

Министерство образования и науки Российской Федерации
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ЭКОНОМИКИ И УПРАВЛЕНИЯ «НИНХ»

Кафедра Статистики

Рег. № 15-14А/01



УТВЕРЖДАЮ:
И.о. проректора по научной
работе и внешним связям
П.А. Новгородов
«12» августа 2014 г.

**РАБОЧАЯ ПРОГРАММА
УЧЕБНОЙ ДИСЦИПЛИНЫ**

Программа составлена в соответствии с федеральным государственным образовательным стандартом высшего образования по соответствующему направлению

Учебная дисциплина

СОВРЕМЕННЫЕ ТЕХНОЛОГИИ АНАЛИЗА МАССИВОВ ДАННЫХ

по направлению подготовки 09.06.01 Информатика и вычислительная техника

направленность (профиль) программы «Информационные системы и процессы»

Квалификация: «Исследователь, преподаватель-исследователь»

Новосибирск 2014

Рабочая программа дисциплины разработана
Серга Людмилой Константиновной, канд. экон. наук, доцентом кафедры
статистики
Иониным Владимиром Георгиевичем, канд. экон. наук, профессором кафедры
статистики

Учебно-методическое обеспечение согласовано с библиотекой университета

Директор библиотеки



Н.Ю. Долгова

Утверждено на заседании кафедры статистики
(протокол от «23» августа 2014 г. № 1).

Заведующий кафедрой,
д-р экон. наук, профессор



В.В. Глинский

СОГЛАСОВАНО:

Ответственный за ОП ВО по программе подготовки научно-педагогических кадров в аспирантуре по направлению подготовки 09.06.01 «Информатика и вычислительная техника» направленности (профилю) «Информационные системы и процессы»

канд.экон.наук, доцент



П.М. Пашков

канд.техн.наук, ст.науч.сотр



А.Л.Осипов

канд.техн.наук, доцент



Т.М. Пестунова

© Новосибирский государственный университет экономики и управления
«НИНХ»

**ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ,
вносимых в Рабочую программу дисциплины
«Современные технологии анализа массивов данных».**

Номер изменения	Дата	Страницы с изменениями	Перечень и содержание измененных разделов Рабочей программы
1	30.06.2015	3	Раздел 1. Организационно-методический п.1.3.Место дисциплины в структуре ОП ВО Обновление элементов ОП ВО в соответствии с изменениями в учебных планах

РАЗДЕЛ 1. ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЙ

1.1. Выписка из ОП ВО по направлению подготовки, определяющая место дисциплины в учебном процессе

Дисциплина ФТД.5 «Современные технологии анализа массивов данных» изучается слушателями направления 09.06.01 «Информатика и вычислительная техника», направленности (профиля) «Информационные системы и процессы», входит в учебный цикл ФТД «Факультативные дисциплины» и отвечает за формирование следующих компетенций: УК-3, ОПК-2, ОПК-4.

1.2. Цели и задачи учебной дисциплины

Основной целью дисциплины «Современные технологии анализа массивов данных» является дополнение и совершенствование знаний в области методов многомерного статистического анализа социально-экономических процессов, изучение современных проблем, направлений и методов на базе современных технологий и систем статистического анализа с использованием ПЭВМ.

Исходя из цели, в процессе изучения дисциплины решаются следующие задачи:

- сформировать у слушателей базу знаний о современных направлениях и возможностях методов многомерной статистики, тенденциях их развития, современных вычислительных технологиях и инструментах многомерного экономико-статистического анализа в социальной среде и экономической жизни общества;
- показать возможные направления современного многомерного статистического анализа при решении научных и практических задач и, в первую очередь, при написании научных квалификационных работ (диссертаций).

1.3. Место дисциплины в структуре ОП ВО

Дисциплина опирается на следующие элементы ОП ВО:

- Б1.Б1 «Иностранный язык»
- Б1.Б2 «История и философия науки»
- Б1.В.ОД.1 «Методология научных исследований»;
- Б1.В.ДВ.1 «Информационный поиск в научных исследованиях: информационные системы, технологии, ресурсы»;

Дисциплина выступает опорой для следующих элементов ОП ВО:

- Б2.1 «Научно-исследовательская практика»;
- Б3. «Научные исследования»;
- Б4. «Государственная итоговая аттестация».

1.4. Компетенции, формируемые в результате изучения дисциплины

Код компетенции	Содержание компетенции	Код структурного элемента компетенции	Содержание структурных элементов компетенции	Пороговый уровень освоения структурных элементов компетенции
Универсальные компетенции				
УК-3	Готовность участвовать в работе российских и международных исследовательских коллективов по решению научных и научно-образовательных задач	УК-3-1	Знать: основы инновационной деятельности, принципы организации инновационных процессов, специфические особенности инновационных процессов в бизнесе и ИТ, промышленные стандарты в области новых информационных технологий, методы выявления бизнес-возможностей.	Знает: основы инновационной деятельности, принципы организации инновационных процессов, специфические особенности инновационных процессов в бизнесе и ИТ, промышленные стандарты в области новых информационных технологий, методы выявления бизнес-возможностей.
		УК-3-2	Уметь: проводить анализ бизнес-возможностей новых информационных технологий, формировать новые бизнес-модели, в том числе на основе новых информационных технологий, проектировать перспективную бизнес-архитектуру, осуществлять изменения архитектуры предприятия.	Умеет: проводить анализ бизнес-возможностей новых информационных технологий, формировать новые бизнес-модели, в том числе на основе новых информационных технологий, проектировать перспективную бизнес-архитектуру, осуществлять изменения архитектуры предприятия.
		УК-3-3	Владеть: методами анализа новых информационных технологий, методами анализа бизнес-возможностей, методами и инструментами построения новых бизнес-моделей, методами управления изменениями	Владеет: методами анализа новых информационных технологий, методами анализа бизнес-возможностей, методами и инструментами построения новых бизнес-моделей, методами управления изменениями
Общепрофессиональный цикл				
ОПК-2	владение культурой научного исследования, в том числе с использованием	ОПК-2-1	Знать: специфику научных исследований в области информационных систем; тенденции в области развития инструментальных	Знает: специфику научных исследований в области информационных систем; тенденции в области развития инструментальных

Код компетенции	Содержание компетенции	Код структурного элемента компетенции	Содержание структурных элементов компетенции	Пороговый уровень освоения структурных элементов компетенции
	современных информационно-коммуникационных технологий		средств разработки информационных систем; пользоваться современными компьютерными технологиями поиска и анализа информации, работать с внутрироссийскими и зарубежными электронными базами данных; основы авторского права.	средств разработки информационных систем; пользоваться современными компьютерными технологиями поиска и анализа информации, работать с внутрироссийскими и зарубежными электронными базами данных; основы авторского права.
		ОПК-2-2	Уметь: выявлять объект и предмет научных исследований; осуществлять постановку проблемы научного исследования прикладных ИС; определять цели и задачи исследования; пользоваться информационными системами для проверки работ на антиплагиат.	Умеет: выявлять объект и предмет научных исследований; осуществлять постановку проблемы научного исследования прикладных ИС; определять цели и задачи исследования; пользоваться информационными системами для проверки работ на антиплагиат.
		ОПК-2-3	Владеть: навыками проводить эксперименты, структурировать материал, объяснять, доказывать и защищать свои идеи, делать умозаключения и выводы по исследованию инноваций.	Владеет: навыками проводить эксперименты, структурировать материал, объяснять, доказывать и защищать свои идеи, делать умозаключения и выводы по исследованию инноваций.
ОПК-4	готовность организовать работу исследовательского коллектива в области профессиональной деятельности	ОПК-4-1	Знать: методы управления коллективом, специфику управления в малых группах, основы конфликтологии, основы лидерства, методы управления научно-исследовательскими работами; методы управления научными проектами в области информационных технологий.	Знает: методы управления коллективом, специфику управления в малых группах, основы конфликтологии, основы лидерства, методы управления научно-исследовательскими работами; методы управления научными проектами в области информационных технологий.

Код компетенции	Содержание компетенции	Код структурного элемента компетенции	Содержание структурных элементов компетенции	Пороговый уровень освоения структурных элементов компетенции
		ОПК-4-2	Уметь: осуществлять набор персонала, отбор и формирование проектной команды, планировать карьерный рост, организовывать эффективные встречи, оценивать производительность, разрабатывать техническое задание на НИР, календарный план НИР, управлять изменениями.	Умеет: осуществлять набор персонала, отбор и формирование проектной команды, планировать карьерный рост, организовывать эффективные встречи, оценивать производительность, разрабатывать техническое задание на НИР, календарный план НИР, управлять изменениями.
		ОПК-4-3	Владеть: навыками правильного взаимодействия и общения с сотрудниками и партнерами по бизнесу, ведения переговоров; методами и инструментальными средствами управления проектами преобразованием.	Владеет: навыками правильного взаимодействия и общения с сотрудниками и партнерами по бизнесу, ведения переговоров; методами и инструментальными средствами управления проектами преобразованием.

Овладение слушателями элементами компетенций «знать» соответствует пороговому уровню освоения компетенции, «знать» и «уметь» соответствует повышенному уровню освоения компетенции, «знать», «уметь» и «владеть» продвинутому уровню освоения компетенции.

РАЗДЕЛ 2. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

2.1. Трудоемкость освоения дисциплины

Трудоемкость дисциплины составляет:

Общая (зачетных единиц)	Общая (часов)	в том числе (часов)			
		контактной работы			самостоятельная работа
		всего	лекционных занятий	практических (семинар.) занятий	
3	108	72	6	66	36

2.2. Образовательные технологии

В процессе освоения дисциплины «Современные технологии анализа массивов данных» используются следующие образовательные технологии:

- лекционные занятия;
- практические (семинарские) занятия;
- самостоятельная работа аспирантов.

В ходе лекционных занятий раскрываются основные вопросы в рамках рассматриваемой темы, делаются акценты на наиболее сложные и важные положения изучаемого материала, которые должны быть приняты аспирантами во внимание. Материалы лекций являются основой для подготовки к практическим (семинарским) занятиям и выполнения заданий самостоятельной работы.

Целью практических (семинарских) занятий является контроль за степенью усвоения пройденного материала, ходом выполнения аспирантами самостоятельной работы и рассмотрение наиболее сложных и спорных вопросов в рамках темы занятия.

Самостоятельная работа аспирантов включает подготовку к практическим (семинарским) занятиям в соответствии с вопросами, представленными в Рабочей программе дисциплины, изучение теоретических основ современного многомерного статистического анализа и их методов практической реализации. Работа выполняется с использованием системы обработки данных «Statistica» и, при необходимости, текстового редактора MS Word, MS Excel и т.п.

При необходимости в процессе работы над заданием аспирант может получить индивидуальную консультацию у преподавателя. Также предусмотрено проведение консультаций аспирантов в ходе изучения материала дисциплины в течение семестра.

Методы обучения с применением интерактивных форм образовательных технологий:

- лекции с применением мультимедийного оборудования;
- консультации с применением мультимедийного оборудования и пакета прикладных программ.

2.3 Тематический план дисциплины

Наименование разделов и тем	Количество часов			Форма текущего контроля	Форма промежуточной аттестации
	Лекции	Практические занятия	Самостоятельная работа		
Тема 1. Версии системы «Statistica», общая технология статистического анализа в системе «Statistica»	0,5	7	6	реферат	
Тема 2. Направления современного многомерного статистического анализа данных в системе «Statistica»	0,5	7	8	тестирование	
Тема 3. Современный многомерный линейный и нелинейный регрессионный анализ в системе «Statistica».	1,0	8	8	тестирование	
Итого 1 часть	2	22	12		зачет
Тема 4. Современный факторный и компонентный анализ в системе «Statistica»	1,0	11	8	тестирование	
Тема 5. Современный анализ канонических корреляций в системе «Statistica»	1,0	11	6	тестирование	
Итого 2 часть	2	22	12		зачет
Тема 6. Многомерное метрическое шкалирования в системе «Statistica»	1,0	11			
Тема 7. Реализация нейросетевого моделирования в системе «Statistica»	1,0	11			
Итого 3 часть	2	22	12		зачет
Итого по дисциплине:	6	66	36		

2.4. Программа самостоятельной работы аспирантов

Самостоятельная работа аспирантов по дисциплине реализуется в следующих формах:

Формы СРС	Количество часов
Подготовка к практическим (семинарским) занятиям	8
Выполнение самостоятельной работы (РГР № 1)	4
Подготовка к практическим (семинарским) занятиям	8
Выполнение самостоятельной работы (РГР № 2)	4
Подготовка к практическим (семинарским) занятиям	8
Выполнение самостоятельной работы (РГР № 3)	4

2.5. Содержание отдельных разделов и тем

Тема 1. Версии системы «Statistica», общая технология статистического анализа в системе «Statistica»

Версии системы «Statistica», общая технология статистического анализа в системе «Statistica». Выбор конфигурации системы «Statistica». Подготовка и преобразование исходных данных, работа с отчетами, редактирование текстов, таблиц и графиков

Тема 2. Направления современного многомерного статистического анализа данных в системе «Statistica»

Характеристика направлений статистического анализа в системе «Statistica», модули статистического анализа, предоставляемые системой «Statistica».

Тема 3. Современный многомерный линейный и нелинейный регрессионный анализ в системе «Statistica»

Варианты нелинейной зависимости признаков, возможности выборы формы нелинейной парной связи в системе «Statistica», задание произвольной функциональной зависимости.

Процедуры построения уравнений множественной линейной и нелинейной регрессии, оценка статистической значимости параметров регрессии и регрессионной модели в целом. Модели исследования робастности исходных данных в системе «Statistica».

Тема 4. Современный факторный и компонентный анализ в системе «Statistica»

Основные теоретические положения линейной модели факторного анализа. Ортогональное факторное отображение. Косоугольное (не ортогональное) факторное отображение, вращение факторных осей. Проблема интерпретируемости главных факторов. Методы выделения факторов и

технология компонентного анализа в системе «Statistica». Использование главных факторов и компонент для многомерной классификации объектов.

Тема 5. Современный анализ канонических корреляций в системе «Statistica»

Каноническая корреляция, как метод разделения совокупности статистических показателей на факторное и результирующее подмножества. Теоретические основы метода канонических корреляций. Основные статистические характеристики в методе анализа канонических корреляций, оценка значимости выделенных канонических взаимосвязей, Использование аппарата канонических корреляций для исследования множественных регрессий. Использование аппарата канонических корреляций для исследования поведения мажоритарных и миноритарных агентов. Реализация метода канонических корреляций в системе «Statistica».

Тема 6. Многомерное метрическое шкалирование в системе «Statistica»

Теоретические основы многомерного метрического шкалирования. Оценка близости объектов, экспертная оценка и построение матрицы расстояний. Результаты многомерного метрического шкалирования совокупности объектов. Реализация техники многомерного шкалирования в системе «Statistica».

Тема 7. Реализация нейросетевого моделирования в системе «Statistica»

Базовая нейросетевая модель. Вероятностная нейронная сеть. Обобщенно регрессионная нейронная сеть. Примеры применения нейросетевого моделирования в системе «Statistica».

РАЗДЕЛ 3. СИСТЕМА ОЦЕНКИ КАЧЕСТВА ОСВОЕНИЯ ОБУЧАЮЩИМИСЯ ДИСЦИПЛИНЫ

3.1. Формы контроля:

- текущий контроль успеваемости;
- промежуточная аттестация.

3.1.1. Текущая аттестация.

Для контроля текущей успеваемости по установленному графику проводится тестирование. Текущий контроль успеваемости осуществляется для своевременной диагностики и возможной корректировки уровня знаний, умений и навыков обучающихся.

Для контроля при проведении практических (семинарских) занятий для в соответствии с учебным планом и графиком учебного процесса преподавателем используются такие формы текущего контроля, как обсуждение результатов решения на ПЭВМ задач по темам дисциплины, проведение опроса по одной или нескольким темам.

Кроме того, ежегодно по изучаемым темам дисциплины аспиранты выполняют самостоятельную расчетную работу по реальным социально-экономическим ситуациям или в рамках темы научной работы, готовят доклад и представляют презентацию своих исследований. К итоговой аттестации допускаются аспиранты, выполнившие все виды самостоятельной работы и получившие положительную оценку. Слушатели, не сдавшие рефераты и расчетную работу, не допускаются кафедрой к сдаче зачета как не выполнившие график учебного процесса по данной дисциплине.

При этом рекомендуется следующая структура доклада:

1. Формулировка проблемы;
2. Определение целей и задач исследования;
3. Логический анализ основных понятий;
4. Выдвижение гипотез;
5. Выбор технологий и составление методического инструментария;
6. Статистический анализ данных;
7. Выводы и рекомендации.

3.1.2. Промежуточная аттестация.

Для контроля усвоения данной дисциплины учебным планом предусмотрен зачет на 1, 2, 3 году обучения аспирантов.

3.2. Список вопросов для подготовки к зачету

Примерный перечень вопросов для подготовки к сдаче зачета по дисциплине.

1. Системы обработки статистических данных. Основные характеристики и возможности.
2. Типы документов в системе «Statistica». Основные направления и терминология статистического анализа данных в системе «Statistica»

3. Показатели многомерного регрессионного и корреляционного анализа в системе «Statistica»
4. Особенности и виды многомерных статистических графиков в в системе «Statistica»
5. Показатели робастности первичных данных в системе «Statistica»
6. Основные процедуры факторного и компонентного анализа в системе «Statistica»
7. Процедуры выбора общностей в факторном анализе в системе «Statistica». Анализ на главных осях.
8. Назначение современного анализа канонических корреляций. Возможности анализа канонических корреляций в системе «Statistica».
9. Теоретические основы метода канонических корреляций. Основные статистические характеристики метода в системе «Statistica».
10. Использование аппарата канонических корреляций для исследования множественных регрессий.
11. Теоретические основы и реализация метода многомерного метрического шкалирования в системе «Statistica».
12. Оценка близости объектов, экспертная оценка и построение матрицы расстояний методом метрического шкалирования в системе «Statistica».
13. Реализация метода многомерного шкалирования в системе «Statistica».
14. Интерпретация результатов процедуры многомерного метрического шкалирования совокупности объектов в системе «Statistica».
15. Определение базовой и вероятностной нейронной сети в системе «Statistica».
16. Построение обобщенной регрессионной нейронной сети в системе «Statistica».
17. Примеры применения нейросетевого моделирования в системе «Statistica».
18. Методы, альтернативные нейросетевому моделированию в системе «Statistica».
19. Многомерный дисперсионный анализ регрессий и корреляций в системе «Statistica».
20. Моделирование многомерных распределений в системе Statistica.

РАЗДЕЛ 4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

4.1. Основное (обязательное) обеспечение

4.1.1. Библиографический список:

а) учебники:

1. Нейронные сети. Statistica Neural Networks : методология и технологии современного анализа данных / под ред. В. П. Боровикова .- 2-е изд., перераб. и доп .- М. : Горячая линия - Телеком, 2008 .- 392 с. 27

б) учебные пособия:

1. Боровиков, Владимир Павлович. Прогнозирование в системе Statistica в среде Windows : Основы теории и интенсивная практика на компьютере : учеб. пособие для высш. учеб. заведений по специальности "Прикладная математика" / В. П. Боровиков, Г. И. Ивченко .- 2-е изд., перераб. и доп .- М. : Финансы и статистика, 2006 .- 368 с. 21

2. Халафян, Алексан Альбертович. Statistica 6. Статистический анализ данных : учеб. пособие / А. А. Халафян .- 2-е изд., [перераб. и доп.] .- М. : БИНОМ, 2010 .- 522 с. 27

3. Ионин, Владимир Георгиевич. Статистический анализ типовых экономических и социальных процессов на ПЭВМ : метод. пособие по курсам "Основы применения ЭВМ в стат. расчетах" и "Многомер. стат. методы" / В. Г. Ионин ; НГАЭиУ .- 2-е изд .- Новосибирск, 2005 .- 169 с. 28

4. Основы статистического анализа. Практ. по стат. мет. и исслед. операций с исп. пакетов STATISTICA и EXCEL: Уч.пос./ Э.А.Вуколов - 2 изд., испр. и доп. - М.: Форум:НИЦ Инфра-М, 2013. - 464 с. – Режим доступа: <http://znanium.com/bookread.php?book=369689>

4.1.2. Информационные источники:

сайт Федеральной службы государственной статистики РФ – www.gks.ru ,
сайт Территориального органа Федеральной службы государственной статистики по Новосибирской области – www.novosibstat.ru ,
сайты Правительства Новосибирской области и Сибирского федерального округа,
официальные сайты Правительства РФ – www.government.ru и других органов исполнительной власти Российской Федерации,
сайт компании «Statsoft».

4.1.3. Программные продукты: профессиональные системы статистического анализа: «Statistica» ,«SPSS» и др.

4.2. Дополнительное обеспечение

4.2.1. Библиографический список:

а) учебники

1. Статистика [Электронный ресурс]: Учебник / Л.П. Харченко, В.Г. Ионин, В.В. Глинский; Под ред. В.Г. Ионина. - 3-е изд., перераб. и доп. - М.: ИНФРА-М, 2010. - 445 с. – Режим доступа: <http://znanium.com/bookread.php?book=206690>
2. Глинский, Владимир Васильевич. Статистический анализ нечисловой информации : курс лекций / В. В. Глинский ; НГУЭУ .- Новосибирск : [Изд-во НГУЭУ], 2013 .- 103 с. 5
3. Вуколов, Эдуард Александрович. Основы статистического анализа : практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL : учеб. пособие / Э. А. Вуколов .- 2-е изд., испр. и доп .- М. : ФОРУМ, 2012 .- 463 с. 72
4. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов и др. - М.: НИЦ ИНФРА-М, 2015. - 890 с. – Режим доступа: <http://znanium.com/bookread.php?book=515227>
5. Иберла К. Факторный анализ. – М.: Статистика 1980.
6. Мандель И.Д. Кластерный анализ. – М.: Финансы и статистика, 1988.
7. Плюта В. Сравнительный многомерный анализ в экономических исследованиях. – М.: Статистика, 1980.
8. Елисеева И.И., Рукавишников В.О. Группировка, корреляция, распознавание образов. – М.: Статистика, 1978.
9. Сирл С., Госман У. Матричная алгебра в экономике.– М.: Статистика, 1974.
10. Уилкс С. Математическая статистика. - М.: Наука, 1967.
11. Глинский В.В., Ионин В.Г. Статистический анализ: Учебное пособие. – М: «ИНФРА-М», 2002. (УМО).
12. Ионин В.Г. Статистический анализ типовых экономических и социальных процессов на ПЭВМ. Изд. 2. Учебно-методическое пособие – Новосибирск: НГУЭУ, 2005.
13. Юзбашев, Михаил Михайлович. Статистический анализ тенденций и колеблемости / М. М. Юзбашев, А. И. Манелля .- М. : Финансы и статистика, 1983 .- 206, [1] с.
14. Бородкин, Леонид Иосифович. Многомерный статистический анализ в исторических исследованиях : [монография] / Л. И. Бородкин .- М. : Изд-во Моск. ун-та, 1986 .- 187, [1] с.
15. Куликов, Евгений Иванович. Прикладной статистический анализ : учеб. пособие / Е. И. Куликов .- 2-е изд., перераб. и доп .- М. : Горячая линия - Телеком, 2008 .- 463 с.

4.2.2. Информационные источники:

поисковые порталы Yandex, Google, Rambler и др.

4.3. Материально-техническое обеспечение, необходимое для реализации работы по дисциплине

Для изучения учебной дисциплины «Современные технологии анализа массивов данных» необходимо наличие компьютерных классов, достаточного количества рабочих компьютерных мест с выходом в Интернет, средства системы Microsoft Office и полных систем статистического анализа «Statistica», «SPSS» и др.