

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Кубанский государственный университет»
Факультет компьютерных технологий и прикладной математики

УТВЕРЖДАЮ:

Проректор по учебной работе,
качеству образования и первый
проректор

Хагуров Т.А.

подпись
« 27 »



2018г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.18 МНОГОМЕРНЫЙ АНАЛИЗ ДАННЫХ

Направление подготовки 01.03.02 Прикладная математика и информатика

Направленность (профиль) Математическое моделирование и вычислительная математика: Математическое моделирование

Программа подготовки академическая

Форма обучения очная

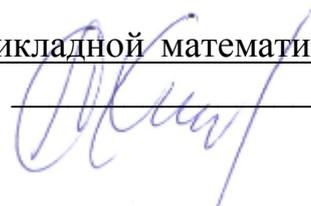
Квалификация (степень) выпускника бакалавр

Краснодар 2018

Рабочая программа дисциплины «МНОГОМЕРНЫЙ АНАЛИЗ ДАННЫХ» составлена в соответствии с Федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению **01.03.02 Прикладная математика и информатика**, утвержденным приказом Министерства образования и науки Российской Федерации № 228 от 12 марта 2015 г.

Программу составил:

Халафян А.А., д-р техн. наук, профессор кафедры прикладной математики КубГУ



Рабочая программа дисциплины «Многомерный анализ данных» утверждена на заседании кафедры математического моделирования протокол № 11 «16» апреля 2018 г.

Заведующий кафедрой математического моделирования акад. РАН, д-р физ.-мат. наук, проф. Бабешко В.А.



Утверждена на заседании учебно-методической комиссии факультета компьютерных технологий и прикладной математики протокол № 1 «20» апреля 2018 г.

Председатель УМК факультета
канд. физ.-мат. наук, доцент Малыхин К.В.



Рецензенты:

Шапошникова Т.Л., д-р пед. наук, канд. физ.-мат. наук, директор института фундаментальных наук (ИФН) ФГБОУ ВО «КубГТУ»

Марков В.Н., д-р техн. наук, профессор кафедры информационных систем и программирования института компьютерных систем и информационной безопасности ФГБОУ ВО «КубГТУ»

1 Цели и задачи изучения дисциплины

1.1 Цель освоения дисциплины

Цели изучения дисциплины определены государственным образовательным стандартом высшего образования и, соотнесены с общими целями ООП ВО по направлению подготовки «**Прикладная математика и информатика**», в рамках которой преподается дисциплина.

Целью освоения учебной дисциплины «Многомерный анализ данных» является изучение основных методов многомерного статистического анализа данных с точки зрения их практического применения; привить навыки работы с соответствующими разделами ППП STATISTICA.

1.2 Задачи дисциплины:

- помочь студентам понять и, освоить методологию многомерного статистического анализа данных;
- привить теоретические и практические знания в области прикладного многомерного анализа данных;
- познакомить студентов и обучить максимально широкому инструментарию многомерного анализа данных в среде ППП STATISTICA;
- выработать в процессе обучения у студентов навыки грамотного использования аппарата вероятно-статистического моделирования посредством применения передовых информационных технологий.

1.3. Место учебной дисциплины в структуре ООП ВО

Дисциплина тесно связана с дисциплинами общенаучного цикла: теория вероятностей и математическая статистика, математический анализ, алгебра и геометрия, дискретная математика, информатика и др.

Материал курса предназначен для использования в дисциплинах, связанных с количественным анализом экономических явлений, таких как, например, математическое моделирование экологических, экономических и технологических процессов. Результаты изучения курса также могут быть использованы при прохождении производственной практики и выполнении выпускной квалификационной работы.

1.4 Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Компетенции обучающегося, формируемые в результате освоения курса «**Многомерный анализ данных**»:

ПК-5	Способностью осуществлять целенаправленный поиск информации о новейших научных и технологических достижениях в информационно-телекоммуникационной сети "Интернет" (далее - сеть "Интернет") и в других источниках		
	Знать	Уметь:	Владеть:
ПК-5	– основы многомерных статистических методов, прикладных и информационных процессов в экономике; – методы и приемы поиска и работы с	– работать с информацией из различных источников, включая сетевые ресурсы сети Интернет, для решения профессиональных и социальных задач.	– навыками поиска информацией из различных источников, включая сетевые ресурсы сети Интернет, для решения профессиональных задач,

информацией из различных источников, включая сетевые ресурсы сети Интернет, для решения профессиональных задач	– формулировать запросы и классифицировать информацию по вероятностно-статистическим моделям данных	– способностью применять системный подход и математические методы к анализу данных, полученных из различных источников
--	---	--

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 3 зачетных единицы, 108 академических часа (из них 72 аудиторных). Курс «Математический анализ данных» лабораторных занятий, сопровождаемых регулярной индивидуальной работой преподавателя со студентами в процессе самостоятельной работы. В конце семестра проводится зачет.

Вид учебной работы		Всего часов Семестр 7
Контактная работа, в том числе:		76,2
Аудиторные занятия (всего)		72
В том числе:		
Занятия лекционного типа		–
Лабораторные занятия		72
Иная контактная работа:		
Контроль самостоятельной работы		4
Промежуточная аттестация (ИКР)		0,2
Самостоятельная работа		
Самостоятельная работа (всего)		31,8
В том числе:		
Проработка учебного (теоретического) материала		31,8
Подготовка к текущему контролю		–
Контроль: зачет		зачет
Общая трудоёмкость	час.	108
	в том числе контактная работа	76,2
	зач. ед	3

2.2 Структура дисциплины:

Распределение видов учебной работы и их трудоёмкости по разделам дисциплины. Разделы дисциплины, изучаемые в 7 семестре

№ n/n	Наименование раздела, темы	Количество часов		
		Всего	Аудиторная работа	Внеаудиторная работа
			ЛР	СРС
1.	Элементарные понятия статистики. Измерительные шкалы	8	6	2
2.	Работа с данными. Основные операции над случаями и переменными	8	6	2

№ n/n	Наименование раздела, темы	Количество часов		
		Всего	Аудиторная работа	Внеаудиторная работа
			ЛР	СРС
3.	Основные статистики, корреляционный анализ	6	4	2
4.	Сравнение средних величин критерием Стьюдента	6	4	2
5.	Непараметрическая статистика. Сравнение средних величин методами непараметрической статистики	6	4	2
6.	Группировка и однофакторная ANOVA	6	4	2
7.	Дисперсионный анализ	6	4	2
8.	Таблицы частот, сопряженности, флагов и заголовков	6	4	2
9.	Канонический анализ	8	6	2
10.	Линейное и нелинейное моделирование взаимосвязей	6	4	2
11.	Дискриминантный анализ	8	4	4
12.	Кластерный анализ	6	4	2
13.	Деревья классификации	6	4	2
14.	Факторный анализ	8	6	2
15.	Многомерное шкалирование	5,8	4	1,8
16.	Обзор пройденного материала и прием зачета	4	4	–
Контроль самостоятельной работы (КСР)		4	–	–
Промежуточная аттестация (ИКР)		0,2	–	–
Итого:		108	72	31,8

Примечание: ЛР – лабораторные занятия, СРС – самостоятельная работа студента

2.3 Содержание разделов дисциплины:

№	Наименование раздела	Форма текущего контроля (по неделям семестра)
1	2	4
1	Основные приемы работы с ППП STATISTICA	1. Устный опрос в конце ЛР 2. Проверка выполнения лабораторного задания
2	Основные статистики и методы сравнения средних величин	1. Устный опрос в конце ЛР 2. Проверка выполнения лабораторного задания
3	Методы многомерного анализа	1. Устный опрос в конце ЛР 2. Проверка выполнения лабораторного задания

2.3.1 Занятия лекционного типа

Учебный план не предусматривает занятий лекционного типа по дисциплине «Многомерный анализ данных».

2.3.2 Занятия семинарского типа

Учебный план не предусматривает занятий семинарского типа по дисциплине «Многомерный анализ данных».

2.3.3 Лабораторные занятия

№	Наименование раздела	Содержание раздела	Форма текущего контроля (по неделям семестра)
1	2	3	4
1	Основные приемы работы с ППП STATISTICA	Тема 1. Элементарные понятия статистики. Измерительные шкалы. Тема 2. Работа с данными. Основные операции над случаями и переменными	1. Проверка выполнения лабораторных работ 2. Проверка выполнения СР
2	Основные статистики и методы сравнения средних величин	Тема 3. Основные статистики, корреляционный анализ Тема 4. Сравнение средних величин критерием Стьюдента Тема 5. Непараметрическая статистика. Сравнение средних величин методами непараметрической статистики Тема 6. Группировка и однофакторная ANOVA Тема 7. Дисперсионный анализ	1. Проверка выполнения лабораторных работ 2. Проверка выполнения СР
3	Методы многомерного анализа	Тема 8. Таблицы частот, сопряженности, флагов и заголовков. Тема 9. Канонический анализ Тема 10. Линейное и нелинейное моделирование взаимосвязей Тема 11. Дискриминантный анализ Тема 12. Кластерный анализ Тема 13. Деревья классификации Тема 14. Факторный анализ. Тема 15. Многомерное шкалирование Тема 16. Обзор пройденного материала и прием зачета	1. Проверка выполнения лабораторных работ 2. Проверка выполнения СР

2.3.4. Курсовые работы – не предусмотрены

2.3.5 Перечень учебно-методического обеспечения для самостоятельной работы обучающегося по дисциплине

Целью самостоятельной работы студента является углубление знаний, полученных в результате аудиторных занятий. Вырабатываются навыки самостоятельной работы. Закрепляются опыт и знания, полученные во время лабораторных занятий.

№	Наименование раздела	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1.	Основные приемы работы с ППП STATISTICA	1. Халафян, А.А. STATISTICA 6. Статистический анализ данных. Второе издание. М: Бином , 2010. 2. Халафян, А.А. STATISTICA 6. Математическая статистика с элементами теории вероятностей. М: Бином , 2011. 3. Туганбаев, А.А. Теория вероятностей и математическая статистика / А.А. Туганбаев, В.Г. Крупин. СПб: Лань, 2011. 320 с. [Электронный ресурс]. – Режим доступа: https://e.lanbook.com/book/652 .
2.	Методы сравнения средних	1. Халафян, А.А. STATISTICA 6. Статистический анализ данных. Второе издание. М: Бином , 2010. 2. Халафян, А.А. STATISTICA 6. Математическая статистика с элементами теории вероятностей. М: Бином , 2011. 3. Туганбаев, А.А. Теория вероятностей и математическая статистика / А.А. Туганбаев, В.Г. Крупин. СПб: Лань, 2011. 320 с. [Электронный ресурс]. – Режим доступа: https://e.lanbook.com/book/652 .
3.	Методы многомерного анализа	1. Халафян, А.А. STATISTICA 6. Статистический анализ данных. Второе издание. М: Бином , 2010. 2. Халафян, А.А. STATISTICA 6. Математическая статистика с элементами теории вероятностей. М: Бином , 2011. 3. Туганбаев, А.А. Теория вероятностей и математическая статистика / А.А. Туганбаев, В.Г. Крупин. СПб: Лань, 2011. 320 с. [Электронный ресурс]. – Режим доступа: https://e.lanbook.com/book/652 .

3. Образовательные технологии

С точки зрения применяемых методов используются как традиционные информационно-объяснительные лекции, так и интерактивная подача материала с мультимедийной системой. Компьютерные технологии в данном случае обеспечивают возможность разнопланового отображения алгоритмов и демонстрационного материала. Такое сочетание позволяет оптимально использовать отведенное время и раскрывать логику и содержание дисциплины.

Лекции представляют собой систематические обзоры многомерного анализа с

подачей материала в виде презентаций.

Лабораторное занятие позволяет научить студента применять теоретические знания при решении и исследовании конкретных задач. Лабораторные занятия проводятся в компьютерных классах, при этом практикуется работа в группах. Подход разбора конкретных ситуаций широко используется как преподавателем, так и студентами при проведении анализа результатов самостоятельной работы. Это обусловлено тем, что в процессе исследования часто встречаются задачи, для которых единых подходов не существует. Каждая конкретная задача при своем исследовании имеет множество подходов, а это требует разбора и оценки целой совокупности конкретных ситуаций.

4. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины

4.1 Фонд оценочных средств, для проведения текущей аттестации

Текущий контроль знаний студентов представляет собой:

- выполнение домашних заданий;
- выполнение самостоятельной работы

Примерные задания на лабораторные работы

Перечень компетенций (части компетенции), проверяемых оценочным средством: ПК–5.

Задание 1. Произвести основные операции над переменными и случаями по представленным файлам данных

Задание 2. Построить двухмерные графики по представленным файлам данных.

Задание 3. Вычислить описательные статистики, просчитать параметрические и непараметрические коэффициенты корреляция по представленным файлам данных.

Задание 4. Произвести подгонку законов распределения по представленным файлам данных, генерацию случайных чисел. Решить задачи.

Задание 5. Произвести сравнение средних посредством t-критерия по представленным файлам данных.

Задание 6. Реализовать метод Группировка и однофакторная ANOVA на представленных файлах данных.

Задание 7. Реализовать метод Дисперсионный анализ на представленных файлах данных.

Задание 8. Реализовать метод Непараметрическая статистика на представленных файлах данных.

Задание 9. Реализовать метод частотный анализ на представленных файлах данных.

Задание 10. Реализовать метод Таблицы кросстабуляции и таблицы флагов и заголовков на представленных файлах данных.

Задание 11. Реализовать методы Множественная регрессия, фиксированная нелинейная регрессия на представленных файлах данных.

Задание 12. Реализовать методы Нелинейная регрессия, модели бинарных откликов на представленных файлах данных.

Задание 13. Реализовать метод Дискриминантами анализ на представленных файлах данных.

Задание 14. Реализовать метод Кластерный анализ на представленных файлах данных.

Задание 15. Реализовать метод Деревья классификации на представленных файлах данных.

Задание 16. Реализовать метод Факторный анализ на представленных файлах данных.

Задание 17. Реализовать метод Канонический анализ на представленных файлах данных.

Задание 18. Реализовать метод Многомерное шкалирование на представленных файлах данных.

4.2 Фонд оценочных средств, для проведения промежуточной аттестации

Перечень компетенций (части компетенции), проверяемых оценочным средством:
ПК–5. Промежуточный контроль осуществляется в конце семестра в форме зачета.

Примерный перечень вопросов к зачету

1. Инструменты для работы с данными
2. Структура электронной таблицы
3. Основные операции над переменными и случаями
4. Основные операции с таблицами данных
5. Обмен данными с другими приложениями WINDOWS
6. Двухмерная графика
7. Трехмерная графика
8. Формирование отчета и рабочей книги
9. Описательные статистики
10. Корреляционная матрица
11. t-критерий сравнения средних
12. Группировка и однофакторная ANOVA
13. Дисперсионный анализ
14. Таблицы частот
15. Таблицы кросстабуляции
16. Таблицы флагов и заголовков
17. Подбор закона распределения
18. Линейная регрессионная модель
19. Модуль множественная регрессия
20. Линеаризующие преобразования
21. Модели бинарных откликов
22. Описание модуля Нелинейное оценивание
23. Экспоненциальная регрессия
24. Кусочно-линейная регрессия
25. Определенная пользователем регрессия
26. Дискриминантный анализ. Описание модуля Дискриминантный анализ
27. Кластерный анализ. Описание модуля Кластерный анализ.
28. Задача факторного анализа. Описание модуля Факторный анализ
29. Деревья классификации
30. Задача канонического анализа. Описание метода. Модуль Канонический анализ
31. Многомерное шкалирование
32. Анализ соответствий

Критерии для получения зачета.

Зачет выставляется, если:

- достаточно полные и систематизированные знания по дисциплине;
- умение ориентироваться в основных теориях, концепциях и направлениях дисциплины и давать им критическую оценку;
- использование научной терминологии, лингвистически и логически правильное изложение ответа на вопросы, умение делать обоснованные выводы;

- владение инструментарием по дисциплине, умение его использовать в постановке и решении научных и профессиональных задач;
- усвоение основной и дополнительной литературы, рекомендованной учебной программой по дисциплине;
- самостоятельная работа на практических занятиях, участие в групповых обсуждениях, высокий уровень культуры исполнения заданий;
- средний уровень сформированности заявленных в рабочей программе компетенций.

Зачет не выставляется, если:

- фрагментарные знания по дисциплине;
- отказ от ответа (выполнения письменной работы);
- знание отдельных источников, рекомендованных учебной программой по дисциплине;
- неумение использовать научную терминологию;
- наличие грубых ошибок;
- низкий уровень культуры исполнения заданий;
- низкий уровень сформированности заявленных в рабочей программе компетенций.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

- при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;
- при проведении процедуры оценивания результатов обучения инвалидов, и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;
- при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов, и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся студентов.

5. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

5.1 Основная литература:

1. Халафян, А.А. STATISTICA 6. Статистический анализ данных. Второе издание. М: Бином, 2010. 522 с.

2. Халафян, А.А. STATISTICA 6. Математическая статистика с элементами теории вероятностей. М: Бином, 2011. 491 с.

3. Туганбаев, А.А. Теория вероятностей и математическая статистика / А.А. Туганбаев, В.Г. Крупин. СПб: Лань, 2011. 320 с. [Электронный ресурс]. – Режим доступа: <https://e.lanbook.com/book/652>.

5.2 Дополнительная литература:

1. Боровиков В.П. Популярное введение в современный анализ данных в системе STATISTICA: методология и технология современного анализа данных. М.: Горячая линия-Телеком, 2013. 288 с. + 1 электрон. опт. диск (CD-ROM).

2. Миркин Б.Г. Введение в анализ данных: учебник и практикум. М.: Юрайт, 2018. 174 с. [Электронный ресурс]. – Режим доступа: www.biblio-online.ru/book/46A41F93-BC46-401C-A30E-27C0FB60B9DE.

6. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

www.statlab.kubsu.ru

<http://www.statsoft.ru/home/textbook/default.htm>

<http://statsoft.ru/solutions/>

<http://statistica.ru/>

<https://e.lanbook.com/>

7. Методические рекомендации по организации изучения дисциплины

По курсу предусмотрено проведение лекционных занятий, на которых дается основной теоретический материал, лабораторных занятий, позволяющих студентам в полной мере ознакомиться с понятиями дисциплины и освоиться в решении практических задач. Важнейшим этапом курса является самостоятельная работа. Целью самостоятельной работы бакалавра является углубление знаний, полученных в результате аудиторных занятий. Вырабатываются навыки самостоятельной работы. Закрепляются опыт и знания, полученные во время лабораторных занятий. Самостоятельная работа студентов в ходе изучения дисциплины состоит в выполнении индивидуальных заданий (отчет в электронной форме), подготовки теоретического материала к лабораторным занятиям, на основе конспектов лекций и учебной литературы, согласно календарному плану и подготовки теоретического материала к тестовому опросу, зачету и экзамену, согласно вопросам к экзамену.

Указания по оформлению работ:

- работа на лабораторных занятиях выполняется за компьютером в компьютерном классе;

- оформление индивидуальных заданий (отчетов) желательно в виде файлов в формате word.

Итогом самостоятельной работы студента является отчет, в котором на оригинальной таблице исходных данных студент самостоятельно проводит анализ данных всеми изученными в рамках курса методами и, представляет его на проверку в электронном виде.

Проверка индивидуальных заданий по темам, разобранным на лабораторных занятиях, осуществляется через неделю на текущем лабораторном занятии, либо в течение недели после этого занятия на консультации.

Для разъяснения непонятных вопросов лектором и ассистентом еженедельно проводятся консультации, о времени которых группы извещаются заранее.

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю) (при необходимости)

8.1 Образовательные технологии

1. Проверка домашних заданий и консультирование посредством электронной почты.
2. Использование мультимедийных презентаций при чтении лекций.

8.2 Перечень необходимого программного обеспечения

1. Операционная система MS Windows.
2. Интегрированное офисное приложение MS Office.
3. Программное обеспечение для организации управляемого коллективного и безопасного доступа в Интернет.

8.3 Перечень необходимых информационных справочных систем

1. Электронная библиотечная система eLIBRARY.RU (<http://www.elibrary.ru/>)
3. Справочно-правовая система «Консультант Плюс» (<http://www.consultant.ru>)
4. Электронная библиотечная система «Университетская библиотека ONLINE» (<http://www.biblioclub.ru>)

9 Материально-техническое обеспечение учебной дисциплины

В качестве материально-технического обеспечения дисциплины используются - проекционное оборудование (цифровой проектор, экран, ноутбук, интерактивная доска). Для проведения занятий используются аудитории с учебной мебелью (столы, стулья), соответствующей количеству студентов и позволяющей осуществлять упражнения по моделированию компьютерные классы.

№	Вид работ	Материально-техническое обеспечение дисциплины и оснащённость
1.	Лабораторные занятия	Компьютерный класс, укомплектованный компьютерами с лицензионным программным обеспечением, необходимой мебелью (доска, столы, стулья). (аудитории: 101, 102, 106, 106а, 105/1, 107(2), 107(3), 107(5), А301).
2.	Групповые (индивидуальные) консультации	Аудитория для семинарских занятий, групповых и индивидуальных консультаций, укомплектованные необходимой мебелью (доска, столы, стулья). (аудитории: 129, 131).
3.	Текущий контроль, промежуточная аттестация	Аудитория для семинарских занятий, текущего контроля и промежуточной аттестации, укомплектованная необходимой мебелью (доска, столы, стулья) (аудитории: 129, 131, 133, А305, А307, 147, 148, 149, 150, 100С, А301б, А512), компьютерами с лицензионным программным обеспечением и выходом в интернет (106, 106а, А301)

№	Вид работ	Материально-техническое обеспечение дисциплины и оснащенность
4.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения, обеспеченный доступом в электронную информационно-образовательную среду университета, необходимой мебелью (столы, стулья). (Аудитория 102а, читальный зал).

Компьютерная поддержка учебного процесса обеспечивается по всем дисциплинам. Факультет компьютерных технологий и прикладной математики, оснащен компьютерными классами, установлена локальная сеть, все компьютеры факультета подключены к сети Интернет. Студентам доступны современные ПЭВМ, современное лицензионное программное обеспечение.

Студенты и преподаватели вуза имеют постоянный доступ к электронному каталогу учебной, методической, научной литературе, периодическим изданиям и архиву статей.