
- виды задач дисперсионного анализа.

уметь:

- использовать свои знания для решения прикладных задач теории вероятностей и математической статистики;
- оценивать корректность постановок задач теории вероятностей и математической статистики;
- самостоятельно находить алгоритмы решения задач, в том числе нестандартных, и проводить их анализ;
- грамотно и корректно представлять решения задач теории вероятностей и математической статистики в устной и письменной форме;
- внедрять математико-статистические методы исследований при решении прикладных задач физики, астрономии, математики и педагогики;
- самостоятельно углублять и расширять знания в области математической статистики.

владеть:

- культурой постановки, анализа и решения задач прикладной статистики, требующих для своего решения использования математических подходов и методов;
- предметным языком теории вероятностей и математической статистики, а также навыками грамотного описания решения задач и представления полученных результатов;
- основными методами математической статистики для построения точечных и доверительных оценок;
- навыками анализа статистических критериев;
- навыками применения теорем математической статистики в прикладных задачах физики, математики и педагогики.

#### 4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

##### 4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Виды задач дисперсионного анализа, примеры	2	2		10
2	Точечные оценки параметров, метод максимального правдоподобия	2	2		10
3	Сравнение оценок	2	2		8
4	Метод Ньютона и его применение для решения уравнения правдоподобия	2	2		8
5	Робастность, асимптотическая толерантность	2	2		7
6	Доверительные интервалы и доверительные области, метод бутстрепа	2	2		5

7	Ядерные оценки плотности	2	2		3
8	Проверка статистических гипотез	2	2		3
9	Критерий Вальда	2	2		3
10	Пример проверки гипотез в задаче исследовании эффективности нового лекарства	2	2		3
11	Поведение критического уровня критерия для разных размеров выборки	2	2		3
12	Общие критерии согласия	2	2		3
13	Линейная регрессия и коэффициенты корреляции	2	2		3
14	Виды задач дисперсионного анализа, примеры	2	2		3
15	Виды альтернатив в непараметрическом случае	2	2		3
Итого часов		30	30		75
Подготовка к экзамену		0 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

#### 4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 2 (Весенний)

##### 1. Виды задач дисперсионного анализа, примеры

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

##### 2. Точечные оценки параметров, метод максимального правдоподобия

Точечные оценки параметров. Статистики и оценки, примеры статистик. Свойства оценок - несмещенность, состоятельность, сильная состоятельность, асимптотическая нормальность. Практический смысл свойств. Примеры. Наследование свойств. Дельта-метод, пример. Метод максимального правдоподобия, примеры и свойства. Задача про гамма-излучение. Метод выборочных квантилей, асимптотическая нормальность выборочной квантили.

##### 3. Сравнение оценок

Сравнение оценок. Функция потерь и функция риска, MSE и MAE, примеры. Асимптотическая эффективность.

##### 4. Метод Ньютона и его применение для решения уравнения правдоподобия

Метод Ньютона и его применение для решения уравнения правдоподобия. Одношаговые оценки и их эквивалентность ОМП. Одношаговая оценка для модели Коши со сдвигом.

##### 5. Робастность, асимптотическая толерантность

Робастность, асимптотическая толерантность. Робастные оценки: усеченное среднее и медиана средних Уолша, их асимптотическая толерантность и асимптотическая эффективность по отношению к выборочному среднему на всем классе гладких симметричных распределений и в частных случаях. Робастность, асимптотическая толерантность. Робастные оценки: усеченное среднее и медиана средних Уолша, их асимптотическая толерантность и асимптотическая эффективность по отношению к выборочному среднему на всем классе гладких симметричных распределений и в частных случаях.

## 6. Доверительные интервалы и доверительные области, метод бутстрепа

Доверительные интервалы и доверительные области. Асимптотические доверительные интервалы, метод построения, пример. Точные доверительные интервалы для параметров в нормальной модели (4 вида). Распределения хи-квадрат и Стюдента, их свойства. Метод бутстрепа. Пример про дисперсию оценки коэффициента асимметрии. Описание метода бутстрепа, пример про оценку дисперсии статистики. Бутстрепные доверительные интервалы.

## 7. Ядерные оценки плотности

Ядерные оценки плотности, виды ядер, связь с эмпирической функцией распределения, сходимость оценки плотности. Оптимальная ширина ядра и оптимальное ядро, подбор оптимальной ширины ядра по выборке.

## 8. Проверка статистических гипотез

Проверка статистических гипотез: гипотезы и критерии, варианты ответов, связь с презумпцией невиновности. Ошибки I и II рода, уровень значимости критерия, мощность критерия, пример. Множественная проверка гипотез, постановка задачи, пример. Контроль FWER, методы Бонферрони и Холма. Нисходящие и восходящие процедуры.

## 9. Критерий Вальда

Критерий Вальда, его разновидности и особенности, функция мощности, пример.

## 10. Пример проверки гипотез в задаче исследовании эффективности нового лекарства

Пример проверки гипотез в задаче исследовании эффективности нового лекарства. p-value – определение в частном и общем случаях. Распределение p-value при справедливости основной гипотезы. Что не является p-value, пример.

## 11. Поведение критического уровня критерия для разных размеров выборки

Поведение критического уровня критерия для разных размеров выборки, следствия. Практическая значимость результата, примеры. Определение необходимого размера выборки на этапе планирования эксперимента.

## 12. Общие критерии согласия

Общие критерии согласия. Критерий Колмогорова, другие критерии, основанные на отличии от ЭФР. Графический способ проверки с помощью Q-Q plot. Критерии проверки нормальности: Жарка-Бера, Шапиро-Уилка. Критерий хи-квадрат, теорема Пирсона, пример.

## 13. Линейная регрессия и коэффициенты корреляции

Линейная регрессия. Постановка задачи линейной регрессии, примеры зависимостей, случай категориальных переменных. Метод наименьших квадратов. Оценка вектора параметров и ее свойства. Геометрический смысл метода. Несмещенная оценка дисперсии ошибки в методе наименьших квадратов. Коэффициенты корреляции Пирсона, Спирмена и Кендалла, их свойства. Таблицы сопряженности 2x2, точный тест Фишера, меры взаимосвязи, определение количества наблюдений. Общий случай таблиц сопряженности, типы вероятностных моделей, критерий хи-квадрат.

#### 14. Виды задач дисперсионного анализа, примеры

Виды задач дисперсионного анализа, примеры. 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.

#### 15. Виды альтернатив в непараметрическом случае

Виды альтернатив в непараметрическом случае. 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.

### 5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория, оснащенная мультимедиа проектором, экраном и микрофоном. Для проведения занятий в формате видеоконференции – ноутбук или персональный компьютер, оснащенный микрофоном и видеокамерой, имеющий выход в сеть Интернет с достаточной для участия в видеоконференции пропускной способностью.

### 6. Перечень рекомендуемой литературы

#### Основная литература

1. Введение в математическую статистику [Текст] : [учебник для вузов] / Г. И. Ивченко, Ю. И. Медведев. — М. : ЛКИ, 2010, 2014, 2015. — 600 с.
2. Наглядная математическая статистика [Текст] : учеб. пособие для вузов / М. Б. Лагутин. — 2-е изд., испр. — М. : Бином. Лаб. знаний, 2009. — 472 с.

#### Дополнительная литература

1. Прикладная математическая статистика [Текст] : для инженеров и научных работников / А. И. Кобзарь. — 2-е изд., испр. / [Научное изд.] — М. : Физматлит, 2012. — 816 с. — (Современные методы в математике). - Библиогр.: с. 737-759. - Предм. указ.: с. 806-810. - Имен. указ.: с. 811-813. - 500 экз. - ISBN 978-5-9221-1375-5 (в пер.) — Полный текст (Доступ из сети МФТИ / Удаленный доступ).

### 7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Не используются

### 8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

На занятиях могут использоваться мультимедийные технологии, включая демонстрацию презентации. Также семинары могут проходить в дистанционном режиме посредством видеоконференций.

## 9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен, с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике.

В результате изучения дисциплины студент должен знать основные определения и понятия, уметь применять полученные знания для решения различных задач.

Успешное освоение курса требует:

- посещения всех занятий, предусмотренных учебным планом по дисциплине;
- ведения конспекта занятий;
- напряжённой самостоятельной работы студента.

Самостоятельная работа включает в себя:

- чтение рекомендованной литературы;
- проработку учебного материала, подготовку ответов на вопросы, предназначенных для самостоятельного изучения;
- решение задач, предлагаемых студентам на занятиях;
- подготовку к выполнению заданий текущей и промежуточной аттестации.

Показателем владения материалом служит умение без конспекта отвечать на вопросы по темам дисциплины.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями преподавателю.

Возможен промежуточный контроль знаний студентов в виде решения задач в соответствии с тематикой занятий.

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)**

<b>по направлению:</b>	Прикладная математика и информатика
<b>профиль подготовки:</b>	Современная комбинаторика центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск" кафедра дискретной математики
<b>курс:</b>	<u>1</u>
<b>квалификация:</b>	магистр

Семестр, формы промежуточной аттестации: 2 (весенний) - Дифференцированный зачет

**Разработчики:**

А.М. Райгородский, д-р физ.-мат. наук, доцент, ассистент  
Н.А. Волков



## 1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-2 Способен управлять проектом на всех этапах его жизненного цикла	УК-2.1 Формулирует в рамках обозначенной проблемы цель, задачи, актуальность, значимость (научную, практическую, методическую и иную в зависимости от типа проекта), ожидаемые результаты и возможные сферы их применения
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
ПК-3 Владеет навыками участия в научных дискуссиях, выступления с сообщениями и докладами, устного, письменного и виртуального (размещение в информационных сетях) характера, представления материалов собственных исследований, проведения корректуры, редактирования, реферирования работ	ПК-3.2 Умеет вести корректную дискуссию в области информационно-коммуникационных технологий и информационных систем, задавать вопросы и отвечать на поставленные вопросы по теме научной работы

## 2. Показатели оценивания компетенций

В результате изучения дисциплины «Прикладная статистика» обучающийся должен:

### знать:

- фундаментальные понятия, законы теории вероятностей и математической статистики;
- понятия, аксиомы, методы доказательств и доказательства основных теорем курса;
- основные свойства объектов математической статистики;
- аналитические и численные подходы и методы для решения типовых прикладных задач теории вероятностей и математической статистики;
- параметрические и непараметрические методы статистического анализа;
- базовые понятия теории проверки статистических гипотез;
- основные подходы к сравнению оценок параметров как случайных величин;
- определение общей линейной регрессионной модели и метод наименьших квадратов;
- многомерное нормальное распределение и его основные свойства;
- методы оценки корреляционной зависимости;
- виды задач дисперсионного анализа.

### уметь:

- использовать свои знания для решения прикладных задач теории вероятностей и математической статистики;
- оценивать корректность постановок задач теории вероятностей и математической статистики;
- самостоятельно находить алгоритмы решения задач, в том числе нестандартных, и проводить их анализ;
- грамотно и корректно представлять решения задач теории вероятностей и математической статистики в устной и письменной форме;
- внедрять математико-статистические методы исследований при решении прикладных задач физики, астрономии, математики и педагогики;
- самостоятельно углублять и расширять знания в области математической статистики.

### владеть:

- культурой постановки, анализа и решения задач прикладной статистики, требующих для своего решения использования математических подходов и методов;
- предметным языком теории вероятностей и математической статистики, а также навыками грамотного описания решения задач и представления полученных результатов;
- основными методами математической статистики для построения точечных и доверительных оценок;
- навыками анализа статистических критериев;
- навыками применения теорем математической статистики в прикладных задачах физики, математики и педагогики.

### 3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

Текущая аттестация по дисциплине «Прикладная статистика» осуществляется в форме контрольных/самостоятельных работ в конце темы.

### 4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Какие основные подходы к статистике рассматриваются?
2. Что такое оценка параметра?
3. Можно ли оценивать сдвиг в модели распределения Коши с помощью выборочного среднего?
4. Как сгенерировать бутстрепную выборку?
5. Какова практическая интерпретация доверительного интервала?
6. Какие ответы возможны в задаче проверки гипотез?
7. Что такое критерий проверки гипотез?
8. Как проверить нормальность выборки?
9. Можно ли с помощью линейной регрессии приблизить параболическую зависимость?
10. Когда имеет смысл применять коэффициент корреляции Пирсона?
11. Что такое ранг наблюдения в выборке?
12. Как проверить независимости вещественного и категориального признака?
13. Как строится критерий Уилкоксона-Манна-Уитни?

#### Критерии оценивания

Оценка отлично 10 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины, проявляющему интерес к данной предметной области, продемонстрировавшему умение уверенно и творчески применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично 9 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично 8 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, правильное обоснование принятых решений, с некоторыми недочетами.

Оценка хорошо 7 баллов - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка хорошо 6 баллов - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности.

Оценка хорошо 5 баллов - выставляется студенту, если он в основном знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач достаточно большое количество неточностей.

Оценка удовлетворительно 4 балла - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он освоил основные разделы учебной программы, необходимые для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

Оценка удовлетворительно 3 балла - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, допускающему ошибки в формулировках базовых понятий, нарушения логической последовательности в изложении программного материала, слабо владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и с трудом применяет полученные знания даже в стандартной ситуации.

Оценка неудовлетворительно 2 балла - выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных принципов и не умеет использовать полученные знания при решении типовых задач.

Оценка неудовлетворительно 1 балл - выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

## **5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности**

Промежуточная аттестация по дисциплине «Прикладная статистика» осуществляется в форме дифференцированного зачета. Дифференциальный зачет проводится в устной форме, на подготовку отводится 30 минут, можно пользоваться программой дисциплины.