

Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования
Московский государственный университет имени М.В. Ломоносова
Механико-математический факультет
Кафедра газовой и волновой динамики



УТВЕРЖДАЮ
Заведующий кафедрой
/Нигматулин Р.И./
« 10 » июня 2019г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Наименование дисциплины (модуля):

Методы многомерного статистического анализа

наименование дисциплины (модуля)

Уровень высшего образования:

специалитет

Направление подготовки (специальность):

01.05.01 Фундаментальные математика и механика

(код и название направления/специальности)

Направленность (профиль) ОПОП:

Фундаментальная механика

(если дисциплина (модуль) относится к вариативной части программы)

Форма обучения:

очная

очная, очно-заочная

Рабочая программа рассмотрена и одобрена
на заседании кафедры газовой и волновой динамики
(протокол № 15, « 10 » июня 20 19 года)

Москва 2019

На обратной стороне титула:

Рабочая программа дисциплины (модуля) разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки / специальности «Фундаментальные математика и механика» (программы бакалавриата, магистратуры, реализуемых последовательно по схеме интегрированной подготовки; программы специалитета; программы магистратуры) в редакции приказа МГУ от 30 декабря 2016 г.

Год (годы) приема на обучение _____ 2014,2015,2016_____

1. Место дисциплины (модуля) в структуре ОПОП ВО, курс относится к вариативной части ОПОП.
2. Входные требования для освоения дисциплины (модуля), предварительные условия (если есть): _____:
знание курса теории вероятности
3. Результаты обучения по дисциплине (модулю), соотнесенные с требуемыми компетенциями выпускников.

Компетенции выпускников (коды)	Планируемые результаты обучения по дисциплине, соответствующие компетенциям
ОПК-1	<i>Знать:</i> основные статистические методы обработки и анализа экспериментальных данных <i>Владеть:</i> современной терминологией математической статистики и классификацией статистических процедур в соответствии со сложностью практических задач
ОПК-3	<i>Уметь:</i> строить статистическую модель изучаемого процесса
<i>ОПК-4</i>	<i>Уметь :</i> применять современные статистические пакеты и соответствующие языки программирования.
<i>ПК-1</i>	<i>Уметь:</i> самостоятельно найти форму представления данных и выбрать необходимый статистический инструмент.
<i>ПК-4</i>	<i>Уметь</i> представить полученные в курсовой работе результаты и проанализировать правильность исходной модели

4. Формат обучения: параллельное сопровождение лекционного курса_практической работой над конкретными задачами с применением современных статистических пакетов.
5. Объем дисциплины (модуля) составляет 2 з.е., в том числе 36 академических часов, отведенных на контактную работу обучающихся с преподавателем, 36 академических часов на самостоятельную работу обучающихся.
6. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий

Наименование и краткое содержание разделов и тем дисциплины (модуля), Форма промежуточной аттестации по дисциплине (модулю)	Всего (часы)	В том числе			
		Контактная работа (работа во взаимодействии с преподавателем) Виды контактной работы, часы			Самостоятельная работа обучающегося, часы ()
		Занятия лекционного типа*	Занятия семинарского типа*	Всего	
Тема _1. Статистическая модель и статистические оценки 1.1.Основные измерительные шкалы и соответствующие им выборочные характеристики . 1.2.Нормальные выборки. 1.3. Точечные статистические оценки. 1.4. Метод максимального правдоподобия 1.5.Доверительный интервал.	8	4	4		
Тема _ . 2. Основные распределения, математической статистики 2.1. Дискретные распределения (биномиальное, полиномиальное, Пуассона) 2.2. Распределения, связанные с нормальным (χ^2 -, Стьюдента, F-распределения) 2.3. Двумерное нормальное распределение 2.4. Распределение Вейбулла, показательное распределение	4	2	2		
Тема...3. Непараметрические методы математической статистики 3.1. Критерии однородности. 3.2. Критерии случайности.	12	6 4	6		

3.3. Критерии согласия. 3.4. Ранговая корреляция 3.5. Нормальная корреляция 3.6. Однофакторный дисперсионный анализ					
Тема 4. Робастное статистическое оценивание 4.1. Грубые ошибки и методы их выявления в статистической совокупности данных 4.2. Методы исчисления устойчивых статистических оценок	4	2	2		<i>Домашнее задание, пример: доказать, что собственные функции операторов физических величин взаимно ортогональны</i> 1 час
Тема 5. Основы корреляционного анализа данных 5.1. Матрица данных и корреляционная матрица 5.2. Свойства корреляционной матрицы 5.2.1. Статистическая модель порождения данных	8	4	4		<i>Домашнее задание, пример найти наиболее вероятное и среднее расстояние от ядра для электрона в невозбужденном атоме водорода</i> 1 час
Тема 6. Многомерный регрессионный анализ 6.1. Введение в множественный корреляционно-регрессионный анализ. 6.2. Простая линейная регрессия. 6.3. Обобщение МНК не случай матричного представления линейной регрессии. 6.4. Структурные уравнения линейной регрессии. 6.5. Нелинейные регрессионные модели.	12	6	6		

Контрольная работа в аудитории		0	2		
Тема 7. Факторный анализ. 7.1. Основная идея методов факторного анализа и их классификация. 7.2. Теорема Герстоуна. 7.3. Общий алгоритм факторного анализа. 7.4. Метод главных компонент. 7.5. Разложение дисперсии в факторном анализе. 7.6. Метод главных факторов. 7.7. Вращение пространства общих факторов. 7.8. Статистическая оценка надежности решений методами главных компонент и факторного анализа.	8	4	4		<i>Домашнее задание, пример: какой химический элемент имеет полностью заполненную К-оболочку и наполовину заполненную 2s оболочку?</i> 1 час
Тема 8. Кластерный анализ. 8.1. Сущность методов кластерного анализа. 8.2. Меры сходства. 8.3. Иерархический кластерный анализ. 8.4. Метод k-средних. 8.5. Метод поиска сгущений. 8.6. Критерии качества классификации.	8	4	4		
Тема 9. Дискриминантный анализ. 9.1. Основные положения дискриминантного анализа. 9.2. Дискриминантные функции и их геометрическая интерпретация. 9.3. Вычисление коэффициентов дискриминантной функции. 9.4. Классификация при наличии двух обучающих выборок. 10. Канонические корреляции	8	4	4		<i>Рефераты по темам, например:</i> 1. Проверка гипотез о тренде параметров распределения 2. Оценивание параметров распределения Вейбулла.

<i>Промежуточная отчетность - экзамен</i>					
<i>Итого 72 часа</i>					

**Внимание! В таблице должно быть зафиксировано проведение текущего контроля успеваемости, который может быть реализован, например, в рамках занятий семинарского типа.*

*** Часы, отводимые на проведение промежуточной аттестации, выделяются из часов самостоятельной работы обучающегося*

7. Фонд оценочных средств (ФОС) для оценивания результатов обучения по дисциплине (модулю)

7.1. Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости.

Примеры:

1. При построении каких статистик применяют эмпирическую функцию распределения?
2. При проверке каких гипотез используют критерий числа серий?
3. С помощью каких критериев можно проверить однородность двух выборок, измеренных в порядковой, количественной шкалах?
4. Привести пример распределения, у которого медиана, мода, совпадают с математическим ожиданием
5. Какие статистические задачи решаются с помощью рангового коэффициента корреляции Спирмена?

7.2. Типовые контрольные задания или иные материалы для проведения промежуточной аттестации.

Рефераты по темам, например:

1. *Проверка гипотез о тренде параметров распределения*
2. *Оценивание параметров распределения Вейбулла:*

Примеры экзаменационных заданий:

1. . Найти засорения и построить устойчивые статистические оценки по Пуанкаре, Винзору и Хуберу математического ожидания разности соответствующих значений третьего и четвертого столбцов таблицы 1.
2. Найти оценки корреляционной матрицы и множественных коэффициентов корреляции и детерминации для данных, приведенных в таблице 2.
3. Найти уравнение множественной регрессии переменной D на L, P, T и X1 таблицы 2. Проверить значимость модели. Каковы оценки дисперсии ошибок наблюдений, коэффициенты детерминации и множественной корреляции. Определить доверительные интервалы для параметров регрессии и проверить гипотезу о значимости параметров.

Определить остатки построить график остатков. Проверить справедливость предположений регрессионного анализа: постоянства дисперсии остатков, их некоррелированность и нормальность распределения остатков.

Сделать вывод об адекватности регрессионной модели результатам наблюдений

- 4 Разбить группу из девяти объектов (текстов) на три кластера. Каждый текст характеризуется пятью признаками. Применить метод k-средних, использовать расстояние между классами, измеряемое по принципу ближайшего соседа, по центрам тяжести групп

Неполный перечень вопросов к экзамену

1. Матрица данных и корреляционная матрица.

2. Свойства корреляционной матрицы.
3. Статистическая модель порождения данных и аппроксимационный подход.
4. Грубые ошибки и методы их выявления в статистической совокупности данных.
5. Методы исчисления устойчивых статистических оценок.
6. Введение в множественный корреляционно-регрессионный анализ.
7. Простая линейная регрессия.
8. Обобщение МНК не случай матричного представления линейной регрессии.
9. Структурные уравнения линейной регрессии.
10. Нелинейные регрессионные модели.
11. Основная идея методов факторного анализа и их классификация.
12. Теорема Терстоуна.
13. Общий алгоритм факторного анализа.
14. Метод главных компонент.
15. Разложение дисперсии в факторном анализе.
16. Метод главных факторов.
17. Вращение пространства общих факторов.
18. Статистическая оценка надежности решений методами главных компонент и факторного анализа.
19. Сущность методов кластерного анализа.
20. Меры сходства.
21. Иерархический кластерный анализ.
22. Метод k-средних.
23. Метод поиска сгущений.
24. Критерии качества классификации.

ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ результатов обучения (РО) по дисциплине (модулю)				
Оценка	2	3	4	5
РО и соответствующие виды оценочных средств				
Знания: экзамен (виды оценочных)	Отсутствие знаний об основных понятиях	Фрагментарные знания об основных понятиях	Общие, но не структурированные знания	Сформированные систематические знания

<i>средств: устные и письменные опросы и контрольные работы, тесты, и т.п.)</i>	математической статистики.	математической статистики, Например. Понятия статистического решения	основ математической статистики и простейших методов многомерного статистического анализа	основ математической статистики и простейших методов многомерного статистического анализа
Умения: домашние задания, контрольные работы, зачет <i>: практические контрольные задания, написание и защита рефератов на заданную тему</i>	Отсутствие умений построения статистической модели предлагаемой задачи	В целом успешное, но не систематическое умение построения статистической модели предлагаемой задачи	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности не принципиального характера) построения статистической модели предлагаемой задачи	Успешное и систематическое умение построения статистической модели предлагаемой задачи
Навыки: экзамен, курсовая работа (владения, опыт деятельности) <i>(виды оценочных средств: выполнение и защита курсовой работы, отчет по практике, отчет по НИР и т.п.)</i>	Отсутствие навыка делать выводы об имеющихся данных на основе построенной статистической модели	Наличие отдельных навыков делать выводы об имеющихся данных на основе построенной статистической модели	Не активные навыки делать выводы об имеющихся данных на основе построенной статистической модели	Сформированные навыки делать выводы об имеющихся данных на основе построенной статистической модели

8. Ресурсное обеспечение:

- Перечень основной и дополнительной литературы

Список литературы.

1. Э.М.Браверман , И.Б.Мучник, «Структурные методы обработки эмпирических данных», Москва, изд-во «Наука», 1983г.
2. Дубов А.М , Мхитарян В.С., Трошин Л.И., «Многомерные статистические методы», Москва, «Финансы и статистика», 2000г.
3. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере

- Перечень лицензионного программного обеспечения (при необходимости)

- Перечень профессиональных баз данных и информационных справочных систем
- Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости)
- Описание материально-технического обеспечения.

9. Язык преподавания.

Русский

10. Преподаватель (преподаватели).

Е.А.Ильюшина

11. Автор (авторы) программы.

Е.А.Ильюшина.