**Санкт-Петербургский государственный университет**

**Р А Б О Ч А Я П Р О Г Р А М М А**

**УЧЕБНОЙ ДИСЦИПЛИНЫ**

Математическая статистика

Mathematical Statistics

**Язык(и) обучения**

русский

Трудоемкость в зачетных единицах: 3

Регистрационный номер рабочей программы: 057535

Санкт-Петербург

2020

**Раздел 1. Характеристики учебных занятий**

**1.1. Цели и задачи учебных занятий**

Изложение основных принципов и результатов математической статистики в объеме, необходимом для общего развития и решения теоретических и прикладных задач соответствующей направленности. Усвоение студентами основных идей, понятий и фактов классической теории оценивания и проверки статистических гипотез, а также знакомство с основными идеями, лежащими в основе современной, ориентированной на интенсивное использование компьютера, статистической методологии.

**1.2. Требования подготовленности обучающегося к освоению содержания учебных занятий (пререквизиты)**

Курсы "Дискретная теория вероятностей" (семестр 2) и "Теория вероятностей" (семестр 4).

**1.3. Перечень результатов обучения (learning outcomes)**

Обучающийся должен овладеть теоретическим материалом в объеме, предусмотренном программой, уметь применять полученные знания при решении теоретических и прикладных задач, на основе владения освоенными разделами: основные числовые характеристики выборки, эмпирические распределения и порядковые статистики, статистические задачи в нормальных выборках, достаточные статистики и их применения, принципы и методы оценивания параметров, регулярные статистические эксперименты и информация Фишера, неравенство Рао-Крамера, асимптотические свойства оценок максимального правдоподобия, робастные оценки и их построение, проверка статистических гипотез по Нейману и Пирсону, критерий отношения правдоподобия и его свойства, критерий хи-квадрат и его разновидности, критерии Колмогорова-Смирнова и Крамера-фон Мизеса, регрессия и корреляция, метод наименьших квадратов и теорема Гаусса-Маркова; современные подходы статистической обработки данных большого объема и/или большой размерности. Обучающийся должен уяснить логику и технику построения математической теории как фундамента самостоятельных научных исследований и применения теории для практических нужд. Дисциплина участвует в формировании компетенций обучающихся по образовательной программе, установленных учебным планом для данной дисциплины.

**1.4. Перечень и объём активных и интерактивных форм учебных занятий**

Лекции 32 часа, практические занятия 30 часов, промежуточная аттестация (зачет, экзамен) 4 часа.

**Раздел 2. Организация, структура и содержание учебных занятий**

**2.1. Организация учебных занятий**

**2.1.1 Основной курс**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Трудоёмкость, объёмы учебной работы и наполняемость групп обучающихся | | | | | | | | | | | | | | | | | | |
| Код модуля в составе дисциплины,  практики и т.п. | Контактная работа обучающихся с преподавателем | | | | | | | | | | | | Самостоятельная работа | | | | Объём активных и интерактивных  форм учебных занятий | Трудоёмкость |
| лекции | семинары | консультации | практические  занятия | лабораторные работы | контрольные работы | коллоквиумы | текущий контроль | промежуточная  аттестация | итоговая аттестация | под руководством преподавателя | в присутствии  преподавателя | сам. раб. с использованием  методических материалов | текущий контроль (сам.раб.) | промежуточная аттестация (сам.раб.) | итоговая аттестация  (сам.раб.) |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | | | | | | | | | | | | | |
| Форма обучения: очная | | | | | | | | | | | | | | | | | | |
| Семестр 5 | 32 |  | 2 | 30 |  |  |  |  | 4 |  |  |  | 14 |  | 26 |  | 36 | 3 |
|  | 2-100 |  | 2-100 | 10-25 |  |  |  |  | 10-25 |  |  |  | 1-1 |  | 1-1 |  |  |  |
| ИТОГО | 32 |  | 2 | 30 |  |  |  |  | 4 |  |  |  | 14 |  | 26 |  |  | 3 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Виды, формы и сроки текущего контроля успеваемости и промежуточной аттестации | | | | | | |
| Код модуля в составе дисциплины, практики и т.п. | Формы текущего контроля успеваемости | | Виды промежуточной аттестации | | Виды итоговой аттестации  (только для программ итоговой аттестации и дополнительных образовательных программ) | |
| Формы | Сроки | Виды | Сроки | Виды | Сроки |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | |
| Форма обучения: очная | | | | | | |
| Семестр 5 |  |  | зачёт, устно, традиционная форма, экзамен, устно, традиционная форма | по графику промежуточной аттестации, по графику промежуточной аттестации |  |  |

**2.2. Структура и содержание учебных занятий**

Период обучения (модуль): **Семестр 7**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № п/п | Наименование темы (раздела, части) | Вид учебных занятий | Количество часов | |
| 1 | **Выборка и ее характеристики, эмпирические меры** | Лекции | 4 | |
| практические занятия |  | |
| в присутствии преподавателя |  | |
| по методическим материалам | 4 | |
| 2 | **Выборка из нормального закона** | Лекции | 4 | |
| практические занятия |  | |
| в присутствии преподавателя |  | |
| по методическим материалам | 6 | |
| 3 | **Достаточные статистики и их применения** | Лекции | 4 | |
| практические занятия |  | |
| в присутствии преподавателя |  | |
| по методическим материалам | 4 | |
| 4 | **Основные понятия теории оценивания и виды оценок** | Лекции | 4 | |
| практические занятия |  | |
| в присутствии преподавателя |  | |
| по методическим материалам | 6 | |
| 5 | **Основы проверки статистических гипотез по Нейману-Пирсону** | Лекции | 4 | |
| практические занятия |  | |
| в присутствии преподавателя |  | |
| по методическим материалам | 4 | |
| 6 | **Непараметрические критерии и их свойства** | Лекции | 4 | |
| практические занятия |  | |
| в присутствии преподавателя |  | |
| по методическим материалам | 6 | |
| 7 | **Множественная линейная регрессия и корреляция** | Лекции | 4 | |
| практические занятия |  |
| в присутствии преподавателя |  | |
| по методическим материалам | 6 | |
| 8 | **Современные методы математической статистики, анализ «больших данных»** | Лекции | 4 | |
| практические занятия |  | |
| в присутствии преподавателя |  | |
| по методическим материалам | 8 | |
| 9 | **Зачет, экзамен** | промежуточная аттестация (ауд) | 4 | |
| промежуточная аттестация (с.р.) | 28 | |
|  |  |  |  | |

**Раздел 1: Выборка и ее характеристики, эмпирические меры**

1. Примеры задач математической статистики.

2. Случайный эксперимент и выборка.

3. Выборочные моменты как оценки генеральных моментов.

4. Порядковые статистики и выборочные квантили.

5. Эмпирические функции распределения, теорема Гливенко-Кантелли и ее обобщения.

**Раздел 2: Выборка из нормального закона**

1. Важнейшие распределения, связанные с нормальным законом. Лемма Фишера.

2. Построение доверительных интервалов для параметров нормального закона.

3. Проверка гипотез о параметрах нормального закона.

4. Задача о равенстве двух средних. Проблема Беренса-Фишера.

**Раздел 3: Достаточные статистики**

1. Эквивалентные определения достаточных статистик. Примеры таких статистик.

2. Теорема факторизации Неймана-Фишера и ее применения.

3. Несмещенное оценивание с помощью достаточных статистик. Теоремы Рао-Блекуэлла-Колмогорова и Лемана-Шеффе.

4. Другие применения достаточных статистик. Теорема Базу.

**Раздел 4: Основные понятия теории оценивания и виды оценок**

1. Постановка задачи теории оценивания, функция потерь и риск.

2. Допустимые, минимаксные и байесовские оценки, связи между ними.

3. Регулярность статистических экспериментов. Информация Фишера и Кульбака-Лейблера.

4. Неравенство Рао-Крамера, R-эффективные оценки, экспонентные семейства.

5. Методы теории оценивания, метод моментов, M- и L-оценки.

6. Метод максимального правдоподобия и его асимптотические свойства.

7. Асимптотическая эффективность по Фишеру. Суперэффективность.

8. Понятие о робастном оценивании, схема загрязнения Хубера-Тьюки.

**Раздел 5: Основы проверки статистических гипотез по Нейману-Пирсону**

1. Постановка задачи проверки гипотез. Тесты, ошибки, критические области.

Уровень значимости, мощность.

2. Лемма Неймана-Пирсона.

3. Проверка гипотез для семейств с монотонным отношением правдоподобия.

4. Критерий отношения правдоподобия и его асимптотические свойства.

5. Критерий хи-квадрат для простых гипотез.

6. Критерий хи-квадрат для сложных гипотез, проверка независимости и однородности.

**Раздел 6:** **Непараметрические критерии и их свойства**

1. Эмпирический процесс и его свойства. Ранги, антиранги и знаки наблюдений.

2. Распределение вектора рангов при альтернативе, теорема Хёффдинга.

3. Локально наиболее мощные ранговые критерии, критерий Вилкоксона и нормальных меток.

4. Критерии Колмогорова-Смирнова и омега-квадрат.

5. Слабая сходимость эмпирического процесса к броуновскому мосту.

**Раздел 7:** **Множественная линейная регрессия и корреляция**

1. Корреляция и регрессия в статистике.

2. Коэффициенты корреляции Пирсона и Спирмена, их вычисление.

3. Модель множественной линейной регрессии.

4. Метод наименьших квадратов и оценка вектора неизвестных параметров.

5. Теорема Гаусса-Маркова как обоснование метода наименьших квадратов.

6. Оценивание параметров шумов регрессии. Нормальная регрессия.

7. Гетероскедастическая модель. Оценка Эйткена.

**Раздел 8. Современные методы математической статистики, анализ «больших данных»**

1. Бутстреп-метод: основные понятия, идеи. Разновидности бутстрепа.

2. Применение бутстрепа для построения точечных оценок и доверительных областей.

3.Метод «складного ножа», бутстреп «без возвращения» (подвыборки).

4. Бутстреп-аппроксимация распределения выборочного среднего. Корректность, корректность второго порядка.

5. Границы применимости бутстрепа. Контр-примеры.

6. Бутстрап для проверки гипотез.

7 .Бутстреп в модели регрессии. Кросс-валидация.

8. Применение бутстрепа при анализе «больших данных».

**3. Обеспечение учебных занятий**

**3.1. Методическое обеспечение**

**3.1.1 Методические указания по освоению дисциплины**

Посещение лекций (и практических занятий – если имеются в учебном плане).

**3.1.2 Методическое обеспечение самостоятельной работы**

Основная и дополнительная литература.

**3.1.3 Методика проведения текущего контроля успеваемости и промежуточной аттестации и критерии оценивания**

**Методика проведения зачета**  
Зачет проводится в устной форме. Для получения зачета необходимо решить 60% задач, предлагаемых в течение семестра. В случае, если к моменту проведения зачета студент решил меньшее количество задач, на зачете ему предлагаются задачи аналогичные по тематике и сложности. Задачи даются в форме домашних заданий с устной сдачей («листочки»), письменных домашних заданий и контрольных. Темы задач фиксированы, количество и форма выдачи остается на усмотрение преподавателя практических занятий. Возможна выдача задач повышенной сложности, решение которых засчитывается в качестве индивидуальных достижений студента (при подаче заявок на именные стипендии, конкурсы и т.п.); сдача таких заданий проводится в устной форме.

**Методика проведения экзамена**

Экзамен проводятся в устной форме. Билет состоит из двух вопросов. Время подготовки ответа на вопросы билета составляет 60 минут.

Использование конспектов и учебников, а также электронных устройств хранения, обработки или передачи информации при подготовке и ответе на вопросы экзамена категорически запрещено. После ответа на вопросы билета преподаватель задает несколько дополнительных вопросов, на основании оценки ответов на которые итоговая оценка по предмету может быть повышена или понижена.

Критерии выставления оценок:

Оценка «отлично» ставится за полностью раскрытый теоретический материал и правильные ответы на дополнительные вопросы преподавателя.

Оценка «хорошо» ставится за изложенный теоретический материал билета (возможно с помощью наводящих подсказок преподавателя) и правильные ответы на важнейшие дополнительные вопросы.

Оценка «удовлетворительно» ставится за знание основных вопросов по каждой теме.

Оценка «неудовлетворительно» выставляется, если не выполняются условия для получения оценок «отлично», «хорошо» и «удовлетворительно».

Соответствие оценки СПбГУ и оценки ECTS (Европейской системы переноса и накопления зачётных единиц):

|  |  |  |  |
| --- | --- | --- | --- |
| Итоговый процент выполнения, % | Оценка СПбГУ при  проведении зачёта | Оценка ECTS | Оценка СПбГУ при  проведении экзамена |
| 90-100 | зачтено | A | отлично |
| 80-89 | зачтено | B | хорошо |
| 70-79 | зачтено | C | хорошо |
| 60-69 | зачтено | D | удовлетворительно |
| 50-59 | зачтено | E | удовлетворительно |
| менее 50 | не зачтено | F | неудовлетворительно |

**3.1.4 Методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации (контрольно-измерительные материалы, оценочные средства)**

Период обучения (модуль): **Семестр 5**

**Список вопросов к экзамену:**

1) Предмет математической статистики и примеры задач.

2) Математическая постановка задач статистики: выборка, статистики, задача оценивания, проверка гипотез.

3) Эмпирическая функция распределения. Теорема Гливенко-Кантелли.

4) Выборочные моменты. Требования, предъявляемые к оценкам параметров.

5) Свойства выборочного среднего и выборочной дисперсии и как оценок генерального среднего и генеральной дисперсии.

6) Порядковые статистики и их распределение (точное и предельное).

7) Предельное распределение выборочных квантилей.

8) Важнейшие распределения, связанные с нормальным законом. Лемма Фишера.

9) Построение доверительных интервалов для параметров нормального закона.

10) Проверка гипотез о параметрах нормального закона.

11) Достаточные статистики: определение, теорема факторизации.

12) Несмещённое оценивание с помощью достаточных статистик (в том числе теорема Базу).

13) Постановка задачи теории оценивания. Допустимые, минимаксные, байесовские оценки и простейшие соотношения между ними.

14) Минимаксность выборочного среднего как оценки нормального среднего.

15) Построение минимаксной оценки для распределения Бернулли.

16) Регулярные статистические эксперименты. Информация Фишера и её свойства.

17) Неравенство Рао-Крамера. R-эффективные оценки и их примеры. Экспонентные семейства.

18) Метод моментов. Примеры его применения. Другие методы оценивания.

19) Метод максимального правдоподобия: определение и примеры применения.

20) Асимптотические свойства оценок максимального правдоподобия.

21) Асимптотическое построение доверительных интервалов.

22) Асимптотическая эффективность оценок по Фишеру. Пример Ходжеса. Суперэффективность. Понятие робастности.

23) Проверка гипотез: постановка задачи и основные определения.

24) Лемма Неймана–Пирсона и ее геометрическая интерпретация.

25) Критерий отношения правдоподобия для сложных гипотез. Критерий Стьюдента как критерий отношения правдоподобия. Теорема Уилкса.

26) Критерий хи-квадрат для простых гипотез (с выводом предельного распределения).

27) Критерий хи-квадрат для проверки сложных гипотез о параметрическом семействе и как критерий независимости признаков.

28) Критерии, основанные на эмпирической функции распределения.

29) Критерий Смирнова для проверки однородности.

30) Статистические свойства рангов. Критерий Вилкоксона.

31) Метод наименьших квадратов для оценки коэффициентов линейной регрессии: построение оценки. Оценка Эйткена.

32) Теорема Гаусса–Маркова.

33) Несмещённое оценивание дисперсии шумов.

34). Нормальная регрессия.

35) Бутстреп-метод, его разновидности, области применения.

36) Бутстреп для среднего. Корректность.

37) Свойство корректности второго порядка бутстреп-аппроксимации распределения выборочного среднего.

38) Бутстреп для построения доверительных областей и проверки гипотез.

39) Бутстреп в модели регрессии, кросс-валидация.

40) Применение бутстрепа для анализа «больших данных».

**3.1.5 Методические материалы для оценки обучающимися содержания и качества учебного процесса**

Анкета для студентов для оценки качества преподавания курса.

Просим Вас заполнить анкету-отзыв по прочитанной дисциплине. Обобщенные данные анкет будут использованы для ее совершенствования. По каждому вопросу проставьте соответствующие оценки по шкале от 1 до 10 баллов (обведите выбранный Вами балл). В

случае необходимости впишите свои комментарии.

1. Насколько Вы удовлетворены содержанием дисциплины в

целом?

1 2 3 4 5 6 7 8 9 10

Комментарий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

2. Насколько Вы удовлетворены общим стилем преподавания?

1 2 3 4 5 6 7 8 9 10

Комментарий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3. Как Вы оцениваете качество подготовки предложенных

методических материалов?

1 2 3 4 5 6 7 8 9 10

Комментарий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

4. Насколько Вы удовлетворены использованием

преподавателями активных методов обучения?

1 2 3 4 5 6 7 8 9 10

Комментарий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

5. Какой из модулей (разделов) дисциплины Вы считаете наиболее полезным, ценным с точки зрения дальнейшего обучения и/или

применения в последующей практической деятельности?

Комментарий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

6. Что бы Вы предложили изменить в методическом и

содержательном плане для совершенствования преподавания данной

дисциплины?

Комментарий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

СПАСИБО!

**3.2. Кадровое обеспечение**

**3.2.1 Образование и (или) квалификация штатных преподавателей и иных лиц, допущенных к проведению учебных занятий**

К чтению лекций должны привлекаться преподаватели, имеющие ученую степень доктора или кандидата наук (в том числе степень PhD, прошедшую установленную процедуру признания и установления эквивалентности) и/или ученое звание профессора или доцента.

**3.2.2 Обеспечение учебно-вспомогательным и (или) иным персоналом**

Не требуется.

**3.3. Материально-техническое обеспечение**

**3.3.1 Характеристики аудиторий (помещений, мест) для проведения занятий**

Учебные аудитории для проведения учебных занятий, оснащенные стандартным оборудованием, используемым для обучения в СПбГУ в соответствии с требованиями материально-технического обеспечения.

**3.3.2 Характеристики аудиторного оборудования, в том числе неспециализированного компьютерного оборудования и программного обеспечения общего пользования**

Стандартное оборудование, используемое для обучения в СПбГУ. MS Windows, MS Office, Mozilla FireFox, Google Chrome, Acrobat Reader DC, WinZip, Антивирус Касперского.

**3.3.3 Характеристики специализированного оборудования**

Не требуется.

**3.3.4 Характеристики специализированного программного обеспечения**

Не требуется.

**3.3.5 Перечень и объёмы требуемых расходных материалов**

Мел — не менее 2 кусков на час лекционных занятий, фломастеры для доски, губка.

**3.4. Информационное обеспечение**

**3.4.1 Список обязательной литературы**

1. А.А. Боровков. Математическая статистика. — М.: Наука, 2007.
2. Г.И. Ивченко, Ю.И.Медведев. Введение в математическую статистику. — М., Изд-во ЛКИ, 2010.
3. Г.И. Ивченко, Ю.И. Медведев, А.В. Чистяков. Сборник задач по математической статистике. — М., Высшая школа, 1989.
4. Л.Н. Большев, Н.В. Смирнов. Таблицы математической статистики. — М.: Наука, 1983.
5. B.Efron, R.J. Tibshirani. An introduction to the bootstrap. — Chapman & Hall, New York, 1993.
6. M. R. Chernick. Bootstrap methods: a guide for practitioners and researchers. — 2nd ed. — John Wiley & Sons, New York 2008.

**3.4.2 Список дополнительной литературы**

1. Г. Крамер. Математические методы статистики. — М.:Мир, 1975.
2. П. Бикел, К. Доксум. Математическая статистика. — М.: Финансы и статистика, 1983.
3. Э. Леман. Теория точечного оценивания. — М., Мир, 1991.
4. Б. Эфрон. Нетрадиционные методы многомерного статистического анализа. — М: Финансы и статистика, 1988.
5. B.Efron. Bootstrap methods: another look at the jackknife, Ann. Stat. 7 (1979), 1–26.
6. B.Efron, T. Hastie. Computer Age Statistical Inference. — Cambridge University Press, 2016.
7. P.J. Bickel, D.A. Freedman. Some asymptotic theory for the bootstrap, Ann. Statist. 9 (1981), no. 6, 1196–1217.
8. P.J. Bickel, F. Goetze, W.R. van Zwet. Resampling fewer than n observations: gains, losses, and remedies for losses. Statistica Sinica 7 (1997), 1–31.
9. M.R. Chernick, R.A. LaBudde. An introduction to bootstrap methods with applications to R. — Wiley, New York, 2011.
10. P. Hall. The bootstrap and Edgeworth expansion. — Springer, New York, 1992.
11. E. Mammen. When does bootstrap work? Asymptotic results and simulations. — Springer, New York, 1992.

**3.4.3 Перечень иных информационных источников**

1. Сайт Научной библиотеки им. М. Горького СПбГУ: http://www.librarv.spbu.ru/

2. Электронный каталог Научной библиотеки им. М. Горького СПбГУ: http://www.librarv.spbu.ru/cgibin/irbis64r/cgiirbis 64.ехе?С21 COM=F&I21 DBN=IBIS&P21 DBN=IBIS

3. Перечень электронных ресурсов, находящихся в доступе СПбГУ: http://cufts.librarv.spbu.ru/CRDB/SPBGU/

4. Перечень ЭБС, на платформах которых представлены российские учебники, находящиеся в доступе СПбГУ: http://cufts.library.spbu.ru/CRDB/SPBGU/browse?name=rures&resource tvpe=8

**Раздел 4. Разработчики программы**

Никитин Яков Юрьевич, доктор физико-математических наук, профессор, заведующий кафедрой теории вероятностей и математической статистики СПбГУ, y.nikitin@spbu.ru