

Правительство Российской Федерации

**Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«Национальный исследовательский университет
«Высшая школа экономики»**

**Факультет Компьютерных наук
Департамент больших данных и информационного поиска**

УТВЕРЖДАЮ
Академический руководитель
образовательной программы
по направлению 01.03.02
«Прикладная математика и информатика»
А.С. Конушин

_____ 2017 г.
«__» _____

Программа дисциплины

Дополнительные главы математической статистики

Для направления 01.03.02 «Прикладная математика и информатика» подготовки бакалавров

Автор программы:

д.ф.-м.н. Шабанов Д.А.

Одобрена на заседании Департамента больших данных и информационного поиска

«__» _____ 2017г.

Руководитель департамента

_____ В.В. Подольский

Рекомендована Академическим советом
образовательной программы
«Прикладная математика и информатика»

«__» _____ 2017 г.

Менеджер Департамента больших данных и информационного поиска

Москва, 2017

*Настоящая программа не может быть использована другими подразделениями
университета и другими вузами без разрешения кафедры-разработчика программы*



1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины «Дополнительные главы математической статистики» (3-ый год обучения) устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов направления подготовки 01.03.02 «Прикладная математика и информатика», обучающихся по бакалаврской программе «Прикладная математика и информатика» изучающих дисциплину «Дополнительные главы математической статистики».

Программа разработана в соответствии с:

- образовательным стандартом федерального государственного автономного образовательного учреждения высшего профессионального образования НИУ ВШЭ подготовки бакалавров по направлению 010400.62 «Прикладная математика и информатика»;
- Образовательной программой подготовки бакалавров по направлению 01.03.02 «Прикладная математика и информатика»,
- Рабочим учебным планом университета подготовки бакалавров по направлению 01.03.02 «Прикладная математика и информатика», утвержденным в 2015г.

2 Цели освоения дисциплины

Цель освоения дисциплины «Дополнительные главы математической статистики» — познакомить слушателей с понятиями, фактами и методами математической статистики, не вошедшими в базовый курс «Теория вероятностей и математическая статистика», а также с различными возможными приложениями для статистической обработки реальных данных.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- Знать основные понятия математической статистики, их основные результаты и математические методы анализа.
- Уметь применять математические методы и модели к анализу случайных явлений для их адекватного описания и понимания.
- Владеть навыками решения стандартных задач математической статистики, а также применением основных аналитических инструментов для анализа вероятностных и статистических задач.

В результате освоения дисциплины студент осваивает следующие компетенции:

Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
Общенаучная	ОНК-1	Способность к анализу и синтезу на основе системного подхода	Стандартные (лекционно-семинарские)
Общенаучная	ОНК-2	Способность перейти от проблемной ситуации к проблемам, задачам и лежащим в их основе противоречиям	Стандартные (лекционно-семинарские)



Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
Общенаучная	ОНК-3	Способность использовать методы критического анализа, развития научных теорий, опровержения и фальсификации, оценить качество исследований в некоторой предметной области	Стандартные (лекционно-семинарские)
Общенаучная	ОНК-4	Готовность использовать основные законы естественнонаучных дисциплин в профессиональной деятельности, применять методы математического анализа и моделирования, теоретического и экспериментального исследования при работе в какой-либо предметной области	Стандартные (лекционно-семинарские)
Общенаучная	ОНК-5	Готовность выявить естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, привлечь их для решения соответствующий аппарат дисциплины	Стандартные (лекционно-семинарские)
Общенаучная	ОНК-6	Способность приобретать новые знания с использованием научной методологии и современных образовательных и информационных технологий	Стандартные (лекционно-семинарские)
Общенаучная	ОНК-7	Способность порождать новые идеи (креативность)	Стандартные (лекционно-семинарские)
Профессиональные	ПК-1	Способность демонстрации общенаучных базовых знаний естественных наук, математики и информатики, понимание основных фактов, концепций, принципов теорий, связанных с прикладной математикой и информатикой	Стандартные (лекционно-семинарские)
Профессиональные	ПК-2	способность понимать и применять в исследовательской и прикладной деятельности современный математический аппарат	Стандартные (лекционно-семинарские)
Профессиональные	ПК-3	способность в составе научно-исследовательского и производственного коллектива решать задачи профессиональной деятельности в соответствии с профилем подготовки, общаться с экспертами в других предметных областях	Стандартные (лекционно-семинарские)
Профессиональные	ПК-4	способность критически оцени-	Стандартные (лекционно-



Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
		вать собственную квалификацию и её востребованность, переосмысливать накопленный практический опыт, изменять при необходимости вид и характер своей профессиональной деятельности	но-семинарские)
Профессиональные	ПК-8	способность решать задачи производственной и технологической деятельности на профессиональном уровне, включая разработку математических моделей, алгоритмических и программных решений	Стандартные (лекционно-семинарские)

4 Место дисциплины в структуре образовательной программы

«Дополнительные главы теории вероятностей» является самостоятельной учебной дисциплиной, относится к математическому и естественнонаучному циклу дисциплин. Для специализации 01.03.02 «Прикладная математика и информатика» настоящая дисциплина является факультативной.

Для освоения учебной дисциплины студенты должны владеть знаниями и навыками в объеме программы средней школы по математике и освоить учебные курсы:

- Дискретная математика,
- Математический анализ-1,-2,-3,-4
- Линейная алгебра и геометрия,
- Теория вероятностей и математическая статистика.

Основные положения дисциплины могут быть использованы в дальнейшем при изучении следующих дисциплин:

- Машинное обучение;
- Эконометрика;
- Статистика случайных процессов.

5 Тематический план учебной дисциплины

№	Название раздела	Всего часов	Аудиторные часы			Самостоятельная работа
			Лекции	Семинары	Практические занятия	
1	Основные понятия математической статистики	4	2			2
2	Основные методы построения оценок	8	4			4
3	Сравнение оценок	12	6			6
4	Метод максимального правдоподобия	8	4			4



5	Достаточные статистики и оптимальные оценки	12	6			6
6	Линейная регрессионная модель	8	4			4
7	Проверка статистических гипотез	20	10			10
8	Ранговые методы	8	4			4
	Итого:	80	40			40

6 Формы контроля знаний студентов

Тип контроля	Форма контроля	2 год				Кафедра	Параметры
		1	2	3	4		
Текущий							
	Домашнее задание			1	1		Всего 30-40 задач
Промежуточный	Экзамен				1		Устный экзамен

6.1 Критерии оценки знаний, навыков

В рамках курса студенты должны выполнить две домашние письменные работы и сдать экзамен (4 модуль). Каждое задание и экзамен оцениваются по 10-балльной шкале. Домашние задания заключаются в решении задач. Проверяется ответ и ход решения. На экзамене также необходимо изложить теоретический лекционный материал.

6.2 Порядок формирования оценок по дисциплине

Промежуточный контроль

Накопленная оценка (3-й и 4-й модули) за текущий контроль формируется из оценок за домашние задания:

$$O_{\text{накопленная}} = 0,5 * O_{\text{д/з1}} + 0,5 * O_{\text{д/з2}}.$$

Итоговая оценка рассчитывается следующим образом:

$$O_{\text{итог}} = 0,8 * O_{\text{накопленная}} + 0,2 * O_{\text{экз}}.$$

Все оценки выставляются по 10-балльной системе, способ округления всегда арифметический.

7 Содержание дисциплины

Раздел 1. Основные понятия математической статистики.

Вероятностно-статистическая модель, понятия наблюдения и выборки. Эмпирическое распределение и эмпирическая функция распределения. Обоснованность основной задачи математической статистики и теорема Гливленко-Кантелли. Вероятностно-статистическая модель. Параметрическая статистическая модель. Моделирование выборки из неизвестного распределения, принадлежащему параметрическому семейству.

Раздел 2. Основные методы построения оценок.



Статистики и оценки. Общая идея построения хороших статистик, примеры: выборочные усреднения, порядковые статистики, выборочные квантили, $SM\$\text{-}$ оценки. Основные свойства оценок: несмещенность, состоятельность, сильная состоятельность, асимптотическая нормальность. Примеры. Наследование асимптотических свойств при взятии непрерывной функции. Методы нахождения оценок, общий принцип подстановки. Метод моментов, состоятельность оценки метода моментов. Выборочные квантили и выборочная медиана. Теорема об асимптотической нормальности выборочной квантили. Примеры.

Раздел 3. Сравнение оценок.

Сравнение оценок, функция потерь и функция риска. Подходы к сравнению оценок: равномерный, байесовский, минимаксный, асимптотический. Допустимые оценки. Доминируемое семейство распределений, условия его регулярности. Неравенство Рао-Крамера (одномерный и многомерный варианты) и эффективные оценки. Критерий эффективности оценки. Теорема о наилучшем квадратичном прогнозе. Байесовская оценка, априорные и апостериорные плотности. Теорема о байесовской оценке, ее оптимальность в байесовском подходе к сравнению оценок.

Раздел 4. Метод максимального правдоподобия.

Метод максимального правдоподобия. Примеры. Экстремальное свойство функции правдоподобия. Теорема о существовании состоятельного решения уравнения правдоподобия. Состоятельность оценки максимального правдоподобия. Теорема об асимптотической нормальности состоятельного решения уравнения правдоподобия. Теорема Бахадура (б/д). Асимптотически эффективные оценки. Асимптотическая эффективность и эффективность оценки максимального правдоподобия.

Раздел 5. Достаточные статистики и оптимальные оценки.

Достаточные статистики. Теорема Колмогорова-Блэкуэлла-Рао об улучшении несмещенных оценок и ее многомерное следствие. Полные статистики. Теорема об оптимальной оценке. Критерий факторизации Неймана-Фишера, доказательство для дискретного случая. Примеры, нахождение оптимальной оценки параметра $a > 0$ в случае выборки из $U(0, a)$. Формула пересчета условных математических ожиданий по двум мерам. Критерий факторизации Неймана-Фишера в общем случае. Экспоненциальное семейство распределений. Теорема о полной достаточной статистике в экспоненциальном семействе. Пример: нахождение оптимальных оценок параметров по нормальной выборке.

Раздел 6. Линейная регрессионная модель.

Линейная регрессионная модель. Оценка наименьших квадратов, ее основные свойства. Лемма об оптимальности в классе несмещенных линейных оценок. Несмещенная оценка для параметра дисперсии в линейной регрессионной модели. Линейная гауссовская модель. Полная достаточная статистика в линейной гауссовской модели. Оптимальные оценки параметров в линейной гауссовской модели. Теорема об ортогональных разложениях гауссовского случайного вектора. Распределения оптимальных оценок в линейной гауссовской модели.

Раздел 7. Проверка статистических гипотез.

Проверка статистических гипотез, общие принципы и основные понятия: гипотеза, критерий, уровень значимости, альтернативы, ошибки первого и второго родов, функция мощности. Несмещенность и состоятельность статистического критерия. Сравнение критериев, равномерно наиболее мощные критерии. Лемма Неймана-Пирсона для проверки простых гипотез. Построение с ее помощью наиболее мощных критериев. Семейства с монотонным отношением правдоподобия. Монотонность функции мощности для таких семейств. Теорема о монотонном отношении правдоподобия. Примеры: проверка гипотез с односторонними альтернативами для бернуллиевских и нормальных выборок. Двойственность доверительного оценивания и проверки



гипотез. Линейные гипотезы в линейной гауссовской модели, F-критерий для проверки линейной гипотезы в гауссовской линейной модели. Пример: проверка однородности двух нормальных выборок. Обобщенный метод наименьших квадратов. Критерии согласия в дискретном случае. Статистика хи-квадрат Пирсона в полиномиальной схеме Бернулли с m исходами. Теорема Пирсона о сходимости статистики хи-квадрат. Параметрический критерий хи-квадрат. Критерий согласия хи-квадрат, его состоятельность. Критерии согласия в непрерывном случае. Критерий Колмогорова для проверки простой гипотезы о виде функции распределения, его состоятельность.

Раздел 8. Ранговые методы.

Ранги. Ранговый метод для проверки однородности двух независимых выборок. Теорема об асимптотической нормальности статистики Уилкоксона.

8 Образовательные технологии

В учебном процессе используются следующие образовательные технологии:

№ п/п	Вид занятия	Форма проведения занятий	Цель
1	Лекция	Изложение теоретического материала.	Получение теоретических знаний по предмету.
2	Самостоятельная работа студента	Решение задач.	Усовершенствование практических навыков.

9 Оценочные средства для текущего контроля и аттестации студента

9.1 Вопросы для оценки качества освоения дисциплины

Раздел 1.

1. Что такое эмпирическая функция распределения?
2. Сформулируйте теорему Гливленко-Кантелли.

Раздел 2.

1. В чем состоит свойство состоятельности оценки?
2. Что такое выборочные квантили?
3. Каковы основные свойства оценки метода моментов?

Раздел 3.

1. Каковы основные подходы к сравнению оценок?
2. Сформулируйте неравенство Рао-Крамера.
3. Что такое байесовская оценка?

Раздел 4.

1. Дайте определение оценки по методу максимального правдоподобия.
2. При каких условиях оценка по методу максимального правдоподобия состоятельна?
3. Что такое асимптотическая эффективность?

Раздел 5.



1. Что такое оптимальная оценка?
2. Сформулируйте критерий факторизации Неймана-Фишера.
3. Приведите пример экспоненциального семейства распределений.

Раздел 6.

1. Дайте определение линейной регрессионной модели.
2. Что такое метод наименьших квадратов?
3. Каковы оптимальные оценки параметров гауссовской линейной регрессии.

Раздел 7.

1. Что такое уровень значимости статистического критерия?
2. Приведите пример семейства распределений с монотонным отношением правдоподобия.
3. Каковы асимптотические свойства критерия хи-квадрат?
4. Сформулируйте теорему Колмогорова о предельном распределении разности настоящей и эмпирической функций распределений.

Раздел 8.

1. Что такое статистика Уилкоксона?
2. В чем состоит ранговый подход к проверке гипотезы однородности двух выборок?

10 Учебно-методическое и информационное обеспечение дисциплины

10.1 Базовые учебники

1. А.А. Боровков. Математическая статистика. - 4-е изд. – СПб: Лань, 2010.
2. Г.И. Ивченко, Ю.И. Медведев. Математическая статистика. – М: Книжный дом "Либликом", 2014.

10.2 Основная литература

1. Севастьянов Б. А. Курс теории вероятностей и математической статистики. - 2-е изд. - М.-Ижевск: Институт компьютерных исследований, 2004.
2. Ширяев А.Н. Вероятность. В 2-х кн., 6-е изд., М.: МЦНМО, 2016.
3. Э. Леман. Теория точечного оценивания. – М: Наука, 1991.
4. Ю.Н. Тюрин. Математическая статистика. Записки лекций - электронная версия.

10.3 Дополнительная литература

1. Бикел П., Доксам К. Математическая статистика. Вып. 1 и 2. – М: Финансы и статистика, 1983.