

**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>	Прикладные математика и физика
<b>профиль подготовки:</b>	Общая и прикладная физика Физтех-школа физики и исследований им. Ландау кафедра дискретной математики
<b>курс:</b>	1
<b>квалификация:</b>	магистр

Семестры, формы промежуточной аттестации:

1 (осенний) - Дифференцированный зачет  
2 (весенний) - Экзамен

Аудиторных часов: 120 всего, в том числе:

лекции: 60 час.  
семинары: 60 час.  
лабораторные занятия: 0 час.

Самостоятельная работа: 75 час.

Подготовка к экзамену: 30 час.

Всего часов: 225, всего зач. ед.: 5

Программу составили:

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

Программа обсуждена на заседании кафедры дискретной математики 04.06.2020

## 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 реализации	УК-2.3 Способен организовать и координировать работу участников проекта, обеспечивать работу команды необходимыми ресурсами
ОПК-1 Владеет системой фундаментальных научных знаний в области физико-математических наук	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты	ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценивать качество разработанной модели

### 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	Классическое (комбинаторное) определение вероятности	4	4		5
2	Геометрические вероятности и их свойства.	4	4		5
3	Условные вероятности, умножение вероятностей, формулы полной вероятности и Байеса.	4	4		5
4	Понятие о случайном блуждании и случайном графе.	6	6		5
5	Общая вероятностная модель. Аксиоматика Колмогорова.	4	4		5
6	Распределение функций от случайных величин.	4	4		5
7	Примеры комбинаторных задач, решаемых за счет линейности математического ожидания	4	4		
8	Виды задач дисперсионного анализа, примеры	2	2		3
9	Точечные оценки параметров, метод максимального правдоподобия	2	2		3
10	Сравнение оценок	2	2		3
11	Метод Ньютона и его применение для решения уравнения правдоподобия	2	2		3
12	Робастность, асимптотическая толерантность	2	2		3

13	Доверительные интервалы и доверительные области, метод бутстрепа	2	2		3
14	Ядерные оценки плотности	2	2		3
15	Проверка статистических гипотез	2	2		3
16	Критерий Вальда	2	2		3
17	Пример проверки гипотез в задаче исследовании эффективности нового лекарства	2	2		3
18	Поведение критического уровня критерия для разных размеров выборки	2	2		3
19	Общие критерии согласия	2	2		3
20	Линейная регрессия и коэффициенты корреляции	2	2		3
21	Виды задач дисперсионного анализа, примеры	2	2		3
22	Виды альтернатив в непараметрическом случае	2	2		3
Итого часов		60	60		75
Подготовка к экзамену		30 час.			
Общая трудоёмкость		225 час., 5 зач.ед.			

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

##### Семестр: 1 (Осенний)

##### 1. Классическое (комбинаторное) определение вероятности

Свойства вероятности при таком определении. Простейшие комбинаторные модели. Примеры комбинаторных задач, для решения которых удобно использовать классическое определение вероятности.

##### 2. Геометрические вероятности и их свойства.

Примеры задач, для решения которых удобно использовать геометрические вероятности: задача о встрече, задача о минимальном и максимальном элементах в случайной выборке и пр. Парадокс Бертрانا.

##### 3. Условные вероятности, умножение вероятностей, формулы полной вероятности и Байеса.

Независимость событий: попарная независимость, независимость в совокупности, независимость события от группы событий. Схема испытаний Бернулли. Полиномиальная схема. Схема серий.

##### 4. Понятие о случайном блуждании и случайном графе.

Порядковые статистики. Закон больших чисел для схемы Бернулли. Предельная теорема Пуассона для схемы серий. Локальная предельная теорема и интегральная предельная теорема Муавра – Лапласа.

##### 5. Общая вероятностная модель. Аксиоматика Колмогорова.

Случайные величины. Закон распределения, функция распределения и ее свойства. Дискретные и абсолютно непрерывные распределения, плотность распределения. Важнейшие распределения: биномиальное, пуассоновское, геометрическое, гипергеометрическое, равномерное, нормальное, Коши, экспоненциальное (показательное), гамма-распределение. Интерпретация предельных теорем Пуассона и Муавра – Лапласа в терминах распределений случайных величин.

#### 6. Распределение функций от случайных величин.

Математическое ожидание случайной величины. Линейность математического ожидания. Математическое ожидание функции от случайной величины.

#### 7. Примеры комбинаторных задач, решаемых за счет линейности математического ожидания

Неравенства Маркова и Чебышёва. Связь между понятием распределения случайной величины и заданием вероятностной меры на прямой

### Семестр: 2 (Весенний)

#### 8. 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

#### 9. 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.

#### 10. Сравнение оценок

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

#### 11. 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.

#### 12. 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.

### 13. 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 Стьюдента, их свойства. Метод бутстрепа. Пример про дисперсию оценки коэффициента асимметрии. Описание метода бутстрепа, пример про оценку дисперсии статистики. Бутстрепные доверительные интервалы.

### 14. 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.

### 15. 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.

### 16. Критерий Вальда

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

### 17. 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.

### 18. 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.

### 19. 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.

### 20. 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.

### 21. 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.

## 22. 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)**

**по направлению:** Прикладные математика и физика  
**профиль подготовки:** Общая и прикладная физика  
Физтех-школа физики и исследований им. Ландау  
кафедра дискретной математики  
**курс:** 1  
**квалификация:** магистр

Семестры, формы промежуточной аттестации:

1 (осенний) - Дифференцированный зачет  
2 (весенний) - Экзамен

**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 реализации	УК-2.3 Способен организовать и координировать работу участников проекта, обеспечивать работу команды необходимыми ресурсами
ОПК-1 Владеет системой фундаментальных научных знаний в области физико-математических наук	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты	ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценивать качество разработанной модели

## 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**

Перечень контрольных вопросов 1 семестра:

1. Классическое (комбинаторное) определение вероятности. Свойства вероятности при таком определении. Простейшие комбинаторные модели. Примеры комбинаторных задач, для решения которых удобно использовать классическое определение вероятности.
2. Геометрические вероятности и их свойства. Примеры задач, для решения которых удобно использовать геометрические вероятности: задача о встрече, задача о минимальном и максимальном элементах в случайной выборке и пр. Парадокс Бертрана.
3. Условные вероятности, умножение вероятностей, формулы полной вероятности и Байеса. Независимость событий: попарная независимость, независимость в совокупности, независимость события от группы событий.
4. Схема испытаний Бернулли. Полиномиальная схема. Схема серий.
5. Понятие о случайном блуждании и случайном графе. Порядковые статистики.
6. Закон больших чисел для схемы Бернулли.
7. Предельная теорема Пуассона для схемы серий. Локальная предельная теорема и интегральная предельная теорема Муавра – Лапласа.
8. Общая вероятностная модель. Аксиоматика Колмогорова.
9. Случайные величины. Закон распределения, функция распределения и ее свойства. Дискретные и абсолютно непрерывные распределения, плотность аспределения. Важнейшие распределения: биномиальное, пуассоновское, геометрическое, гипергеометрическое, равномерное, нормальное, Коши, экспоненциальное (показательное), гамма-распределение. Интерпретация предельных теорем Пуассона и Муавра – Лапласа в терминах распределений случайных величин.
10. Связь между понятием распределения случайной величины и заданием вероятностной меры на прямой.
11. Понятие о сингулярном распределении. Теорема Лебега о разложении произвольной функции распределения (б/д).
12. Распределение функций от случайных величин.
13. Математическое ожидание случайной величины. Линейность математического ожидания. Математическое ожидание функции от случайной величины.
14. Примеры комбинаторных задач, решаемых за счет линейности математического ожидания.
15. Моменты. Дисперсия. Вычисление моментов для распределений из п. 7.
16. Неравенства Маркова и Чебышёва.

Перечень контрольных вопросов 2-го семестра (экзамен)

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?

Примеры контрольных заданий 2 семестра (экзамен):

1. Пусть  $X_1, \dots, X_n$  --- выборка из распределения Лапласа со сдвигом  $\theta$ , то есть плотность имеет вид  $p_\theta(x) = \frac{1}{2} e^{-|x-\theta|}$ . Сравните в асимптотическом подходе выборочное среднее  $\overline{X}$ , усеченное среднее  $\overline{X}_\alpha$ , выборочную медиану  $\widehat{\mu}$ , медиану средних Уолша  $W$ .
2. Найдите асимптотическую толерантность медианы средних Уолша  $W$ .
3. Пусть  $X = (X_1, \dots, X_n)$  --- выборка из неизвестного распределения  $\text{Prob}$  с плотностью  $p(x)$  и  $\widetilde{p}_n(x)$  --- построенная по ней ядерная оценка плотности. Пусть также  $\xi$  --- случайная величина из распределения  $\text{Prob}$ , независимая с выборкой  $X$ . Рассмотрим ожидаемую среднеквадратичную ошибку  $\mathbb{E} \left( \widetilde{p}_n(\xi) - p(\xi) \right)^2$ . Какова асимптотика оптимальной ширины ядра, минимизирующей эту ошибку?
4. Пусть  $X_j = (X_{j1}, \dots, X_{jn_j})$ ,  $j \in \{1, 2\}$  --- две выборки. Для каждой выборки проверяются гипотезы  $H_j$  vs.  $H'_j$  с помощью критерия  $S_j$  уровня значимости  $\alpha$ . Предположим, гипотезы  $H_j$  верны. Какие значения может принимать FWER? Чему соответствуют пограничные значения, а также случай независимости выборок?
5. Пусть  $X_1, \dots, X_n$  --- выборка из распределения Бернулли с параметром  $\theta$ . Вычислите  $MSE_{\widehat{\theta}}(\theta)$ , где  $\widetilde{\theta} = \overline{X} + \frac{1}{1 + \sqrt{n}} \left( \frac{1}{2} - \overline{X} \right)$  --- оценка Ходжеса-Лемана.
6. Медицинская лаборатория проводит испытания нового препарата для лечения некоторого заболевания. Для исследований были отобраны 2500 больных. Некоторые из них принимали новый препарат, а другие --- плацебо. В первой группе значимое улучшение состояния наблюдается среди 853 пациентов из 1719 пациентов, принимавших новый препарат. Во второй группе значимое улучшение наблюдается среди 369 пациентов из 781 пациентов, принимавших плацебо. Влияет ли новый препарат на улучшение состояния у пациентов?
7. Рассмотрим критерий Уилкоксона-Манна-Уитни. Докажите, что при отсутствии совпадений среди  $X_i$  и  $Y_j$  для статистики Манна-Уитни справедливо  $U = V - \frac{m(m+1)}{2}$ . Найдите  $\mathbb{E} U$  при справедливости гипотезы  $H_0$  об однородности выборок.

Пример экзаменационного билета

1. Как сгенерировать бутстрепную выборку?
2. Пусть  $X = (X_1, \dots, X_n)$  --- выборка из неизвестного распределения  $\text{Prob}$  с плотностью  $p(x)$  и  $\widetilde{p}_n(x)$  --- построенная по ней ядерная оценка плотности. Пусть также  $\xi$  --- случайная величина из распределения  $\text{Prob}$ , независимая с выборкой  $X$ . Рассмотрим ожидаемую среднеквадратичную ошибку  $\mathbb{E} \left( \widetilde{p}_n(\xi) - p(\xi) \right)^2$ . Какова асимптотика оптимальной ширины ядра, минимизирующей эту ошибку?

#### 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 в 1-м семестре. Дифференциальный зачет проводится в устной форме по билетам.

Итоговая аттестация по дисциплине «Прикладная статистика» осуществляется в форме экзамена во 2-м семестре. Экзамен проводится в устной форме по билетам.

При проведении зачёта и экзамена обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося не должен превышать одного астрономического часа.