

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное учреждение
высшего профессионального образования
"Казанский (Приволжский) федеральный университет"
Институт вычислительной математики и информационных технологий



УТВЕРЖДАЮ

Проректор
по образовательной деятельности КФУ
Проф. Таюрский Д.А.

"__" _____ 20__ г.

Программа дисциплины

Дисперсионный анализ Б1.В.ОД.7

Направление подготовки: 01.03.02 - Прикладная математика и информатика

Профиль подготовки: Теория вероятностей и математическая статистика

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Автор(ы):

Симушкин С.В.

Рецензент(ы):

Чупрунов А.Н.

СОГЛАСОВАНО:

Заведующий(ая) кафедрой: Турилова Е. А.

Протокол заседания кафедры No ____ от " ____ " _____ 201__ г

Учебно-методическая комиссия Института вычислительной математики и информационных технологий:

Протокол заседания УМК No ____ от " ____ " _____ 201__ г

Регистрационный No

Казань
2015

Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре основной образовательной программы
3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля
4. Структура и содержание дисциплины/ модуля
5. Образовательные технологии, включая интерактивные формы обучения
6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов
7. Литература
8. Интернет-ресурсы
9. Материально-техническое обеспечение дисциплины/модуля согласно утвержденному учебному плану

Программу дисциплины разработал(а)(и) доцент, к.н. (доцент) Симушкин С.В. кафедра математической статистики отделение прикладной математики и информатики ,
Sergey.Simushkin@kpfu.ru

1. Цели освоения дисциплины

изучить методы обработки статистических данных, зависящих от ряда контролируемых факторов, возникающих при анализе физических, экономических, медицинских процессов естествознания;

ознакомить с основными принципами проведения статистических экспериментов;

научить решать практические задачи с использованием методов дисперсионного анализа;

дать представление о математических (теоретико-вероятностных) основаниях построения процедур проверки многомерных гипотез и оценивания многомерных характеристик

2. Место дисциплины в структуре основной образовательной программы высшего профессионального образования

Данная учебная дисциплина включена в раздел " Б1.В.ОД.7 Дисциплины (модули)" основной образовательной программы 01.03.02 Прикладная математика и информатика и относится к обязательные дисциплины. Осваивается на 3 курсе, 6 семестр.

Курс "Дисперсионный анализ" входит в число курсов по выбору профессионально-го цикла подготовки бакалавра по направлению "Прикладная математика и информатика" (профиль "Теория вероятностей и математическая статистика") и является разделом Б3.В.4 ООП.

Изучению курса предшествует изучение базовых дисциплин "Линейная алгебра", "Математический анализ" и "Теория вероятностей и математическая статистика". В результате освоения предшествующих дисциплин студент должен:

знать:

□ основные понятия и методы линейной алгебры (решение систем линейных уравнений, обращение матриц, приведение квадратичных форм к главным осям, формирование базиса линейных пространств);

□ основные способы решения оптимизационных задач (метод Лагранжа), основные интегральные соотношения (интегралы Эйлера, Дирихле, Лапласа);

□ основные вероятностные законы (нормальный, Фишера, хи-квадрат), метод характеристических функций и дельта-метод отыскания асимптотического распределения;

□ принципы формирования статистических гипотез и их основных вероятностных характеристик (ошибки 1-го и 2-го рода, мощность критерия, среднеквадратическая ошибка, распределение статистик, асимптотическое распределение);

уметь:

□ производить алгебраические операции над матрицами и векторами;

□ решать линейные уравнения в матричной форме;

□ решать задачи на экстремум функций многих переменных;

□ использовать разложение в ряд Тейлора функцию многих переменных;

□ использовать основные алгебраические и тригонометрические тождества для преобразования алгебраических выражений;

□ находить распределение статистик и их основные характеристики (среднее значение, ковариация, корреляция);

владеть:

□ навыками использования математических справочников и таблиц;

□ приемами работы в основных пакетах прикладных программ ("Excel", "Mathematica").

Дисциплина изучается на 3 курсе обучения.

3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОК-1 (общекультурные компетенции)	владение культурой мышления, способность к обобщению, анализу, восприятию информации, постановке цели и выбору путей ее достижения
ОК-14 (общекультурные компетенции)	способность использовать в научной и познавательной деятельности, а также в со-циальной сфере профессиональные навыки работы с информационными и компьютерны-ми технологиями
ОК-15 (общекультурные компетенции)	способность работы с информацией из различных источников, включая ресурсы сети Интернет, для решения профессиональных и социальных задач
ПК-20 (профессиональные компетенции)	использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования
ПК-6 (профессиональные компетенции)	способность собирать, обрабатывать, анализировать и систематизировать научно-техническую информацию по тематике исследования, использовать достижения отечест-венной и зарубежной науки, техники и технологии

В результате освоения дисциплины студент:

1. должен знать:

основные вероятностные законы, описывающие поведение случайных векторов (нормальный закон, мультиномиальный закон);

основные принципы построения моделей регрессионных связей между исследуемыми характеристиками (регрессия с линейной зависимостью от неизвестных параметров);

основные формы представления многомерных данных и особенности работы с ними;

методы, применяемые для анализа и обработки многомерных совокупностей (метод наименьших квадратов, критерий Фишера, ранговые критерии, несмещенные оценки параметров модели, таблицы регрессионных моделей);

основные методы построения оптимальных планов проведения факторных экспериментов (полный факторный план, дробные реплики, блочные схемы, латинские планы, D-, G-, A-, L-, Q-оптимальные планы);

методы построения оптимальных планов для полиномиальной и тригонометрической регрессий

2. должен уметь:

применять методы дисперсионного анализа для обработки реальных числовых данных, учитывая границы применимости математической модели;

применять специализированные программные продукты для проведения вычислительных процедур дисперсионного анализа;

реализовывать процедуры дисперсионного анализа в рамках имеющихся средств обработки данных;

выбирать методику статистического исследования экспериментальных данных

3. должен владеть:

методами построения математической модели профессиональных задач и содержательной интерпретации полученных результатов.

навыками выбора математических методов обработки экспериментальных данных, адекватных целям исследования;

□ навыками реализации математических методов обработки экспериментальных данных в виде прикладных программных продуктов;

навыками составления отчетов по методикам исследования и их реализации в виде ПО, анализа результатов обработки

4. должен демонстрировать способность и готовность:

способность и готовность применять на практике основные методы дисперсионного анализа

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 4 зачетных(ые) единиц(ы) 144 часа(ов).

Форма промежуточного контроля дисциплины экзамен в 6 семестре.

Суммарно по дисциплине можно получить 100 баллов, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов. Минимальное количество для допуска к зачету 28 баллов.

86 баллов и более - "отлично" (отл.);

71-85 баллов - "хорошо" (хор.);

55-70 баллов - "удовлетворительно" (удов.);

54 балла и менее - "неудовлетворительно" (неуд.).

4.1 Структура и содержание аудиторной работы по дисциплине/ модулю

Тематический план дисциплины/модуля

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Классическая линейная модель регрессии. Критерий Фишера	6	1-3	0	0	12	коллоквиум
2.	Тема 2. Модели дисперсионного анализа	6	4-11	0	0	32	коллоквиум
3.	Тема 3. Дисперсионный анализ многомерных данных	6	12-14	0	0	12	коллоквиум
4.	Тема 4. Оптимизация факторных планов	6	15-18	0	0	16	коллоквиум
.	Тема . Итоговая форма контроля	6		0	0	0	экзамен

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
Итого				0	0	72	

4.2 Содержание дисциплины

Тема 1. Классическая линейная модель регрессии. Критерий Фишера

лабораторная работа (12 часа(ов)):

Классическая линейная модель. Оценки метода наименьших квадратов. Теорема Гаусса-Маркова. Канонические переменные и распределение оценок. Критерий Фишера

Тема 2. Модели дисперсионного анализа

лабораторная работа (32 часа(ов)):

Однофакторный дисперсионный анализ. Непараметрические критерии однородности. Двухфакторный дисперсионный анализ с одним наблюдением в ячейке. Двухфакторный дисперсионный анализ с равным числом наблюдений в ячейке. Полный многофакторный анализ с взаимодействиями. Формальный дисперсионный анализ. Дисперсионный анализ со случайными факторами. Возможности пакета Statgraphics

Тема 3. Дисперсионный анализ многомерных данных

лабораторная работа (12 часа(ов)):

Задача проверки гипотезы о центре нормально распределенных данных. Критерий Фишера. Сравнение двух групп по большому числу признаков. Сравнение более двух групп. Статистика Уилкса. Элементы дискриминантного анализа

Тема 4. Оптимизация факторных планов

лабораторная работа (16 часа(ов)):

Полный факторный план. Дробные реплики. Анализ определяющих соотношений. Латинские планы. Блочные схемы. Оптимальные факторные планы. Информационная матрица. Примеры сравнения планов. Критерии оптимальности. Теоремы эквивалентности. Теорема Кифера-Волфовица. Примеры применения. Теоремы эквивалентности для L-оптимальных планов. D-оптимальные планы для полиномиальной и тригонометрической регрессии. Оптимальные планы первого порядка

4.3 Структура и содержание самостоятельной работы дисциплины (модуля)

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
1.	Тема 1. Классическая линейная модель регрессии. Критерий Фишера	6	1-3	Изучение методов классической регрессии по литературным источникам	6	коллоквиум

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
2.	Тема 2. Модели дисперсионного анализа	6	4-11	Изучение методов однофакторного анализа по литературным источникам, применение методов анализа к реал	16	коллоквиум
3.	Тема 3. Дисперсионный анализ многомерных данных	6	12-14	Изучение методов анализа многомерных данных по литературным источникам, применение методов анализа	6	коллоквиум
4.	Тема 4. Оптимизация факторных планов	6	15-18	Изучение методов факторного анализа по литературным источникам, применение методов анализа к реальны	8	коллоквиум
	Итого				36	

5. Образовательные технологии, включая интерактивные формы обучения

Часть материала теоретического характера преподается в виде лекций у доски или в виде презентаций на мультимедийном экране. Предполагается использование диалоговой формы ведения занятий с постановкой и решением проблемных задач, обсуждением дискуссионных моментов.

Некоторые разделы курса студенты изучают самостоятельно по указанным методическим материалам или по Интернет-источникам с последующим докладом в виде презентаций и дискуссией.

При проведении практических занятий, на которых теоретический материал применяется к конкретным данным, используются готовые статистические процедуры стандартных пакетов, а также предлагается студентам разработать самостоятельные вычислительные процедуры.

Перед каждым занятием, как лекционной, так и практической направленности проводится экспресс-опрос по пройденному теоретическому материалу.

В целях выработки навыков работы в коллективе и развития коммуникативных способностей, часть вычислительных заданий, а также проработку новейших методов дисперсионного анализа студенты выполняют, разбившись на творческие группы

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

Тема 1. Классическая линейная модель регрессии. Критерий Фишера

коллоквиум , примерные вопросы:

Ответить на вопросы теоретического характера: 1. Формализуйте модель линейной регрессии. 2. Дайте определение оценок наименьших квадратов. 3. Сформулируйте теорему Гаусса-Маркова 4. Опишите структуру канонических переменных 5. Сформулируйте теорему о распределении ОНК 6. Опишите критерий Фишера

Тема 2. Модели дисперсионного анализа

коллоквиум , примерные вопросы:

Ответить на вопросы теоретического характера: 1. Опишите предположения однофакторного (двухфакторного, многофакторного) дисперсионного анализа 2. Как строятся неапраметрические критерии однородности? 3. Таблица двухфакторного дисперсионного анализа с одним (несколькими равными) наблюдением в ячейке 4. Таблица полного многофакторного анализа с взаимодействиями. Обработать реальные данные изученными методами.

Тема 3. Дисперсионный анализ многомерных данных

коллоквиум , примерные вопросы:

Ответить на вопросы теоретического характера: 1. Дать формальную структуру многомерной модели дисперсионного анализа. 2. Описать критерий Фишера для задач проверки гипотезы о центре нормально распределенных данных, сравнения двух групп по большому числу признаков, сравнения более двух групп. 3. Статистика Уилкса и ее асимптотическое распределение. 4. Элементарные и неэлементарные дискриминантные признаки. Применение для дискриминации. Провести обработку средствами пакета Excel реальных медицинских данных.

Тема 4. Оптимизация факторных планов

коллоквиум , примерные вопросы:

Ответить на вопросы теоретического характера: 1. Описать полный факторный план и его дробные реплики по заданным параметрам модели и генерирующим соотношениям. 2. Латинские, греко-латинские планы, блочные схемы. 3. Дать определения различным критериям оптимальности. 4. Описать причину, по которой сравнение происходит по информационной матрице. 5. Рандомизированные планы и их дисперсионная матрица. Дисперсия поверхности отклика. 6. Теорема эквивалентности Кифера-Волфовица. 7. Теоремы эквивалентности для L-оптимальных, D-оптимальных и G-оптимальных планов для полиномиальной и тригонометрической регрессии Провести сравнение нескольких планов по различным критериям оптимальности.

Тема . Итоговая форма контроля

Примерные вопросы к экзамену:

По первым трем разделам дисциплины проводится коллоквиум в виде опроса по пройденному материалу.

Экзаменационная оценка по дисциплине (из расчета 100 баллов) складывается из текущей работы

(50 баллов = 36 баллов за ответы на коллоквиумах + 14 баллов за подготовку доклада по выбранной теме) и ответа на экзаменационный билет (50 баллов).

Экзаменационный билет включает в себя три блока вопросов.

В первом блоке содержится 5 вопросов, для правильного ответа на которые необходимо привести формулировку

определений и утверждений из списка определений и основных теорем (30% экзаменационной оценки).

Во втором блоке необходимо сформулировать теоретическое утверждение из списка основных теорем и

привести подробную схему его доказательства (50% оценки).

В третьем блоке решить задачу с применением любых доступных средств обработки статистических данных (20% экзаменационной оценки)

Список вопросов (вопросы теоретического характера отмечены *).

1. Формализация модели линейной регрессии.
2. Определение оценок наименьших квадратов.
- 3*. Вид оценок НК для матрицы плана полного ранга.
- 4*. Доказать несмещенность ОМНК.
5. Определение параметрической функции, допускающей оценку.
- 6*. Теорема Гаусса-Маркова
7. Описать структуру канонических переменных
- 8.* Теорема о распределении ОМНК и суммы ошибок.
9. Доверительная область для вектора параметрических функций.
- 10*. Описать критерий Фишера
11. Описать модель однофакторного (двухфакторного, многофакторного) дисперсионного анализа
12. Неапараметрические критерии однородности
13. Таблица двухфакторного дисперсионного анализа с одним наблюдением в ячейке
14. Таблица двухфакторного дисперсионного анализа с несколькими наблюдениями в ячейке
15. Таблица полного многофакторного анализа с взаимодействиями.
16. Структура многомерной модели дисперсионного анализа.
- 17*. Критерий Фишера для задач проверки гипотезы о центре нормально распределенных данных, сравнения двух групп по большому числу признаков, сравнения более двух групп.
18. Критерий Фишера для задач сравнения двух групп по большому числу признаков.
19. Критерий Фишера для задач сравнения более двух групп по большому числу признаков.
20. Статистика Уилкса и ее асимптотическое распределение.
- 21*. Элементарные и неэлементарные дискриминантные признаки. Применение для дискриминации.
22. Описать полный факторный план и его дробные реплики по заданным параметрам модели и генерирующим соотношениям.
23. Латинские, греко-латинские планы, блочные схемы.
24. Дать определения различным критериям оптимальности.
25. Описать причину, по которой сравнение происходит по информационной матрице.
26. Рандомизированные планы и их дисперсионная матрица. Дисперсия поверхности отклика.
- 27*. Теорема эквивалентности Кифера-Волфовица.
28. Теоремы эквивалентности для L-оптимальных, D-оптимальных и G-оптимальных планов для полиномиальной регрессии
29. Теоремы эквивалентности для L-оптимальных, D-оптимальных и G-оптимальных планов для тригонометрической регрессии

7.1. Основная литература:

1. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики. - СПб.: Лань, 2011. - 256с.
ЭБС "Лань": http://e.lanbook.com/books/element.php?pl1_id=2026
2. Боровков А.А. Математическая статистика. - СПб.: Лань, 2010. - 704 с.
ЭБС "Лань": http://e.lanbook.com/books/element.php?pl1_id=3810

3. Свешников А.А. Прикладные методы теории вероятностей. - СПб.: Лань, 2012. - 480 с.
ЭБС "Лань": http://e.lanbook.com/books/element.php?pl1_id=3184
4. Свешников А.А. Прикладные методы теории марковских процессов. - СПб.: Лань, 2007. - 192 с.
ЭБС "Лань": http://e.lanbook.com/books/element.php?pl1_id=590
5. Свешников А.А. Прикладные методы теории случайных функций. - СПб.: Лань, 2011. - 464с
ЭБС "Лань": http://e.lanbook.com/books/element.php?pl1_id=656

7.2. Дополнительная литература:

1. Коралов Л.Б., Синай Я.Г. Теория вероятностей и случайные процессы. - М.: МЦНМО, 2014. - 408 с.
ЭБС "Лань": <http://e.lanbook.com/view/book/56404/>
2. Королев В.Ю., Бенинг В.Е., Шоргин С.Я. Математические основы теории риска. - М.: Физматлит, 2011. - 620 с.
ЭБС "Лань": <http://e.lanbook.com/view/book/2742/>
3. Стоянов Й. Контрпримеры в теории вероятностей. - М.: МЦНМО, 2014. - 294 с.
ЭБС "Лань": <http://e.lanbook.com/view/book/56414/>
4. Алон Н., Спенсер Дж. Вероятностный метод. - М.: БИНОМ. Лаборатория знаний, 2013. - 320 с.
ЭБС "Лань": <http://e.lanbook.com/view/book/8716/>
5. Лагутин М.Б. Наглядная математическая статистика. - М.: БИНОМ. Лаборатория знаний, 2013. - 472 с.
ЭБС "Лань": <http://e.lanbook.com/view/book/56887/>

7.3. Интернет-ресурсы:

- MashineLearning - <http://www.machinelearning.ru/>
StatPlus - http://www.statplus.net.ua/ru/help/source/a_anova_single.htm
StatSoft - электронный учебник по Statistica -
<http://www.statsoft.ru/home/textbook/modules/stanman.html>
Анализ экономических данных с использованием Statistica 6,0 -
http://cdo.bseu.by/stat1/lab2_1.htm
Федеральный образовательный портал - <http://ecsocman.hse.ru/>

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Дисперсионный анализ" предполагает использование следующего материально-технического обеспечения:

Теоретическая часть лекционного материала представляется с помощью презентаций в формате PowerPoint или PDF, для чего используются аудитории учебной доской и мелом или с мультимедийным проектором; численная реализация изучаемых методов осуществляется с помощью пакетов программ "Excel" и "Mathematica" в компьютерных классах с соответствующим про-граммным обеспечением и возможностью выхода в Интернет

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 01.03.02 "Прикладная математика и информатика" и профилю подготовки Теория вероятностей и математическая статистика .

Автор(ы):

Симушкин С.В. _____

"__" _____ 201__ г.

Рецензент(ы):

Чупрунов А.Н. _____

"__" _____ 201__ г.