

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Кубанский государственный университет»
Факультет математики и компьютерных наук

УТВЕРЖДАЮ
Проректор по учебной работе
качеству образования – первый
проректор

Хагуров Т.А.

29 мая 2020 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.02

МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ

Направление подготовки

02.04.01 Математика и компьютерные науки

Программа магистратуры

«Математическое и компьютерное моделирование»

Форма обучения

очная

Квалификация

магистр

Краснодар 2020

Рабочая программа дисциплины «Многомерный статистический анализ» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.04.01 Математика и компьютерные науки (уровень высшего образования: магистратура)

Программу составила:
к.ф.-м.н., доцент кафедры МКМ

Качанова И. А. _____

Рабочая программа дисциплины утверждена на заседании кафедры математических и компьютерных методов,
протокол №11 от 21.04.2020.

Заведующий кафедрой математических и компьютерных методов

Лежнев А. В. _____

Утверждена на заседании учебно-методической комиссии
факультета математики и компьютерных наук,
протокол № 2 от 30.04.2020.

Председатель УМК
факультета математики и компьютерных наук

Шмалько С. П. _____

Рецензенты:

Савенко И. В., коммерческий директор ООО «Росглавино»

Никитин Ю.Г., доцент кафедры теоретической физики и компьютерных технологий
ФГБОУ ВО «Кубанский государственный университет»

1 Цели и задачи изучения дисциплины

1.1 Цель дисциплины состоит в формировании у студентов представления о сущности многомерного статистического анализа и его роли в вероятностно-статистическом моделировании; познании методологических основ и практическое овладение приемами многомерного статистического анализа.

1.2 Задачи дисциплины:

В результате освоения курса «Многомерный статистический анализ» студенты должны

- понять необходимость и область применения многомерных статистических методов;
- научиться организовывать статистическое наблюдение и обрабатывать статистические данные с использованием современных компьютерных технологий;
- обучиться многомерным статистическим методам, позволяющим среди множества возможных вероятностно-статистических моделей обоснованно выбрать ту, которая наилучшим образом соответствует исходным статистическим данным, характеризующим реальное поведение исследуемой совокупности объектов, оценить надежность и точность выводов, сделанных на основе ограниченного статистического материала;
- усвоить основные правила расчета обобщающих статистических показателей;
- уметь формулировать выводы, необходимые для проведения научных исследований и осуществления практической деятельности.

1.3 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина Б1.В.02 «Многомерный статистический анализ» относится к части, формируемой участниками образовательных отношений Блока 1. Дисциплины (модули).

Успешное усвоение студентом данного предмета предполагает у него наличие базовых знаний в области математической статистики, теории вероятностей, владение навыками работы в пакете MS Excel. Знания и умения, практические навыки, приобретенные студентами в результате изучения дисциплины, могут использоваться при выполнении дипломной работы, связанной с построением вероятностно-статистических моделей для социально-экономических и других процессов.

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

Изучение учебной дисциплины «Многомерный статистический анализ» направлено на формирование у обучающихся следующих профессиональных компетенций

№ п.п.	Индекс компетенции	Содержание компетенции (или её части)	В результате изучения учебной дисциплины обучающиеся должны		
			знать	уметь	владеть
1.	ПК-1	Способен демонстрировать фундаментальные знания математических и естественных наук, программирования и информационных технологий	- интерфейс пакетов прикладных программ для работы со статистическими данными	-обрабатывать эмпирические и экспериментальные данные; -осуществлять выбор инструментальных средств для обработки статистических данных в соответствии с поставленной задачей; -содержательно интерпретировать результаты расчетов.	-навыками работы в среде пакетов прикладных программ для работы с многомерными статистическим и данными
2.	ПК-2	Способность проводить научные исследования, на основе существующих методов в конкретной области профессиональной деятельности	-основные источники получения официальных статистических данных; -основные методы обработки и анализа первичных статистических данных	- открывать новые признаки неизвестного - анализировать явления - исследовать известное и неизвестное - систематизировать, формулировать проблему исследования - определять программу практических действий - предусматривать ход событий и	-методами обработки и анализа статистических данных в соответствии с поставленными задачами

№ п.п.	Индекс компетенции	Содержание компетенции (или её части)	В результате изучения учебной дисциплины обучающиеся должны		
			знать	уметь	владеть
				последствия тех или иных этапов	

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 4 зач.ед. (144 часов), их распределение по видам работ представлено в таблице

Вид учебной работы	Всего часов	Семестры (часы)
		1-й
Контактная работа, в том числе:	64,3	64,3
Аудиторные занятия (всего)	64	64
Занятия лекционного типа	32	32
Занятия семинарского типа (семинары, практические занятия)		
Лабораторные занятия	32	32
Иная контактная работа:		
Контроль самостоятельной работы (КСР)		
Промежуточная аттестация (ИКР)	0,3	0,3
Самостоятельная работа, в том числе:		
Проработка учебного (теоретического) материала	53	53
Подготовка к текущему контролю	26,7	26,7
Общая трудоемкость	час.	144
	в том числе контактная работа	54,3
	зач. ед	4

2.2 Структура дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины. Разделы дисциплины, изучаемые в семестре 1 (*очная форма*)

№	Наименование разделов	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1	Множественный регрессионный анализ		6		6	10
2	Кластерный анализ		6		6	10
3	Дисперсионный анализ		6		6	11
4	Факторный анализ		7		7	11
5	Дискриминантный анализ		7		7	11
<i>Итого по дисциплине:</i>			32		32	53

Примечание: Л – лекции, ПЗ – практические занятия / семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

2.3 Содержание разделов дисциплины:

2.3.1 Занятия лекционного типа.

№	Наименование раздела	Темы лекционных занятий	Форма текущего контроля
1	2	3	4
1.	Множественный регрессионный анализ	Понятие множественной регрессии. Отбор факторов при построении множественной регрессии. Мультиколлинеарность. Выбор формы уравнения регрессии. Оценка параметров уравнения линейной множественной регрессии. Качество оценок МНК линейной множественной регрессии. Теорема Гаусса-Маркова. Проверка качества уравнения регрессии. F-критерий Фишера. Частные уравнения регрессии. Частная корреляция. Обобщенный метод наименьших квадратов. Гетероскедастичность. Построение регрессионных моделей при наличии автокорреляции остатков. Регрессионные модели с переменной структурой. Фиктивные переменные.	У
2.	Кластерный анализ	Введение в кластерный анализ: измерение близости объектов; характеристики близости объектов. Методы кластерного анализа. Иерархические алгоритмы. Расстояния между кластерами. Процедуры эталонного типа.	У

3.	Дисперсионный анализ	Основные понятия. Методика дисперсионного анализа. Описание процедуры Factorial ANOVA. Описание Repeat measures ANOVA.	У
4.	Факторный анализ	Основные понятия. Методика факторного анализа. Выделение факторов. Критерий Кайзера. Критерий каменистой осыпи. Задача вращения общих факторов. Интерпретация результатов.	У
5.	Дискриминантный анализ	Введение в дискриминантный анализ. Дискриминация: коэффициенты канонической дискриминантной функции; число дискриминантных функций. Классифицирующие функции: применение элементарных классифицирующих функций; классификация объектов с помощью выборочного расстояния Махаланобиса; классификационная матрица.	У

2.3.2 Занятия семинарского типа.

Занятия семинарского типа не предусмотрены

2.3.3 Лабораторные занятия.

№	Наименование лабораторных работ	Форма текущего контроля
1	3	4
1.	Множественный регрессионный анализ.	ЛР
2.	Кластерный анализ.	ЛР
3.	Дисперсионный анализ.	ЛР
4.	Факторный анализ.	ЛР
5.	Дискриминантный анализ.	ЛР К

Защита лабораторной работы (ЛР), выполнение курсового проекта (КП), курсовой работы (КР), расчетно-графического задания (РГЗ), написание реферата (Р), эссе (Э), коллоквиум (К), тестирование (Т) и т.д.

2.3.4 Примерная тематика курсовых работ (проектов)

Курсовые работы не предусмотрены

2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1	Проработка учебного (теоретического) материала	Литература из основного и дополнительного списков
2	Подготовка к текущему контролю	Образцы проектов в пакете Statistica по темам лабораторных занятий в электронном виде

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

– в форме электронного документа,

Для лиц с нарушениями слуха:

– в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

– в форме электронного документа,

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. Образовательные технологии.

Образовательный процесс по дисциплине строится на основе комбинации следующих образовательных технологий.

Интегральную модель образовательного процесса по дисциплине формируют технологии методологического уровня: технология поэтапного формирования умственных действий, технология развивающего обучения, элементы технологии развития критического мышления.

Реализация данной модели предполагает использование следующих технологий стратегического уровня (задающих организационные формы взаимодействия субъектов образовательного процесса), осуществляемых с использованием определенных тактических процедур:

– лекционные (вводная лекция, информационная лекция, обзорная лекция, лекция-консультация, проблемная лекция);

– лабораторные (углубление знаний, полученных на теоретических занятиях, решение задач);

– тренинговые (формирование определенных умений и навыков, формирование алгоритмического мышления);

– активизации познавательной деятельности (приемы технологии развития критического мышления через чтение и письмо, работа с литературой, подготовка презентаций по темам домашних работ);

– самоуправления (самостоятельная работа студентов, самостоятельное изучение материала).

Рекомендуется использование информационных технологий при организации коммуникации со студентами для представления информации, выдачи рекомендаций и консультирования по оперативным вопросам (электронная почта), использование мультимедиа-средств при проведении лекционных и лабораторных занятий.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

4. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации.

4.1 Фонд оценочных средств для проведения текущего контроля.

Текущий контроль осуществляется в ходе проведения лабораторных занятий в виде устного опроса, выполнения лабораторных работ, коллоквиумов. Перечень заданий к лабораторным занятиям приведен в фонде оценочных средств по дисциплине «Многомерный статистический анализ».

4.2 Фонд оценочных средств для проведения промежуточной аттестации.

4.2.1. Пример (вариант) для промежуточной аттестации (экзамена) по итогам освоения дисциплины

Задача. Имеются следующие данные о ценах и дивидендах по обыкновенным акциям, а также данные о доходности компании.

№	цена акции, долл. США	доходность капитала, %	уровень дивидендов, %
1	25	15,2	2,6
2	20	13,9	2,1
3	15	15,8	1,5
4	34	12,8	ЗД
5	20	6,9	2,5
6	33	14,6	ЗД
7	28	15,4	2,9
8	30	17,3	2,8
9	23	13,7	2,4
10	24	12,7	2,4
11	25	15,3	2,6
12	26	15,2	2,8
13	26	12,0	2,7
14	20	15,3	1,9
15	20	13,7	1,9
16	13	13,3	1,6
17	21	15,1	2,4
18	31	15,0	3,0
19	26	11,2	ЗД
20	11	12,1	2,0

1) Проверить факторы на наличие коллинеарности.

- 2) Построить линейное уравнение множественной регрессии и пояснить экономический смысл его параметров.
- 3) Определить значения коэффициента множественной корреляции и коэффициента детерминации.
- 4) Проверить значимость уравнения при заданном уровне значимости (для каждого варианта свой уровень значимости).
- 5) Проверить значимость коэффициентов при заданном уровне значимости. Для значимых коэффициентов построить доверительный интервал.
- 6) Определить средние частные коэффициенты эластичности. Проинтерпретировать их.
- 7) Найти коэффициенты частной корреляции. Проанализировать их.
- 8) Построить уравнение линейной регрессии с учетом только значимых факторов.

Задача. Больные гипертиреозом (увеличение щитовидной железы) общим числом 23 человека были разделены на три группы.

Группа 1. Лечение оказалось успешным; проведенное через большой промежуток времени клиническое обследование показало, что пациент здоров.

Группа 2. Лечение безуспешно, т. е. состояние больного осталось без изменения.

Группа 3. Исход лечения успешен, но в дальнейшем возможен рецидив.

По результатам обследования 23 пациентов имеются следующие измерения:

y₆ – йод, регистрируемый через 3 часа после принятия испытательной дозы;

y₉ – йод, регистрируемый через 48 часов после принятия испытательной дозы;

y₁₀ – содержание в крови белковосвязанного йода (РВ¹³¹J) через 48 часов;

k_l – номер группы.

Конкретные результаты приведены в табл.

Таблица

ДАННЫЕ О 23 БОЛЬНЫХ ГИПЕРТИРЕОЗОМ, РАЗДЕЛЕННЫЕ НА ТРИ ГРУППЫ

№	k _l	y ₆	y ₉	y ₁₀	№	k _l	y ₆	y ₉	Y ₁₀
1	1	14.4	25.1	0.20	1	1	54.0	57.0	0.19
2	1	20.1	40.1	0.11	3	1	16.1	20.6	0.22
					4				
3	1	24.1	32.1	0.17	1	1	57.5	74.5	0.49
					5				
4	1	11.1	16.9	0.12	1	1	37.8	63.0	0.32
					6				
5	1	16.3	32.1	0.36	1	2	55.8	48.0	2.74
					7				
6	1	40.5	64.4	0.21	1	2	75.0	60.0	1.37
					8				
7	1	52.7	50.0	0.53	1	2	72.0	65.0	0.70
					9				

8	1	20.8	22.3	0.13	2	2	70.6	45.0	1.40
9	1	14.0	3.1	0.18	0	3	24.1	45.0	0.22
1	1	27.0	41.7	0.19	1	3	33.2	55.0	0.01
0	1	44.3	63.8	0.22	2	3	30.4	44.6	0.09
1	1	47.5	50.1	0.29	2				
1	1				3				
2									

По матрице исходных данных найти средние и стандартные отклонения дискриминантных переменных, общую T и внутригрупповую W матрицы сумм квадратов и перекрестных произведений. Найти также коэффициенты канонической дискриминантной функции, коэффициент канонической корреляции, Λ -статистику Уилкса, статистику хи-квадрат, уровень значимости. Построить классифицирующие функции.

Вопросы к экзамену

1. Понятие корреляционной матрицы.
2. Линейная множественная регрессионная модель.
3. Качество оценок МНК линейной множественной регрессии. Теорема Гаусса-Маркова.
4. Проверка качества уравнения линейной множественной регрессии. F-критерий Фишера.
5. Точность коэффициентов регрессии. Доверительные интервалы.
6. Методы кластерного анализа.
7. Иерархические алгоритмы.
8. Расстояния между кластерами.
9. Процедуры эталонного типа.
10. Методика дискриминантного анализа.
11. Что характеризует Лямбда Уилкса?
12. Что показывают квадраты расстояний Махаланобиса?
13. Какое максимальное число канонических дискриминантных функций допустимо в дискриминантном анализе?
14. Какую информацию дают стандартизованные и структурные коэффициенты дискриминантной функции?
15. Опишите процедуру отбора переменных с помощью стандартизованных и структурных коэффициентов.
16. Какова интерпретация канонического коэффициента корреляции?.
17. В каком случае учет априорных вероятностей может сильно изменить результаты классификации?
18. Методика факторного анализа.
19. Суть задачи вращения общих факторов
20. Критерий Кайзера.
21. Критерий Каменистой осыпи.
22. Метод главных компонент.

4.2.2. Критерии оценки знаний

Экзамен состоит из 2 частей:

- 1) Теоретическая часть (вопросы приведены выше).

2) Практическая часть (решение задач).

Критерии оценки качества освоения студентами дисциплины:

- пороговый («оценка «удовлетворительно»);
- стандартный (оценка «хорошо»);
- эталонный (оценка «отлично»);

Критерий	В рамках формируемых компетенций студент демонстрирует
пороговый	знание и понимание теоретического содержания курса с незначительными пробелами; несформированность некоторых практических умений при применении знаний в конкретных ситуациях, низкое качество выполнения учебных заданий (не выполнены, либо имеют низкую оценку); низкий уровень мотивации учения;
стандартный	полное знание и понимание теоретического содержания курса, без пробелов; недостаточную сформированность некоторых практических умений при применении знаний в конкретных ситуациях; достаточное качество выполнения всех предусмотренных программой обучения учебных заданий (ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками); средний уровень мотивации учения;
эталонный	полное знание и понимание теоретического содержания курса, без пробелов; сформированность необходимых практических умений при применении знаний в конкретных ситуациях, высокое качество выполнения всех предусмотренных программой обучения учебных заданий (оценены числом баллов, близким к максимальному); высокий уровень мотивации учения.

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 (модулю) 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.

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

5.1 Основная литература:

1. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для прикладного бакалавриата / В. Е. Гмурман. — 11-е изд., перераб. и доп. — М. : Издательство Юрайт, 2018. — 404 с. — (Серия : Бакалавр. Прикладной курс). — ISBN 978-5-534-00247-8. — Режим доступа : www.biblio-online.ru/book/AC41B7DD-F936-4105-9511-9BD045A42CFD .

2. Дубина, И. Н. Основы математического моделирования социально-экономических процессов : учебник и практикум для бакалавриата и магистратуры / И. Н. Дубина. — М. : Издательство Юрайт, 2018. — 349 с. — (Серия : Бакалавр и магистр. Академический курс). — ISBN 978-5-534-00501-1. — Режим доступа : www.biblio-online.ru/book/AE81649F-D411-4FF5-8733-614106E0D831 .

Для освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья имеются издания в электронном виде в электронно-библиотечных системах.

5.2 Дополнительная литература:

1. Симчера В.М. Методы многомерного анализа статистических данных [Электронный ресурс] : учеб. пособие — Электрон. дан. — Москва : Финансы и статистика, 2008. — 400 с. — Режим доступа: <https://e.lanbook.com/book/1005> — Загл. с экрана.

2. Халафян, А.А. Statistica 6. Статистический анализ данных [Текст] : учеб. пособие для вузов / А. А. Халафян. – 3-е изд. – М.: Бином пресс, 2008. – 508 с. – (Сер. "Учебник"). - Ц. - ISBN 978-5-9518021-5-6

3. Черткова, Е. А. Статистика. Автоматизация обработки информации : учебное пособие для вузов / Е. А. Черткова. — 2-е изд., испр. и доп. — М. : Издательство Юрайт, 2018. — 195 с. — (Серия : Университеты России). — ISBN 978-5-534-01429-7. — Режим доступа: www.biblio-online.ru/book/0CBA0F5B-1227-46F3-8C8E-D9BAB4AC306A .

5.3. Периодические издания:

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

7. Методические указания для обучающихся по освоению дисциплины (модуля).

По курсу предусмотрено проведение лекционных занятий, на которых дается основной теоретический материал, рассматриваются основные приёмы решения задач и решаются примеры практических задач.

На лабораторных занятиях студенты, решая семестровые задания, приобретают практические навыки применения многомерных статистических методов для анализа и прогнозирования конкретных экономических процессов с использованием реальной статистической информации (данных), выявления количественной связи между изучаемыми показателями и влияющими на них факторами.

Важнейшим этапом курса является самостоятельная работа по дисциплине «Многомерный статистический анализ», во время которой студенты осуществляют проработку необходимого материала, используя литературу из основного и дополнительного списков, готовятся к текущему контролю, изучая примеры задач, рассмотренных на лекциях и на лабораторных занятиях, и образцы моделей по темам лабораторных занятий (выдаются студентам в электронном виде).

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).

8.1 Перечень информационных технологий.

8.2 Перечень необходимого программного обеспечения.

Пакеты MS Excel, Statistica 10.

8.3 Перечень информационных справочных систем:

1. Справочно-правовая система «Консультант Плюс» (<http://www.consultant.ru>)
2. Электронная библиотечная система eLIBRARY.RU (<http://www.elibrary.ru/>)

9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)

№	Вид работ	Материально-техническое обеспечение дисциплины (модуля) и оснащенность
1.	Лекционные занятия	Лекционная аудитория
2.	Лабораторные занятия	Лаборатория, укомплектованная компьютерами для работы студентов и компьютером для преподавателя, подключенным к интерактивной доске.
3.	Текущий контроль, промежуточная аттестация	Лаборатория, укомплектованная компьютерами для работы студентов и компьютером для преподавателя, подключенным к интерактивной доске.
4.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета.

