

Министерство образования и науки Российской Федерации
 Федеральное государственное автономное образовательное учреждение
 высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
 ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Физико-технический институт
 Направление подготовки: Прикладная математика и информатика
 Кафедра Высшей математики и Математической физики

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Многомерный статистический анализ социально-экономической деятельности промышленных предприятий России в разрезе форм собственности

УДК 519.233:658,5:332.012.3

Студент

Группа	ФИО	Подпись	Дата
0BM51	Новосельцева Дарья Андреевна		30.05.17

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент каф. ВММФ	Михальчук А.А.	Кандидат ф-м. наук		30.05.17

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Рыжакина Т.Г.	Кандидат экономических наук		18.05.17

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор	Федорчук Ю.М.	Доктор технических наук		25.05.17

ДОПУСТИТЬ К ЗАЩИТЕ:

Зав. кафедрой	ФИО	Ученая степень, звание	Подпись	Дата
ВММФ	Трифонов А.Ю.	Профессор ф-м. наук		08.06.2017

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Институт Физико-технический
Направление подготовки 01.04.02 Прикладная математика и информатика
Кафедра Высшей математики и математической физики

УТВЕРЖДАЮ:
Зав. кафедрой ВММФ


А. Ю. Трифонов

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

В форме:

магистерской диссертации

Студенту:

Группа	ФИО
0ВМ51	Новосельцевой Дарье Андреевне

Тема работы:

Утверждена приказом проректора-директора (директора) (дата, номер)	§ 2777 /с
--	-----------

Срок сдачи студентом выполненной работы:

30.05.2017


ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	<ul style="list-style-type: none">– Научно-техническая литература по многомерному статистическому анализу;– Научно-техническая литература по построению математических моделей оценивания различий предприятий регионов в РС и ИСС.– Данные статистики по предприятиям регионов России подраздела DM в разрезе форм собственности (Материалы сайта Росстата — URL: http://www.gks.ru/)
Перечень подлежащих исследованию, проектированию и разработке вопросов	<ul style="list-style-type: none">– Анализ научно-технической литературы, формирование базы исходных социально-экономических показателей;– Проведение кластерного, корреляционного и факторного анализов исходных показателей;– Проведение кластерного анализа регионов и выявление различий между кластерами с

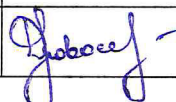
	помощью методов дисперсионного анализа; – Оформление реферата, пояснительной записки.
Консультанты по разделам выпускной квалификационной работы:	
Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Рыжакина Т.Г.
Социальная ответственность	Федорчук Ю.М.
Названия разделов, которые должны быть написаны на иностранном языке:	
ПриложениеА Теоретические основы многомерного статистического анализа	

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику:	16.02.2017
--	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент кафедрой ВММФ	Михальчук А.А.	к.ф.-м.н., доцент		16.02.2017

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0ВМ51	Новосельцева Дарья Андреевна		16.02.2017

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
0BM51	Новосельцевой Дарье Андреевне

Институт	ФТИ	Кафедра	ВММФ
Уровень образования	Магистр	Направление/специальность	Прикладная математика и информатика

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	1. Стоимость расходных материалов 2. Стоимость расхода электроэнергии 3. Норматив заработной платы
2. Нормы и нормативы расходования ресурсов	1. Тариф на электроэнергию 2. Коэффициенты для расчета заработной платы
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	1. Отчисления во внебюджетные фонды (27,1%) 2. Расчет дополнительной заработной платы (12%)

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала, перспективности и альтернатив проведения НИ с позиции ресурсоэффективности и ресурсосбережения	1. Потенциальные потребители результатов исследования; 2. Анализ конкурентных технических решений; 3. SWOT – анализ.
2. Планирование и формирование бюджета научных исследований	1. Структура работ в рамках научного исследования; 2. Определение трудоемкости выполнения работ и разработка графика проведения научного исследования; 3. Бюджет научно - технического исследования (нти).
3. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования	1. Определение интегрального финансового показателя разработки; 2. Определение интегрального показателя ресурсоэффективности разработки; 3. Определение интегрального показателя эффективности

Перечень графического материала (с точным указанием обязательных чертежей):

1. Оценка конкурентоспособности технических решений
2. Матрица SWOT
3. Альтернативы проведения НИ
4. График проведения и бюджет НИ
5. Оценка ресурсной, финансовой и экономической эффективности НИ

Дата выдачи задания для раздела по линейному графику	10.02.17г
--	-----------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Рыжакина Татьяна Гавриловна	Доцент, кандидат экономических наук		10.02.17г

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0BM51	Новосельцева Дарья Андреевна		10.02.17г

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа ОВМ51	ФИО Новосельцевой Дарье Андреевне
-----------------	--------------------------------------

Институт	Физико-технический	Кафедра	Высшей математики и математической физики
Уровень образования	Магистр	Направление/специальность	Прикладная математика и информатика
Тема выпускной квалификационной работы		Многомерный статистический анализ социально-экономической деятельности промышленных предприятий России в разрезе форм собственности	

Исходные данные к разделу «Социальная ответственность»:

<p>1. Описание рабочего места (рабочей зоны, технологического процесса, механического оборудования) на предмет возникновения:</p> <ul style="list-style-type: none"> – вредных проявлений факторов производственной среды (метеоусловия, вредные вещества, освещение, шумы, вибрации, электромагнитные поля, ионизирующие излучения) – опасных проявлений факторов производственной среды (механической природы, термического характера, электрической, пожарной и взрывной природы) – негативного воздействия на окружающую природную среду (атмосферу, гидросферу, литосферу) – чрезвычайных ситуаций (техногенного, стихийного, экологического и социального характера) 	<p>1. Характеристика помещения, где был разработан дипломный проект: ширина, составляет $b = 3$ м, длина комнаты $a = 5$ м, высота $h = 2,8$ м. Площадь помещения составляет 15 м^2, объем 42 м^3. В помещении одно окно, с параметрами: ширина 1,5 м, высота 1,5 м. Количество рабочих мест $n = 1$. Опасные и вредные факторы:</p> <ul style="list-style-type: none"> - микроклимат; -освещенность; -шум; -электростатические поля; -электромагнитные поля; -интенсивная напряженность трудового процесса.
<p>2. Знакомство и отбор законодательных и нормативных документов по теме</p>	<p>2. Часть социальной ответственности ВКР выполнена на основе законодательно-нормативной базы, которая включает в себя государственные стандарты и санитарные правила и нормы о допустимых значениях вредных и опасных факторов на рабочем месте.</p>


Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>1. Анализ выявленных вредных факторов проектируемой производственной среды в следующей последовательности:</p> <ul style="list-style-type: none"> – физико-химическая природа вредности, её связь с разрабатываемой темой; – действие фактора на организм человека; – приведение допустимых норм с необходимой размерностью (со ссылкой на соответствующий нормативно-технический документ); – предлагаемые средства защиты (сначала коллективной защиты, затем – индивидуальные защитные средства) 	<p>1. Анализ опасных и вредных факторов:</p> <ul style="list-style-type: none"> - действие фактора на организм человека; - приведение допустимых норм с необходимой размерностью
<p>2. Анализ выявленных опасных факторов проектируемой произведённой среды в следующей последовательности</p>	<p>2. Производственная безопасность:</p>

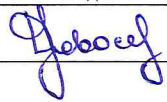
<ul style="list-style-type: none"> - механические опасности (источники, средства защиты); - термические опасности (источники, средства защиты); - электробезопасность (в т.ч. статическое электричество, молниезащита – источники, средства защиты); - пожаровзрывобезопасность (причины, профилактические мероприятия, первичные средства пожаротушения) 	<p>-характеристика помещения, микроклимат.</p> <p>-производственный шум</p> <p>-освещенность рабочей зоны</p> <p>Рабочее место, оценка условий труда на рабочем месте:</p> <ul style="list-style-type: none"> - психофизические факторы - анализ опасных производственных факторов и обоснование мероприятий по их устранению (техника безопасности) - пожарная безопасность
<p>3. Охрана окружающей среды:</p> <ul style="list-style-type: none"> - защита селитебной зоны - анализ воздействия объекта на атмосферу (выбросы); - анализ воздействия объекта на гидросферу (сбросы); - анализ воздействия объекта на литосферу (отходы); - разработать решения по обеспечению экологической безопасности со ссылками на НТД по охране окружающей среды. 	<p>3. Экологическая безопасность</p> <ul style="list-style-type: none"> -перечень отходов: бумага(черновики), орг. техника -методы их обезвреживания и утилизации
<p>4. Защита в чрезвычайных ситуациях:</p> <ul style="list-style-type: none"> - перечень возможных ЧС на объекте; - выбор наиболее типичной ЧС; - разработка превентивных мер по предупреждению ЧС; - разработка мер по повышению устойчивости объекта к данной ЧС; - разработка действий в результате возникшей ЧС и мер по ликвидации её последствий 	<p>4. Защита в чрезвычайных ситуациях:</p> <ul style="list-style-type: none"> - Электробезопасность - Чрезвычайные ситуации природного характера (сильный мороз) - Чрезвычайные ситуации, связанные с проникновением посторонних на производственную территорию
<p>5. Правовые и организационные вопросы обеспечения безопасности:</p> <ul style="list-style-type: none"> - специальные (характерные для проектируемой рабочей зоны) правовые нормы трудового законодательства; - организационные мероприятия при компоновке рабочей зоны 	<p>5. Законодательная и нормативная база</p>
<p>Перечень графического материала:</p>	
<p>При необходимости представить эскизные графические материалы к расчётному заданию (обязательно для специалистов и магистров)</p>	<p>1. План помещения и размещения светильников с люминесцентными лампами</p> <p>2. План эвакуации людей</p>

Дата выдачи задания для раздела по линейному графику	10.03.17г
--	-----------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор	Федорчук Юрий Митрофанович	Профессор, доктор технических наук		10.03.17г

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
ОВМ51	Новосельцева Дарья Андреевна		10.03.17г

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ООП

Код результата	Профессиональные компетенции
	<i>1) В области научно-исследовательской деятельности:</i>
(ПК-1)	Способность проводить научные исследования и получать новые научные и прикладные результаты самостоятельно и в составе научного коллектива
(ПК-2)	Способность разрабатывать концептуальные и теоретические модели решаемых научных проблем и задач
<i>2) В проектной и производственно-технологической:</i>	
(ПК-3)	Способность понимания углубленного анализа проблем, постановки и обоснования задач научной и проектно-технологической деятельности
(ПК-4)	Способность разрабатывать концептуальные и теоретические модели решаемых задач проектной и производственно-технологической деятельности
<i>3) В организационно-управленческой деятельности:</i>	
(ПК-5)	Способность управлять проектами, планировать научно-исследовательскую деятельность, анализировать риски, управлять командой проекта
(ПК-6)	Способность разрабатывать и оптимизировать бизнес-планы научно-прикладных проектов
<i>4) В педагогической деятельности:</i>	
(ПК-7)	Способность к преподаванию математических дисциплин и информатики в образовательных организациях основного общего, среднего общего, среднего профессионального и высшего образования
(ПК-8)	Способность разрабатывать учебно-методические комплексы для электронного и мобильного обучения
<i>5) В консалтинговой деятельности:</i>	
(ПК-9)	Способность разрабатывать аналитические обзоры состояния области прикладной математики и информационных технологий
<i>6) В консорциумной деятельности:</i>	
(ПК-10)	Способность к взаимодействию в рамках международных проектов и сетевых сообществ

РЕФЕРАТ

Выпускная квалификационная работа 114 страниц, 17 рисунков, 39 таблиц, 36 источников, 3 приложения.

Ключевые слова: многомерный статистический анализ, параметрический и непараметрический дисперсионный анализ, уровень значимости, корреляция, кластеризация.

Объектом исследования являются социально-экономические показатели (СЭП) предприятий России подраздела ДМ «Производство транспортных средств и оборудования» в разрезе форм собственности.

Цель работы – исследование деятельности предприятий подраздела «Производство транспортных средств и оборудования» обрабатывающей промышленности России в разрезе форм собственности с использованием методов многомерного статистического анализа.

Актуальность работы заключается в том, что в связи с интенсивными процессами создания предприятий в иностранной и совместной собственности в отдельных отраслях промышленности России возникает потребность сравнения предприятий в российской собственности, иностранной и совместно собственности по СЭП. В частности, наиболее интенсивной отраслью создания таких предприятий является подраздел ДМ «Производство транспортных средств и оборудования» обрабатывающей промышленности России. Объемы производства предприятий в иностранной и совместной собственности в этом подразделе достигли 42% в 2013 году и стали сопоставимыми с предприятиями в российской собственности. В связи с вышесказанным, актуальной темой является проблема сравнительного анализа поведения предприятий подраздела ДМ в разрезе форм собственности, показателей и регионов России.

В процессе исследования проводился многомерный статистический анализ: корреляционный, кластерный, факторный

анализы СЭП, кластерный анализ регионов РС и ИСС, параметрический и непараметрический анализы кластеров регионов.

В результате исследования на основе корреляционных связей построена кластерная модель СЭП, уменьшающая количество 9 исходных показателей до 6-ти; факторная модель показателей, объясняющая общую изменчивость исходных показателей на 99%; кластерные модели регионов в РС и ИСС; получены оценки уровней значимости различий предприятий в РС и ИСС по СЭП.

Для проведения расчетов и построения моделей использованы пакеты MS Excel, Statistica.

Магистерская диссертация выполнена в текстовом редакторе Microsoft Office Word 2010.

ОПРЕДЕЛЕНИЯ

В данной работе применены следующие термины с соответствующими определениями [8]:

Многомерный статистический анализ: раздел математической статистики, посвященный математическим методам построения оптимальных планов сбора, систематизации и обработке многомерных статистических данных, направленным на выявление характера и структуры взаимосвязей между компонентами исследуемого многомерного признака и предназначенных для получения научных и практических выводов.

Дисперсионный анализ: статистический метод, позволяющий анализировать влияние различных факторов (категориальных, группирующих, независимых переменных) на результаты эксперимента (зависимые переменные).

Уровень значимости: это ограничение сверху на вероятность ошибки первого рода (вероятность отвергнуть нулевую гипотезу, когда она на самом деле верна).

Корреляция: статистическая взаимосвязь двух или более случайных величин (либо величин, которые можно с некоторой допустимой степенью точности считать таковыми).

Кластерный анализ: анализ, позволяющий получить разбиение большого объема данных на классы или группы (от англ. cluster=класс) согласно некоторому критерию или их совокупности.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	8
ОБЗОР ЛИТЕРАТУРЫ.....	9
1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА.....	11
1.1. Дисперсионный анализ.....	11
1.2. Факторный анализ.....	20
1.3. Кластерный анализ.....	25
2. МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ПОДРАЗДЕЛА ДМ	30
2.1. Корреляционный анализ показателей ДМ	31
2.2. Факторный анализ показателей ДМ.....	33
2.3. Кластерный анализ регионов	36
2.4. Выводы	53
3. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ.....	56
3.1. Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения	56
3.1.1. Потенциальные потребители результатов исследования	56
3.1.2. Анализ конкурентных технических решений	57
3.1.3. SWOT- анализ.....	59
3.1.4. Оценка готовности проекта к коммерциализации.....	61
3.1.5. Методы коммерциализации результатов научно-технического исследования	62
3.1.6. Инициация проекта.....	64
3.1.7. Организационная структура проекта	65
3.1.8. Ограничения и допущения проекта	65
3.2. Планирование научно - исследовательских работ.....	66
3.2.1. Иерархическая структура работ проекта.....	66
3.2.2. Структура работ в рамках научного исследования	67
3.2.3. Определение трудоемкости выполнения работ и разработка графика проведения научного исследования	68
3.3. Бюджет научно - технического исследования (НТИ)	72
3.3.1. Расчет материальных затрат НТИ.....	72
3.3.2. Основная заработная плата	72
3.3.3. Дополнительная заработная плата исполнителей темы	74
3.3.4. Отчисления во внебюджетные фонды (страховые отчисления)....	74
3.3.5. Расчет затрат на научные и производственные командировки	75
3.3.6. Контрагентные расходы.....	75
3.3.7. Накладные расходы.....	75

3.3.8. Формирование бюджета затрат НТИ.....	76
3.4. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования.....	77
3.4.1. Динамические методы экономической оценки инвестиций	77
3.4.2. Чистая текущая стоимость (NPV)	78
3.4.3. Дисконтированный срок окупаемости.....	79
3.4.4. Внутренняя ставка доходности (IRR)	79
3.4.5. Индекс доходности (рентабельности) инвестиций (PI)	80
3.5. Вывод.....	83
4. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ.....	84
4.1. Анализ опасных и вредных факторов	85
4.2. Производственная безопасность.....	86
4.2.1. Характеристика помещения, микроклимат	86
4.2.2. Характеристика помещения, микроклимат	87
4.2.3. Производственный шум	88
4.2.4. Освещенность рабочей зоны.....	90
4.3. Рабочее место, оценка условий труда на рабочем месте	93
4.3.1. Психофизиологические факторы	95
4.3.2. Анализ опасных производственных факторов и обоснование мероприятий по их устранению (техника безопасности)	96
4.4. Электробезопасность	96
4.5. Экологическая безопасность.....	99
4.6. Пожарная безопасность	101
4.7. Чрезвычайные ситуации	103
4.8. Выводы и рекомендации	105
4.9. Законодательная и нормативная база.....	105
ЗАКЛЮЧЕНИЕ	107
СПИСОК ПУБЛИКАЦИЙ СТУДЕНТА	108
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	112
ПРИЛОЖЕНИЕ А.....	115
ПРИЛОЖЕНИЕ Б.....	134
ПРИЛОЖЕНИЕ В.....	135

ВВЕДЕНИЕ

С 2006 года в Российской экономике наблюдаются высокоинтенсивные процессы создания предприятий в иностранной и совместной собственности в отдельных отраслях промышленности. В частности, наиболее интенсивной отраслью создания таких предприятий является подраздел ДМ «Производство транспортных средств и оборудования» обрабатывающей промышленности России. Объемы производства предприятий в иностранной и совместной собственности в 2013 году достигли 42% и стали сопоставимыми с предприятиями в российской собственности в этом подразделе. Возникает проблема сравнительного анализа поведения предприятий подразделения ДМ в разрезе форм собственности, показателей и регионов России.

Целью настоящей работы является исследование деятельности предприятий подразделения «Производство транспортных средств и оборудования» обрабатывающей промышленности России в разрезе форм собственности с использованием методов многомерного статистического анализа, таких как: корреляционный, кластерный, факторный, параметрический и непараметрический дисперсионный анализы.

Выявление закономерностей процессов развития промышленных предприятий в российской (РС) и иностранной + совместной (ИСС) собственности подразделения ДМ в разрезе регионов России является составной задачей, предполагающей поэтапное решение:

- 1) построение факторного экономического пространства регионов России на основании корреляционного и факторного анализов социально-экономических показателей развития промышленности предприятий подразделения ДМ регионов России;

- 2) проведение статистического сравнительного анализа результативности развития предприятий промышленности регионов в РС и ИСС в рамках дисперсионного анализа соответствующих факторных экономических пространств;

- 3) интерпретация полученных результатов экономическими выводами.

ОБЗОР ЛИТЕРАТУРЫ

Либерализация торговли и интернационализация рынка оказывает значительное влияние на внутренние рынки. Иностранные инвестиции, новые технологии, высокий уровень конкуренции и международная интеграция бизнеса являются основными стимулами развития промышленности. Тем не менее, отечественные и зарубежные предприятия часто оказываются на неравных позициях. Они могут использовать различные бизнес подходы, методы управления и имеют асимметричный доступ к международным рынкам, что естественно создает рыночную дифференциацию [22, 23]. Более того, социальные факторы, такие как различия в корпоративной культуре и бизнес-философии оказывают влияние на деятельность отечественных и иностранных фирм [24].

В некоторых исследованиях рассмотрено негативное влияние этих различий на отечественные предприятия. Байс показывает, что инновационная производительность автомобильной промышленности иностранных южно-африканских предприятий лучше [22]. Беллак утверждает «иностранные фирмы демонстрируют отличные показатели по сравнению со своими коллегами из развитых стран» [24]. Барнес также поддерживают эту точку зрения, заявляя, что отечественные фирмы обычно не могут конкурировать с их зарубежными конкурентами [25]. Определенное положительное влияние иностранного присутствия - это улучшение результатов и увеличение реальной заработной платы в фирмах, преодоление технологического отставания, улучшение взаимодействия между международными и отечественными игроками [26, 27, 28]. Влияние иностранных владельцев бизнеса на различные показатели национального рынка труда было детально изучено в современной экономической науке [29, 30]. Исследователи проанализировали статистическую информацию на уровне компании, региона или страны в основных отраслях производства и услуг товаров.

Результаты исследования показали:

- Иностранные предприятия нанимают самых квалифицированных сотрудников; почти все зарплаты в этих компаниях выше, чем на отечественных предприятиях.

- Высокотехнологичные иностранные конкуренты негативно влияют на производительность труда национальных предприятий и на уровень занятости в стране, но оказывают положительное влияние на уровень качества товаров и наполненность товара на национальном рынке.

С 2006 года прослеживается значительное изменение в структуре сектора производства транспортных средств и оборудования России. Это изменение обусловлено созданием предприятий в иностранной и совместной собственности. Производство транспортных средств и оборудования это подраздел DM согласно Статистической классификации продукции по видам экономической деятельности Европейского Экономического Сообщества. Этот сектор производит около 11% (более чем 2800 миллиардов рублей) в общем объеме производства и составляет 13,5% (более чем 1 миллион занятых) в общем количестве занятого населения обрабатывающей промышленности России. Объемы производства предприятий в иностранной и совместной собственности в этом подразделе достигли 42% в 2013 году и стали сопоставимыми с предприятиями в российской собственности. Однако это резкое увеличение не привело к более быстрому росту объемов производства в подразделе DM в целом. Предприятия в иностранной и совместной собственности ориентированы на определенную часть российского рынка и начинают конкурировать с отечественными предприятиями. Созданные предприятия были ориентированы на фиксированный объем российского рынка и стали конкурировать с российскими предприятиями. В этих условиях возникает проблема сравнительного анализа поведения предприятий подраздела DM в разрезе форм собственности, показателей и регионов России.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА

1.1. Дисперсионный анализ

Дисперсионный анализ [1, 3] – метод статистики, который позволяет анализировать влияние различных факторов (категориальных, группирующих, независимых переменных) на результаты эксперимента (зависимые переменные). Для того, чтобы провести дисперсионный анализ необходимо, чтобы независимая переменная была категориальной, а зависимая – метрической.

Суть дисперсионного анализа (analysis of variance — сокращенно ANOVA) заключается в разложении дисперсии измеряемого признака на независимые слагаемые, из которых каждое характеризует влияние того или иного фактора или их взаимодействия. Последующее сравнение таких слагаемых позволяет оценить значимость каждого изучаемого фактора, а также их комбинации.

Анализ основывается на расчете *F-статистики* (статистика Фишера), которая представляет собой отношение двух *дисперсий*: межгрупповой и внутригрупповой. *F-тест* в однофакторном дисперсионном анализе определяет, значимо ли различаются средние нескольких независимых выборок. Он заменяет *t-тест* для независимых выборок при наличии более двух выборок и дает тот же результат в случае двух выборок.

В случае применения дисперсионного анализа строят математическую модель, т.е. математическое соотношение, представляющее каждую зависимую переменную в виде суммы ошибки и среднего значения. В свою очередь, среднее значение каждого наблюдения представляется в виде суммы генерального среднего и «эффекта» от каждого фактора. Возникающие здесь статистические задачи связаны с оценкой этих эффектов и проверкой статистических гипотез о них. Задача проверки выдвинутых гипотез может быть решена

только при введении дополнительных предположений о вероятностной структуре погрешностей наблюдений. Обычно предполагают, что они независимы и подчиняются нормальному закону с нулевым средним и постоянной дисперсией, что позволяет использовать развитую теорию метода наименьших квадратов (МНК). Менее жёсткие предположения требуют соответственно достаточно большого числа наблюдений, при котором становится оправданным обращение к результатам асимптотической теории.

Классификация моделей дисперсионного анализа основана прежде всего на характере анализируемых факторов: различают модели с фиксированными факторами (модель I), со случайными факторами (модель II).

Техника дисперсионного анализа меняется в зависимости от числа изучаемых независимых факторов.

Рассмотрим *параметрический однофакторный одномерный дисперсионный анализ*.

Однофакторный дисперсионный анализ используется в тех случаях, когда есть в распоряжении три или более независимые выборки (группы), полученные из одной генеральной совокупности путем изменения какого-либо фактора A , имеющего, соответственно, три или более уровня A_i , $i = 1, \dots, k$. Предполагается, что эти выборки распределены по нормальному закону и имеют разные выборочные средние и одинаковые выборочные дисперсии σ^2 . Поэтому необходимо ответить на вопрос, оказал ли этот фактор существенное влияние на разброс выборочных средних группы или разброс является следствием случайностей, вызванных, например, небольшим объемом выборок.

Пусть Y – случайная величина, определенная на этой генеральной совокупности, а μ – ее среднее. В представлении $\mu_i = \mu + \alpha_i$ для среднего величины Y на i -й выборке объема n_i величина α_i есть дифференциальный эффект, соответствующий уровню A_i . Обозначим через y_{ij} значение Y ,

определенное j -м наблюдением ($j = 1, \dots, n_i$) i -ой выборки. Рассматриваемая здесь модель дисперсионного анализа описывает именно такую ситуацию в предположении, что распределение Y на каждой выборке нормально с одной и той же дисперсией σ^2 , т.е. подчинено закону $N(\mu_i, \sigma^2)$. Каждое значение y_{ij} равно сумме генерального среднего μ (единого для всех k уровней фактора), дифференциального эффекта α_i , определяемого уровнем A_i , и случайной ошибки ε_{ij} , вызванной влиянием неконтролируемых факторов:

$$y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij}; \quad (1.1)$$

$$j = 1, \dots, n_i; \quad i = 1, \dots, k,$$

где $\{\varepsilon_{ij}\}$ независимы и распределены по нормальному закону $N(0, \sigma^2)$.

Фактор будет интерпретироваться по модели I, то есть экспериментатора интересуют только выборки, отвечающие именно данным уровням фактора. Модель однофакторного дисперсионного анализа с фиксированными эффектами (модель I) задается соотношением (1.1). Проверяется гипотеза H_0 : все $\alpha_i = 0$. Из представления модели в виде (1.1) можно вывести МНК-оценки $\hat{\mu}_i$ для μ_i , $i = 1, \dots, k$. Из них вытекают МНК-оценки величин μ и α_i . Для того чтобы обеспечить единственность МНК-оценок, нам придется наложить дополнительное ограничение на параметры $\alpha_1, \dots, \alpha_k$. Обычное требование состоит в том, чтобы взвешенная сумма эффектов равнялась нулю:

$$\sum_{i=1}^k n_i \alpha_i = 0.$$

Таким образом, задача сводится к минимизации суммы квадратов

$$S = \sum_{i=1}^k \sum_{j=1}^{n_i} (\varepsilon_{ij})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [y_{ij} - \mu_i]^2$$

по переменным μ_i , $i = 1, \dots, p$.

Основная идея дисперсионного анализа заключена в разложении (1.2): общая вариация переменной, порожденная влиянием фактора и измеренная суммой SS_T , складывается из двух компонент: SS_B и SS_R , характеризующих изменчивость этой переменной между уровнями фактора (SS_B) и изменчивость внутри уровней (SS_R).

$$SS_T = SS_B + SS_R, \quad (1.2)$$

где

$$SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} [y_{ij} - \bar{y}_{**}]^2$$

– общая, или полная, сумма квадратов отклонений;

$$SS_B = \sum_{i=1}^k \sum_{j=1}^{n_i} [\bar{y}_{i*} - \bar{y}_{**}]^2$$

– сумма квадратов отклонений групповых средних от общей средней, или межгрупповая (межуровневая факторная) сумма квадратов отклонений;

$$SS_R = \sum_{i=1}^k \sum_{j=1}^{n_i} [y_{ij} - \bar{y}_{i*}]^2$$

– сумма квадратов отклонений наблюдений от групповых средних, или внутригрупповая (остаточная) сумма квадратов отклонений.

В дисперсионном анализе анализируются не сами суммы квадратов отклонений, а так называемые средние квадраты, которые получаются делением сумм квадратов отклонений на соответствующее число степеней свободы (MS_B , MS_R). Для среднего квадрата MS_B , являющегося несмещенной оценкой межгрупповой дисперсии, число степеней свободы $\nu_B = k - 1$, так как при его расчете используются k групповых средних, связанных между собой одним уравнением. А для среднего квадрата MS_R , являющегося несмещенной оценкой внутригрупповой дисперсии, число степеней свободы $\nu_R = n - k$, т.к. при её расчете используются все n наблюдений, связанных между собой k уравнениями. Таким образом, имеем:

$$MS_B = SS_B / (k - 1); MS_R = SS_R / (n - k).$$

Для проверки гипотезы $H_0: \alpha_1 = \dots = \alpha_k = 0$ о том, что все дифференциальные эффекты равны нулю, вычислим наблюдаемое F -отношение:

$$F_H = MS_B / MS_R,$$

Для того чтобы сделать окончательный вывод, необходимо найти критическое значение $F_{кр}$ по таблице распределения Фишера с учётом количества степеней свободы (v_B и v_R) и соответствующего уровня значимости. Если наблюдаемое значение критерия F_H , вычисленное по данным выборки, принадлежит критической области ($F_H > F_{кр}$), то гипотезу отвергают; если наблюдаемое значение не принадлежит критической области, то нет оснований отвергать гипотезу. Принятие (не отвержение) гипотезы $H_0: \text{все } \alpha_i = 0$ означает справедливость эквивалентной гипотезы, т.е. $H_0: \mu_1 = \dots = \mu_k = \mu$, что все k групповых средних равны генеральному среднему, что практически означает не значимое различие групповых и генерального средних. Результат является значимым, если гипотезу отвергают, поскольку это говорит о наличии существенных различий между средними значениями по группам (уровням фактора).

В случае нескольких независимых групп непараметрическими аналогами (альтернативами) однофакторного дисперсионного анализа являются **ранговый дисперсионный анализ Краскела–Уоллиса**. Критерий Краскела–Уоллиса основан на рангах, а не на исходных наблюдениях и предполагает, что рассматриваемая переменная непрерывна и измерена, как минимум, в порядковой шкале. Критерий проверяет гипотезу: сравниваемые выборки имеют одно и то же распределение.

Рассмотрим теоретические основы критерия Краскела–Уоллиса на примере следующей задачи [5]. Данные состоят из $N = \sum_{j=1}^k n_j$ наблюдений, по n_j наблюдений на j -ю выборку (группу), $j = 1, \dots, k$. Предполагается

следующая исходная математическая модель непараметрического однофакторного дисперсионного анализа:

$$x_{ij} = \mu + \tau_j + \varepsilon_{ij}, \quad i = 1, \dots, n_j, \quad j = 1, \dots, k,$$

где μ – неизвестное общее среднее; τ_j – неизвестный эффект j -ой выборки ($\sum_{j=1}^k \tau_j = 0$). Все ε_{ij} (случайные ошибки) взаимно независимы и извлечены

из одной и той же непрерывной совокупности. По имеющимся предположениям требуется проверить гипотезу: $H_0: \tau_1 = \tau_2 = \dots = \tau_k$. Для этого надо выполнить следующие действия.

1) Проранжировать все N наблюдений вместе от меньшего к большему. Пусть r_{ij} обозначает ранг x_{ij} в этой совместной ранжировке.

2) Положить для $j = 1, \dots, k$

$$R_j = \sum_{i=1}^{n_j} r_{ij}, \quad R_{*j} = \frac{R_j}{n_j}, \quad R_{**} = \frac{N+1}{2}.$$

Например, R_1 – сумма рангов, присвоенных выборке 1, а R_{*1} – средний ранг, который получила эта выборка.

3) Вычислить статистику

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k n_j (R_{*j} - R_{**})^2 = \left[\frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} \right] - 3(N+1).$$

4) На уровне значимости α отклонить H_0 , если $H \geq h(\alpha, k, (n_1, \dots, n_k))$, принять H_0 , если $H < h(\alpha, k, (n_1, \dots, n_k))$, где постоянная $h(\alpha, k, (n_1, \dots, n_k))$ удовлетворяет условию $P\{H \geq h(\alpha, k, (n_1, \dots, n_k))\} = \alpha$. Если гипотеза H_0 верна, то статистика H имеет асимптотическое (при $\min(n_1, \dots, n_k) \rightarrow \infty$) распределение χ^2 с $k-1$ степенями свободы. Приближенный критерий уровня α таков: отклонить H_0 , если $H \geq \chi^2(k-1, \alpha)$, принять H_0 , если $H < \chi^2(k-1, \alpha)$, где $\chi^2(k-1, \alpha)$ – верхняя $\alpha\%$ -ная точка χ^2 -распределения с $k-1$ степенями свободы.

Рассмотрим, наконец, **однофакторный многомерный дисперсионный анализ**.

Пусть $n_r (n = \sum_{r=1}^k n_r)$ – число наблюдений, полученных для r -й выборки из k выборок (групп) для каждой из m переменных Y_1, Y_2, \dots, Y_m . Обозначим через $y_{ij}^{(r)}$ ($i = \overline{1, m}, j = \overline{1, n_r}, r = \overline{1, k}$) значение переменной Y_i , определенное j -м наблюдением r -ой выборки. Тогда

$$\bar{Y}_{ir} = \frac{1}{n_r} \sum_{j=1}^{n_r} y_{ij}^{(r)} \quad (i = \overline{1, m}, r = \overline{1, k})$$

– выборочные (групповые) среднее значение переменной Y_i для r -й выборки, а

$$\bar{Y}_i = \frac{1}{n} \sum_{r=1}^k \sum_{j=1}^{n_r} y_{ij}^{(r)} \quad (i = \overline{1, m})$$

– общее среднее значение для объединенной выборки по каждой Y_i из m переменных. Кроме того, пусть

$$S_{ij}^{(r)} = \sum_{l=1}^{n_r} \sum_{q=1}^{n_r} [y_{il}^{(r)} - \bar{Y}_{ir}] [y_{jq}^{(r)} - \bar{Y}_{jr}] \quad (i, j = \overline{1, m})$$

– значение остаточной суммы квадратов и произведений для r -й выборки с $n_r - 1$ степенями свободы.

Аналогично SS_B в одномерном дисперсионном анализе, определим в многомерном случае

$$B_{ij} = \sum_{r=1}^k n_r \bar{Y}_{ir} \bar{Y}_{jr} - n \bar{Y}_i \bar{Y}_j \quad (i, j = \overline{1, m})$$

как межгрупповую сумму квадратов отклонений, и аналогично SS_R

$$W_{ij} = \sum_{r=1}^k S_{ij}^{(r)} \quad (i, j = \overline{1, m})$$

внутригрупповая (остаточная) сумма квадратов и произведений отклонений.

Матрицы (B_{ij}) и (W_{ij}) формируют соответствующие (межгрупповой и внутригрупповой) источники дисперсии системы m случайных величин

$\{Y_1, Y_2, \dots, Y_m\}$ с соответственно $k-1$ и $n-k$ числами степеней свободы. Эти величины являются основой таблицы многомерного дисперсионного анализа. Тогда многомерный Λ -критерий Уилкса (*лямбда-критерий Уилкса*) для проверки гипотезы о равенстве средних значений для k выборок по совокупности m переменных имеет вид

$$\Lambda = |W|/|W + B|,$$

где $|W|$ и $|B + W|$ – определители матриц (W_{ij}) и $(B_{ij})+(W_{ij})$ соответственно.

Статистика Λ имеет U -распределение с m , $k-1$ и $n-k$ степенями свободы [2]. За исключением специальных случаев, процентилю U -распределения бывает трудно вычислять и поэтому на практике обычно используется одна из двух аппроксимаций. Так, вопрос о том, следует ли отвергать проверяемую гипотезу, можно решить сравнением критерия Бартлетта

$$\chi^2 = -[n - 1 - (m+k)/2] \ln \Lambda$$

с процентилями χ^2 -распределения с $m(k - 1)$ степенями свободы. С другой стороны, можно использовать F -аппроксимацию U , предложенную Рао [10]:

$$F = \frac{[n - 1 - (m + k) / 2] \cdot s - m(k - 1) / 2 + 1}{m(k - 1)} \cdot \frac{1 - \Lambda^{1/s}}{\Lambda^{1/s}}, \quad s = \sqrt{\frac{m^2 (k - 1)^2 - 4}{m^2 + (k - 1)^2 - 5}}.$$

Гипотеза отвергается, если $F > F_{1-\alpha}$ со степенями свободы, определенными числителем и знаменателем первой дроби в выражении F .

Таким образом, если одновременно имеется некоторое число зависимых переменных, то возрастает сложность вычислений, в то время как содержание и базовые принципы не меняются. Вместо одномерного F -критерия используется многомерный F -критерий (*лямбда-критерий Уилкса*), основанный на сравнении ковариационной матрицы ошибок и межгрупповой ковариационной матрицы. Если общий многомерный критерий значим, можно заключить, что соответствующий эффект значим

по совокупности всех переменных. Иными словами, если гипотеза о равенстве средних по совокупности всех переменных отвергается на основании Λ -критерия, то далее проверяется гипотеза о равенстве средних по каждой переменной на основании одномерного F -критерия. Если и она отвергается, то мы делаем вывод, что некоторые μ не совпадают. Так как F -критерий не дает информации о том, какие именно из средних не равны, необходимо провести **множественный анализ средних**.

Выделяют три процедуры множественного сравнения:

1) метод Шеффе – для проверки гипотезы $H_0: \sum_{r=1}^k c_r \mu_r = 0$ против альтернативы $H_1: \sum_{r=1}^k c_r \mu_r \neq 0$ с уровнем значимости α образуют следующий $100(1 - \alpha)\%$ -ный доверительный интервал:

$$\sum_{r=1}^k c_r \bar{y}_{r*} \pm \sqrt{k \cdot MS_R \cdot F_{1-\alpha}(k, n-k) \sum_{r=1}^k \frac{c_r^2}{n_r}},$$

где MS_R – средний квадрат, является несмещенной оценкой внутригрупповой дисперсии с числом степеней свободы $\nu_R = n - k$ берется из таблицы однофакторного дисперсионного анализа, а $F_{1-\alpha}(k, n - k)$ есть $100(1 - \alpha)\%$ -я процентиль распределения $F(k, n - k)$. Если этот интервал не содержит нуль, то H_0 отвергается с уровнем α . Этот процесс повторяется для каждой интересующей нас линейной комбинации, причем общим для всех критериев уровнем значимости остается α .

2) метод Тьюки, который применим только для контрастов и только в случае равных объемов выборок, т.е. при $n_1 = n_2 = \dots = n_k = n/k$. Для проверки гипотезы $H_0: \sum_{r=1}^k \lambda_r \mu_r = 0$ против альтернативы $H_1: \sum_{r=1}^k \lambda_r \mu_r \neq 0$ образуют следующий доверительный интервал:

$$\sum_{r=1}^k \lambda_r \bar{y}_{r*} \pm \frac{1}{2} \sqrt{\frac{k \cdot MS_R}{n}} \cdot q_{1-\alpha} \sum_{r=1}^k |\lambda_r|,$$

а $q_{1-\alpha}$ есть $100(1 - \alpha)\%$ -я процентиль распределения студентизованного размаха с k и $v = n - k$ степенями свободы. Если этот интервал не содержит нуль, то H_0 отвергается с уровнем значимости α .

3) t -метод. Пусть k_0 — число заранее выбранных контрастов. Тогда для проверки гипотезы $H_0: \sum_{r=1}^k \lambda_r \mu_r = 0$ против альтернативы $H_1: \sum_{r=1}^k \lambda_r \mu_r \neq 0$

следует построить доверительный интервал:

$$\sum_{r=1}^k \lambda_r \bar{y}_{r*} \pm t_{1-\alpha/2k_0}(v_R) \sqrt{MS_R \sum_{r=1}^k \frac{\lambda_r^2}{n_r}},$$

а $t_{1-\alpha/2k_0}$ есть $100(1 - \alpha/2k_0)\%$ -я процентиль t -распределения v_R степенями свободы. Если этот интервал не содержит нуль, то мы должны отвергнуть H_0 .

1.2. Факторный анализ

Если объект (регион) описывается с помощью n признаков (СЭП), то в результате действия метода получается факторная модель, зависящая от меньшего числа переменных [13, 14, 18]. При этом предполагается, что на исходные измеряемые данные $\{X_1, X_2, \dots, X_n\}$ оказывает влияние небольшое число латентных (скрытых) признаков (факторов) $\{U_1, \dots, U_m\}$, $m < n$. Цель факторного анализа заключается в выявлении этих скрытых характеристик (факторов) и оценивании их числа.

Факторный анализ используется как метод редукции данных (сокращения) или как метод классификации:

Факторный анализ как метод редукции данных основан на использовании зависимости между переменными: вводится новая переменная на основе регрессии, то есть включающая в себя наиболее существенные черты исходных переменных, так что ее использование (замена нескольких старых коррелированных переменных одной новой) приводит к сокращению числа переменных. При этом новый фактор

(переменная) представляется линейной комбинацией исходных переменных. После выделения первого фактора, то есть построения первой линии регрессии, для которой дисперсия максимальна, определяется следующая линия, максимизирующая остаточную вариацию (разброс данных вокруг первой прямой), то есть выделяется второй фактор, и т.д. Факторы выделяются один за другим так, чтобы максимизировать изменчивость, оставшуюся от предыдущих, то есть оказываются независимыми друг от друга. Другими словами, некоррелированными или *ортогональными*.

Решение о том, когда надо прекратить процедуру выделения факторов, в основном зависит от того, что считается малой «случайной» изменчивостью. Это решение достаточно произвольно, однако имеются некоторые общие рекомендации, позволяющие рационально выбрать число факторов: критерий накопленной или кумулятивной дисперсии, критерий каменистой осыпи. Поэтому обычно исследуется несколько решений с большим или меньшим числом факторов, и затем выбирается одно наиболее «осмысленное».

Факторный анализ как метод классификации основан на оценках корреляций (факторных нагрузок) между исходными переменными и факторами в рамках выбранной факторной модели и позволяет узнать значимость факторов, то есть можно ли интерпретировать их разумным образом и как это сделать. Можно развернуть оси в любом направлении, не изменяя относительного положения точек; однако действительные координаты точек, то есть факторные нагрузки, должны, без сомнения, меняться. Целью вращения является получение понятной (интерпретируемой) матрицы нагрузок, то есть факторов, которые ясно отмечены высокими нагрузками для некоторых переменных и низкими – для других, что и позволяет провести классификацию переменных.

Факторный анализ, как раздел многомерного статистического анализа, объединяет математико-статистические методы решения задач, связанных с построением линейной модели:

$$X=AU + \varepsilon, \quad (1.3)$$

где $X=\{X_1, X_2, \dots, X_n\}^T$ – исходная наблюдаемая n -мерная случайная величина с нулевыми математическими ожиданиями $M[X] = 0$ и корреляционной матрицей $C=M[XX^T]$;

$A=(a_{ij})$ – неизвестная $(n \times m)$ -матрица нагрузок общих факторов на наблюдаемые величины;

$U=\{U_1, \dots, U_m\}^T$ – ненаблюдаемая многомерная случайная величина общих факторов с нулевыми математическими ожиданиями $M[U] = 0$ и единичной корреляционной матрицей $M[UU^T]=E$;

$\varepsilon=\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}^T$ – n -мерная случайная величина ошибок с $M[\varepsilon] = 0$, $M[\varepsilon U^T]=0$ и неизвестной диагональной ковариационной $(D[\varepsilon_i] = \sigma_i^2)$ матрицей $\Psi, = M[\varepsilon\varepsilon^T]$.

Из модели факторного анализа следует, что

$$C=M[XX^T] = M[(AU + \varepsilon)(AU + \varepsilon)^T] = AM[UU^T] A^T + M[\varepsilon\varepsilon^T] = AA^T + \Psi,$$

то есть, корреляции c_{ij} между исходными случайными величинами связаны с факторными нагрузками a_{ij} и дисперсиями ошибок σ_i^2 с помощью соотношений:

$$c_{ij} = \sum_{l=1}^m a_{il} a_{jl}, \quad 1 = c_{ii} = \sum_{l=1}^m a_{il}^2 + \sigma_i^2 = h_i^2 + \sigma_i^2, \quad (1.4)$$

где $h_i^2 = \sum_{l=1}^m a_{il}^2$ называется общностью исходной случайной величины X_i и определяется суммой квадратов факторных нагрузок, то есть вкладом всех m общих факторов в дисперсию величины X_i .

Проведем дальнейшее рассмотрение в рамках метода главных факторов, под которым понимают приложение метода главных компонент к редуцированной корреляционной матрице C_h в случае линейной модели:

$$X = AU \text{ или } X_i = \sum_{j=1}^m a_{ij} U_j \quad (i = \overline{1, n}); \quad m < n, \quad (1.5)$$

где $X = \{X_1, X_2, \dots, X_n\}^T$ с $M[X] = 0$ и корреляционной матрицей C_h .

Представим матрицу нагрузок общих факторов в виде

$$A = (a_{ij}) = \{\alpha_1, \alpha_2, \dots, \alpha_k, \dots, \alpha_m\},$$

где $\alpha_k = \{a_{1k}, a_{2k}, \dots, a_{nk}\}^T$ – k -ый столбец матрицы факторных нагрузок A . С другой стороны,

$$A = (a_{ij}) = \{\beta_1, \beta_2, \dots, \beta_k, \dots, \beta_n\}^T,$$

где $\beta_k = \{a_{k1}, a_{k2}, \dots, a_{km}\}$ – k -ая строка матрицы A . Тогда линейную модель (1.5) можно представить в виде:

$$X_k = \beta_k U \quad (k = \overline{1, n}) \quad (1.6)$$

Согласно методу главных компонент, общий фактор U_k должен вносить максимальный вклад в суммарную общность:

$$\sum_{l=1}^n a_{lk}^2 = \alpha_k^T \alpha_k \rightarrow \max$$

при условии $AA^T = C_h$, выраженном через β_k :

$$c_{ij} = \sum_{l=1}^m a_{il} a_{jl} = \beta_i \beta_j^T, \quad (1.7)$$

где $c_{ij} = c_{ji}$ и $c_{ii} = h_i^2$

Данная задача условного экстремума решается с помощью метода неопределенных множителей Лагранжа:

$$L = \alpha_k^T \alpha_k - \sum_{i,j=1}^n \mu_{ij} (\beta_i \beta_j^T - c_{ij}),$$

где $\mu_{ij} = \mu_{ji}$ – множители Лагранжа.

Находя частные производные от функции Лагранжа по компонентам вектора α_k^T и приравнивая их нулю, получаем систему n линейных однородных алгебраических уравнений, которую с учетом

$$\text{обозначения } \sum_{l=1}^n a_{lk}^2 = \alpha_k^T \alpha_k = \lambda_k$$

после ряда алгебраических преобразований удастся представить независимо от k в матричной форме:

$$(C_h - \lambda_k E) \alpha_k = 0.$$

Последнее равенство определяет задачу на собственные значения (числа) λ_k и собственные вектора α_k редуцированной корреляционной матрицы C_h .

Условие нетривиальности решения задачи на собственные вектора (вырожденность соответствующей матрицы)

$$|C_h - \lambda_k E| = 0$$

приводит к алгебраическому уравнению n -го порядка относительно собственного значения. Таким образом, получаем n упорядоченных собственных чисел $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq \dots \geq \lambda_n$. При этом каждому собственному значению (числу) λ_k ($k = \overline{1, n}$) соответствует собственный вектор α_k – k -ый столбец матрицы факторных нагрузок A , соответствующий фактору U_k , то есть столбец коэффициентов при факторе U_k в (1.5). С учетом упорядоченности собственных чисел и нормировки соответствующих собственных векторов ($\alpha_k^T \alpha_k = \lambda_k$) удается ранжировать факторы по их вкладу в суммарную общность.

На основе исходной наблюдаемой величины X и рассчитанной матрицы нагрузок общих факторов A возможно получить оценки общих факторов U . Например, в рамках метода главных компонент ($m=n$) матрица A будет квадратной и разрешение матричного уравнения (1.5) относительно U не вызывает затруднений, если ранг A равен n , то есть существует A^{-1} – обратная к A матрица:

$$U = A^{-1} X.$$

В рамках метода главных факторов ($m < n$) прямоугольная матрица A не имеет обратной и процедура оценки U усложняется:

$$U = (A^T A)^{-1} A^T X,$$

где матрица $A^T A$ уже является квадратной порядка m .

По полученным оценкам U можно судить о каждом объекте наблюдения по m общим факторам.

Для проверки значимости построенной модели факторного анализа можно воспользоваться критерием Бартлетта по проверке нулевой гипотезы H_0 о том, что m общих факторов достаточно для объяснения выборочных коэффициентов корреляции. Наблюдаемое статистическое значение критерия вычисляют по формуле

$$\chi^2_H = \left(N - \frac{2n+5}{6} - \frac{2}{3}m \right) \ln \frac{|AA^T|}{|C|},$$

где N – число наблюдений (объектов исследования),
 $|AA^T|$ – определитель воспроизведенной из модели матрицы корреляций,

$|C|$ – определитель исходной корреляционной матрицы

При заданном уровне значимости α H_0 отвергают, если $\chi^2_H > \chi^2_{кр}$, где $\chi^2_{кр}$ определяется с помощью калькулятора распределения

вероятности по закону χ^2 с $\nu = \frac{(n-m)^2 - n - m}{2}$ степенями

свободы так, чтобы $P(\chi^2_H > \chi^2_{кр}) = \alpha$. Это значит, что m общих факторов не достаточно и следует выделить хотя бы $m+1$ фактор. После этого процедура проверки повторяется.

Далее, на основе результатов факторного анализа, можно провести, например, классификацию объектов наблюдений по общим факторам, число которых значительно меньше числа исходных показателей.

На основе исходной наблюдаемой величины X и рассчитанной матрицы нагрузок общих факторов A возможно получить оценки общих факторов U .

По полученным оценкам U можно судить о каждом объекте наблюдения по m общим факторам.

1.3. Кластерный анализ

Кластерный анализ [12, 17] – это статистический анализ, позволяющий получить разбиение большого объема данных на классы

или группы (от англ. *cluster*– класс) согласно некоторому критерию или их совокупности.

Для проведения классификации данных X_1, \dots, X_n используют понятие метрики или расстояния. В теории кластерного анализа используются различные метрики для измерения расстояния между отдельными точками (векторами), вот некоторые из них:

1) евклидово расстояние

$$\rho^2(X^{(i)}, X^{(j)}) = \sum_{k=1}^m (X_k^{(i)} - X_k^{(j)})^2;$$

где w_k – веса, пропорциональные важности признака в задаче классификации. Веса задают после проведения дополнительных исследований и полагают, что $\sum_{k=1}^m w_k = 1$;

2) расстояние городских кварталов (или *city-block*) – расстояние по карте между кварталами в городе

$$\rho(X^{(i)}, X^{(j)}) = \sum_{k=1}^m |X_k^{(i)} - X_k^{(j)}|;$$

3) расстояние Чебышева

$$\rho(X^{(i)}, X^{(j)}) = \max \left\{ |X_k^{(i)} - X_k^{(j)}| \right\}$$

4) корреляционное расстояние ($1 - r$ Пирсона). Корреляционное расстояние вычисляется по формуле $(1 - r)$, где r Коэффициент корреляции Пирсона, его называют также линейной корреляцией, из-за того, что он измеряет степень линейных связей между переменными. Данный тип расстояния различается со всеми другими типами расстояний использованием методов корреляционного анализа для оценки расстояний между кластерами.

В ряде алгоритмов наряду с расстояниями между точками (векторами) используются расстояниями между кластерами и объединениями кластеров (расстояние между кластерами по принципу

«ближнего соседа», «дальнего соседа», центрами тяжести групп, по принципу «средней связи», расстояние по Колмогорову).

Методы кластерного анализа подразделяются на агломеративные (объединяющие), дивизимные (разделяющие) и итеративные.

Первые последовательно объединяют отдельные объекты в кластеры, вторые, наоборот, расчлняют кластеры на объекты. Третьи объединяют первые два. Их особенностью является формирование кластеров, исходя из условий разбиения (так называемых параметров), которые могут быть изменены в процессе работы алгоритма для улучшения качества разбиения. Итеративные методы обычно используются для классификации больших объемов информации.

Агломеративные методы наиболее простые и распространенные среди алгоритмов кластерного анализа. На первом шаге каждый вектор или объект X_1, \dots, X_n исходных данных рассматривается как отдельный кластер или класс. По вычисленной матрице расстояний выбираются наиболее близкие друг к другу и объединяются. Очевидно, что процесс завершится через $(n - 1)$ шаг, когда в результате все объекты будут объединены в один кластер.

Последовательность объединений можно представить в виде дендрограммы, или дерева (рис. 1.1).

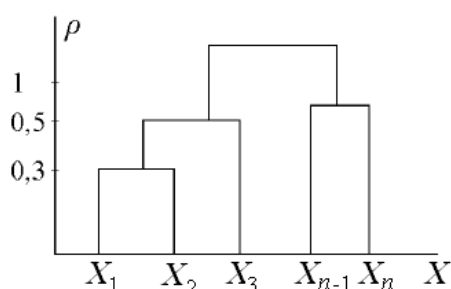


Рисунок 1.1 – Дендограмма

К агломеративным относят методы полной, средней связи, и метод Варда (Уорда).

1) **Метод одиночной связи.** Пусть ξ_1, \dots, ξ_n — данные, причем каждый вектор образует один кластер. Затем вычисляется матрица расстояний

между этими кластерами. Используя эту матрицу, выбираются два наиболее близких вектора, которые и образуют первый кластер S_1 . На следующем шаге между S_1 и оставшимися векторами (которые мы считаем кластерами) вычисляется следующая матрица расстояний. Ближайший к полученному ранее классу S_1 кластер объединяется с ним, образуя S_2 и т.д.

2) **Метод полной связи.** Метод практически повторяет метод одиночной связи за исключением того, что включение нового объекта в кластер происходит тогда и только тогда, когда расстояние между объектами (векторами или кластерами) меньше некоторого наперед заданного числа. Число задается пользователем.

3) **Метод средней связи.** Как и в методе полной связи, все найденные между кластерами расстояния сравниваются с задаваемым пользователем числом. И если оно меньше заданного числа, новый объект включается в старый класс. Таким образом, метод средней связи отличается от метода полной связи только способом вычисления расстояния между кластерами.

4) **Метод ВАРДА (УОРДА).** В данном методе используются методы дисперсионного анализа для оценки расстояний между кластерами. В нем минимизируется сумма квадратов (SS) для любых двух (гипотетических) кластеров, которые могут быть сформированы на каждом шаге. В целом метод стремится создавать кластеры малого размера.

К итеративным методам кластеризации относится метод k - средних.

1) **Метод k -средних.** Пусть имеются векторы $\xi_1, \dots, \xi_n \in R^p$, которые необходимо разбить на k кластеров. На нулевом шаге из n векторов случайным образом выбираем k из них, считая, что каждый образует один кластер. Получаем множество кластеров-эталонов $e_1^{(0)}, \dots, e_k^{(0)}$ с весами $w_1^{(0)}, \dots, w_k^{(0)}$, определяющими число элементов в них. Нуль сверху обозначает номер итерации. На этом этапе все веса равны единице. На следующем шаге из оставшегося набора данных выбираем некоторый

вектор, например ξ_i , и вычисляем матрицу расстояний между ξ_i и эталонами $e_1^{(0)}, \dots, e_k^{(0)}$ по некоторой метрике, например по евклидовой:

$$\rho^2(\xi^{(i)}, e_m^{(0)}) = \sum_{j=1}^p (\xi_j^{(i)} - \xi_{m,j}^{(0)})^2, m = \overline{1, k}$$

На основе знания вычисленной матрицы расстояний вектор ξ_i помещается в тот эталон, расстояние до которого минимально. Допустим для определенности, что это $e_m^{(0)}$. Он заменяется новым, пересчитанным с учетом присоединенной точки, по формуле

$$e_m^{(1)} = \begin{cases} \frac{w_m^{(0)} e_m^{(0)} + \xi_i}{w_m^{(0)} + 1}, & \xi_i \text{ включен} \\ e_m^{(0)}, & \xi_i \text{ не включен} \end{cases}$$

Кроме того, пересчитывается и вес:

$$w_m^{(1)} = \begin{cases} w_m^{(0)} + 1, & \xi_i \text{ включен} \\ w_m^{(0)}, & \xi_i \text{ не включен} \end{cases}$$

Если в матрице встречается два или более минимальных расстояния, то ξ_i включают в кластер с наименьшим порядковым номером. На следующем шаге выбирают следующий вектор из оставшихся, и процедура повторяется. Таким образом, через $(n - k)$ шагов каждому эталону $e_m^{(n-k)}$ будет соответствовать вес $w_m^{(n-k)}$ и процедура кластеризации завершится.

Заметим, что задача кластеризации не имеет единственного решения. Как следствие, перебрать все допустимые разбиения данных на классы достаточно сложно и не всегда представляется возможным. Для того чтобы оценить качество различных способов кластеризации вводят понятие функционала качества разбиения, который принимает минимальное значение на наилучшем (с точки зрения исследователя) разбиении.

2. МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ПОДРАЗДЕЛА ДМ

В данном исследовании использованы методы многомерного статистического анализа. Корреляционный анализ СЭП и построение факторного экономического пространства СЭП позволили выявить взаимосвязи показателей, сформировать на их основе факторные корреляционно независимые показатели и определить различия их значений между предприятиями регионов России в разрезе форм собственности: российской (РС), иностранной и совместной (ИСС). Кластерный анализ позволил сформировать типовые кластеры регионов в разрезе форм собственности и провести сравнение их факторных показателей.

Объект исследования: предприятия российской, иностранной и совместной собственности регионов России подраздела ДМ «Производство транспортных средств и оборудования».

Период исследования: 2010 - 2013 гг.

Информационная база анализа: данные статистики по предприятиям регионов России подраздела ДМ в разрезе форм собственности [21]. Таким образом, созданная база данных далее использовалась в пакете Statistica для многомерного статистического анализа данных.

На первом этапе исследования социально - экономической деятельности предприятий в разрезе форм собственности (Российская собственность (РС) – 58 регионов; иностранная собственность + совместная собственность (ИСС) – 18 регионов) была сформирована система 9-ти социально-экономических показателей (таблица 2.1).

Таблица 2.1– Система социально-экономических показателей

Номер показателя	Показатель
	<i>Абсолютные показатели</i>

№1	Отгруженная продукция, среднее 2010-2013
№2	Зарплата (средняя за месяц)
№3	Труд (занятые)
№4	ФОТ (годовой фонд оплаты труда)
	Расчетные (относительные) показатели
№5	Труд (занятые) / отгрузка
№6	ФОТ / отгрузка
	Темпы роста (среднегодовые темпы роста)
№7	Зарплата (средняя заработная плата за год)
№8	Труд (количество занятого населения)
№9	ФОТ (годовой фонд оплаты труда)

По четырем абсолютным показателям использовались средние значения за 2010-2013 гг. Относительные показатели рассчитаны на основе средних значений этих показателей. Среднегодовые темпы роста рассчитаны на основании среднегодовых значений абсолютных показателей за 2010 и 2013 года. Критерием включения региона в исследование являлся среднегодовой объем отгруженной продукции, который по соответствующей форме собственности должен превышать 1 млрд. руб. При этом отдельные регионы с нетипичным поведением (аномальными значениями) предприятий были исключены из исследования.

2.1. Корреляционный анализ показателей ДМ

С помощью методов корреляционного анализа – коэффициентов парных корреляций: параметрического Пирсона и непараметрического Спирмена, были выявлены статистические взаимосвязи между показателями [2].

Таблица 2.2–Матрицы коэффициентов парных корреляций СЭП (Пирсона (r) – над диагональю в верхнем правом треугольнике и Спирмена (R) –под диагональю в нижнем левом треугольнике).

Показатель	№1	№2	№3	№4	№5	№6	№7	№8	№9
№1	1.00	0.51	0.67	0.73	-0.41	-0.28	-0.09	0.31	0.23
№2	0.35	1.00	0.10	0.27	-0.34	0.04	-0.24	0.51	0.37
№3	0.86	0.21	1.00	0.96	-0.06	0.03	0.11	-0.15	-0.11

№4	0.90	0.43	0.97	1.00	-0.12	0.03	0.05	-0.09	-0.07
№5	-0.46	-0.35	-0.01	-0.11	1.00	0.86	0.17	-0.44	-0.34
№6	-0.27	0.12	0.15	0.15	0.83	1.00	0.10	-0.31	-0.24
№7	0.03	-0.19	0.15	0.08	0.12	0.09	1.00	-0.27	0.26
№8	0.01	0.30	-0.24	-0.14	-0.45	-0.32	-0.04	1.00	0.84
№9	0.04	0.28	-0.16	-0.06	-0.39	-0.27	0.29	0.92	1.00

Согласно критериям Спирмена и Пирсона выявлено высоко значимые (на уровне значимости $p < 0,0005$) положительные корреляционные связи между показателями №3 и №4 (коэффициент Пирсона $r = 0,96$, коэффициент Спирмена $R = 0,97$), показателями №5 и №6 (коэффициент Пирсона $r = 0,86$, коэффициент Спирмена $R = 0,83$) и показателями №8 и №9 (коэффициент Пирсона $r = 0,84$, коэффициент Спирмена $R = 0,92$), а также между показателями №1 и №3, №1 и №4.

На основе кластерного анализа системы 9-ти исходных показателей построена дендрограмма, позволяющая выделять разные группы корреляционно связанных показателей в зависимости от выбранного расстояния объединения d : $d=0,3$ соответствует 6 групп корреляционно связанных показателей F1-F6 (рис. 2.1). Поскольку корреляционный анализ исходных показателей показал значимые корреляционные связи разных пар показателей, то, при проведении их кластеризации целесообразно было использовать корреляционное расстояние в качестве меры близости показателей. В качестве правила объединения двух кластеров использован метод Варда (Уорда). Данный метод отличается от всех других методов использованием методов дисперсионного анализа для оценки расстояний между кластерами.

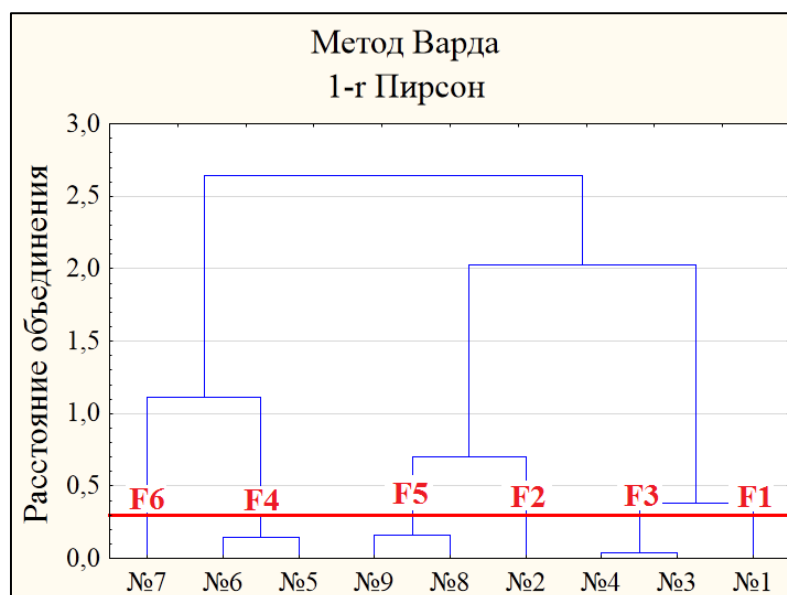


Рисунок 2.1 – Вертикальная дендрограмма корреляционной матрицы показателей

2.2. Факторный анализ показателей ДМ

Наличие корреляционной зависимости исходных показателей допускает использование факторного анализа. Основными целями факторного анализа являются редукция данных числа показателей (сокращение числа показателей) и определение структуры взаимосвязей между показателями, т.е. классификация показателей. Факторный анализ как метод классификации основывается на оценках корреляций (факторных нагрузок) между исходными показателями и факторами (или "новыми" показателями) в рамках выбранной факторной модели и позволяет определить значимость факторов. Цель состоит в том, чтобы посредством относительно небольшого числа факторов воспроизвести большую часть дисперсии показателей.

Графический метод («критерию каменистой осыпи») позволяет определить оценки значимости собственных значений: справа от критической точки, где убывание собственных значений слева направо максимально замедляется, находится только «факториальная осыпь»,

которой пренебрегают. Согласно графическому методу можно оставить 6 факторов (рис. 2.2).

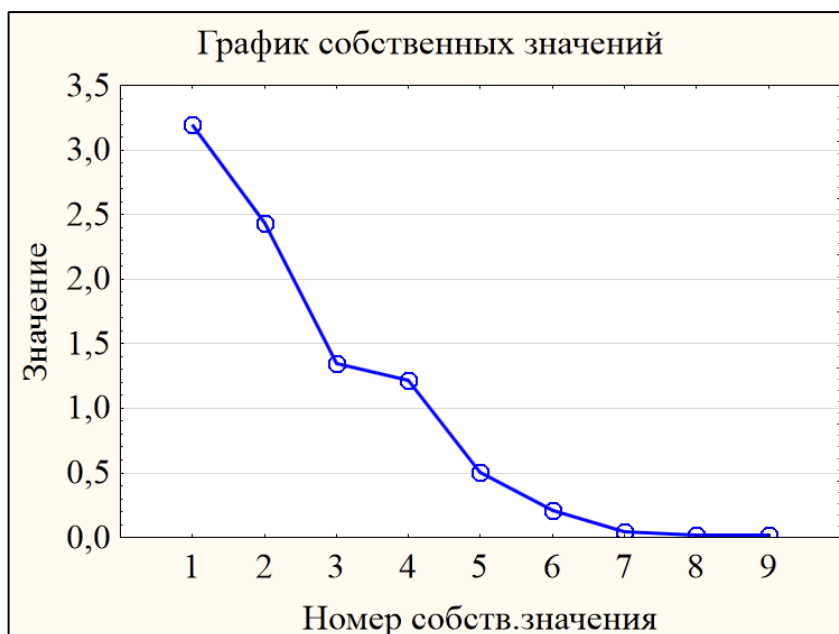


Рисунок 2.2– График собственных значений

С помощью метода главных компонент получена 6-мерная факторная модель экономического пространства (табл. 2.3).

Таблица 2.3–6-мерная факторная модель экономического пространства

Переменная	F1	F2	F3	F4	F5	F6
№1	0.536	0.311	0.711	-0.256	0.195	-0.066
№2	0.065	0.928	0.158	-0.041	0.289	-0.140
№3	0.008	-0.034	0.989	0.009	-0.081	0.058
№4	0.003	0.149	0.980	-0.009	-0.067	0.023
№5	0.060	-0.251	-0.100	0.927	-0.202	0.075
№6	-0.137	0.145	0.016	0.960	-0.133	0.038
№7	-0.017	-0.119	0.048	0.079	0.025	0.987
№8	0.066	0.218	-0.057	-0.205	0.911	-0.255
№9	0.013	0.143	-0.042	-0.151	0.935	0.271
Общ. дисп.	0.318	1.147	2.487	1.920	1.895	1.147
Доля общ.	0.035	0.127	0.276	0.213	0.211	0.127

В таблице 2.3 красным выделены более значимые (основные) повернутые факторные нагрузки (частные коэффициенты корреляции) показателей на факторы. Это позволяет по совокупности показателей интерпретировать соответствующие факторы, приписывая им наиболее

существенные черты значимых показателей. В нижней строчке приведены доли объясненной данным фактором дисперсии исходных показателей, иными словами, весовые коэффициенты факторов. Накопленная дисперсия первыми 6-ю факторами $\approx 99\%$.

Согласно табл. 2.3, высокие факторные нагрузки показателей распределились по факторам, имеющим наибольшие веса, следующим образом:

Фактор F1 ($\approx 0,04$), характеризуется положительной корреляционной связью с абсолютным показателем №1 и интерпретируется как фактор объёма отгруженной продукции;

Фактор F2 ($\approx 0,13$), характеризуется положительной корреляционной связью с абсолютным показателем №2 и интерпретируется как фактор размера средней заработной платы;

Фактор F3 ($\approx 0,28$), характеризуется положительной корреляционной связью с абсолютными показателями №1, №3 и №4 и интерпретируется как фактор абсолютных показателей;

Фактор F4 ($\approx 0,21$), характеризуется положительной корреляционной связью с расчетными показателями №5 и №6 и интерпретируется как фактор расчетных показателей;

Фактор F5 ($\approx 0,21$), характеризуется положительной корреляционной связью с показателями темпов роста №8-№9 и интерпретируется как фактор темпов роста ФОТа и количества занятого населения;

Фактор F6 ($\approx 0,13$), характеризуется положительной корреляционной связью с показателем темпа роста №7 и интерпретируется как фактор темпов роста средней заработной платы.

В дальнейшем, для проведения кластеризации были получены значения факторных переменных.

2.3. Кластерный анализ регионов

Кластеризации регионов проводилась по важнейшим факторам, характеризующих социально-экономическую деятельность предприятий, а также их развитие. В качестве таких факторов были выбраны:

- F3 и F5 – первый вариант кластеризации – отражает размеры Занятого населения, ФОТа и их темпы роста, позволяет выделить регионы с наиболее большим количеством занятого населения, а также изменение данного показателя;

- F1 и F3 – второй вариант кластеризации – отражает размеры отгруженной продукции, ФОТа и количества занятого населения, что позволяет определить регионы с наиболее высокими заработной платой и количеством работников, а также эффективность их работ;

- F2 и F6 – третий вариант кластеризации – отражает размеры ЗП и её темпы роста, позволяет выделить регионы с наиболее высокой средней заработной платой, а также отследить её изменение.

Кластеризация F3 и F5.

Методом К-средних и методами иерархической классификации проведена отдельная (по РС и ИСС) кластеризация регионов. Разными методами кластеризации получены близкие результаты на уровне 6-ти кластерной модели 58-ти регионов РС (рис.2.3) и 6-ти кластерной модели 18-ти регионов ИСС (рис.2.4).

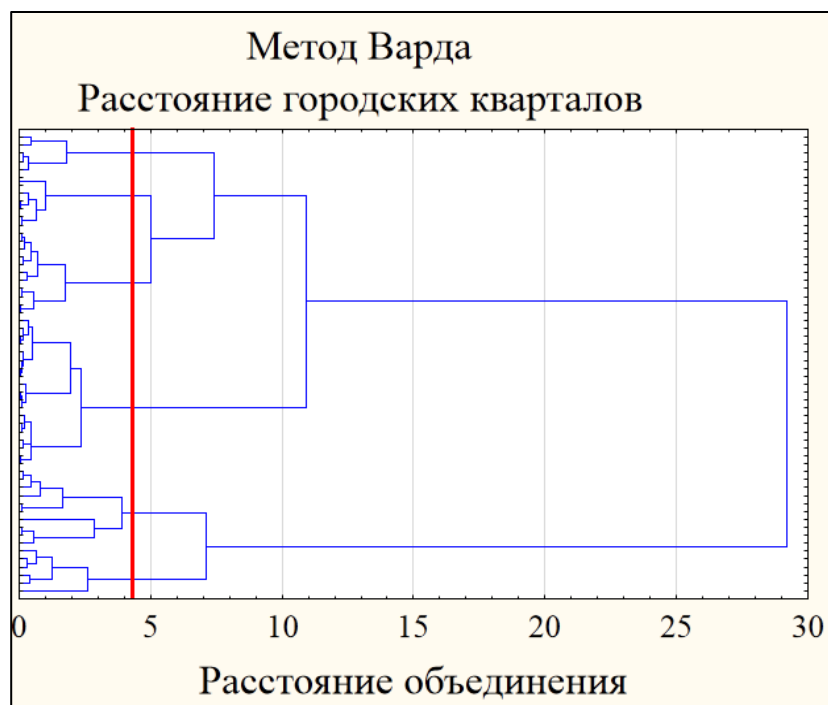
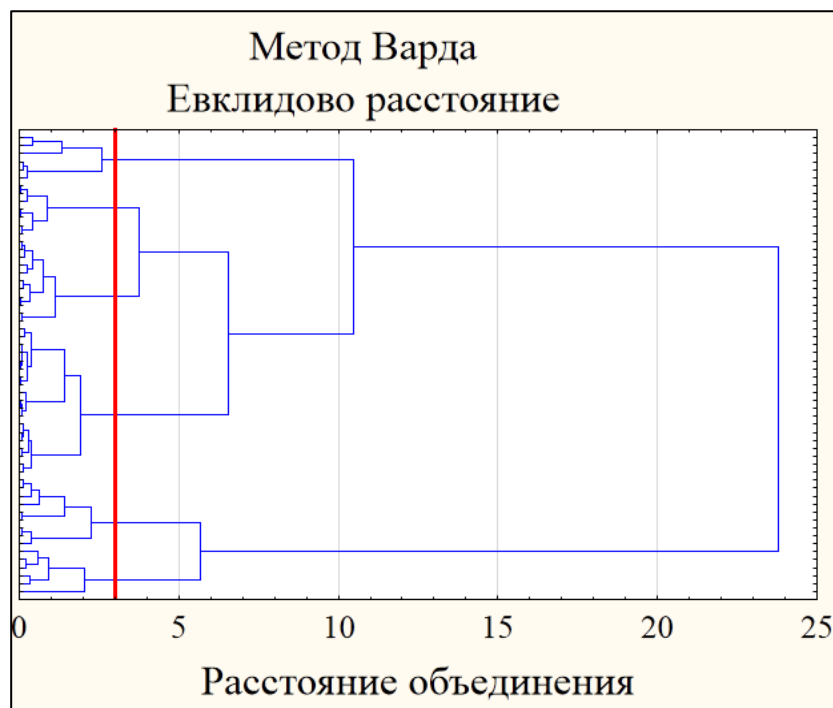


Рисунок 2.3– Горизонтальные дендрограммы 58 регионов РС

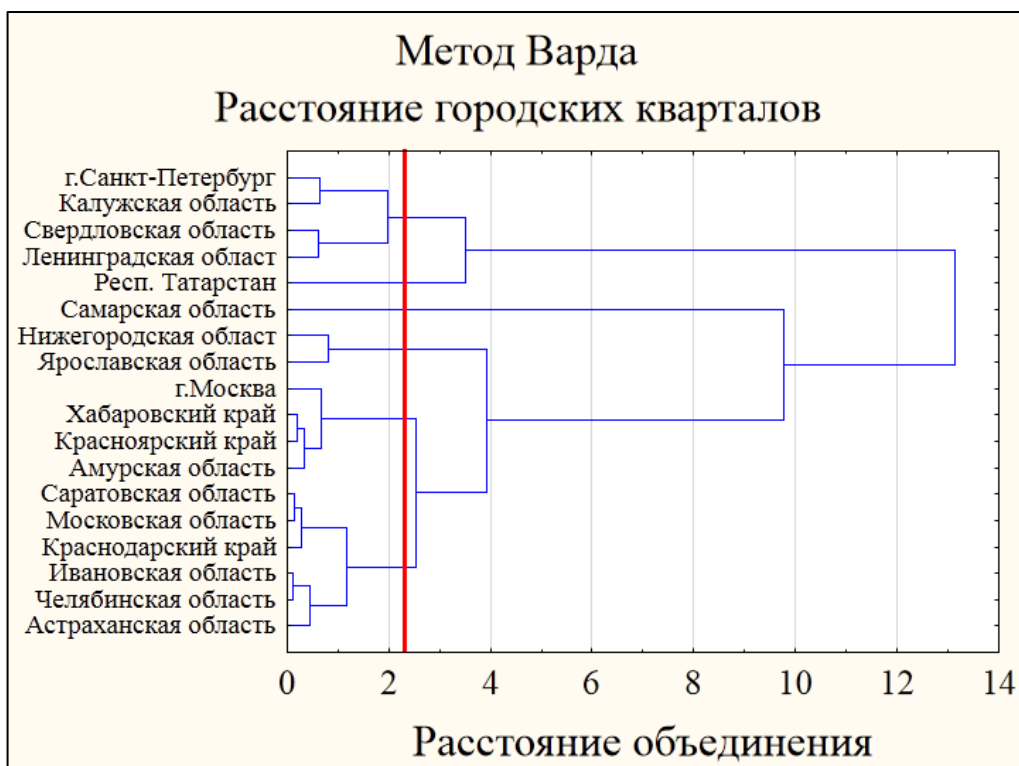
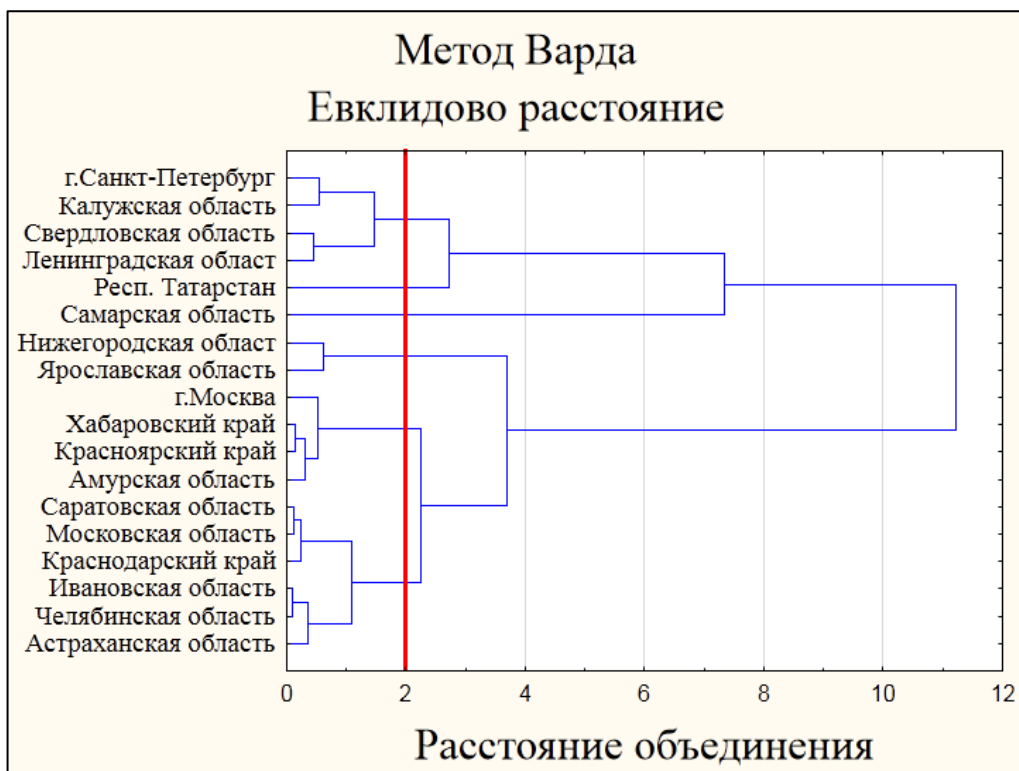


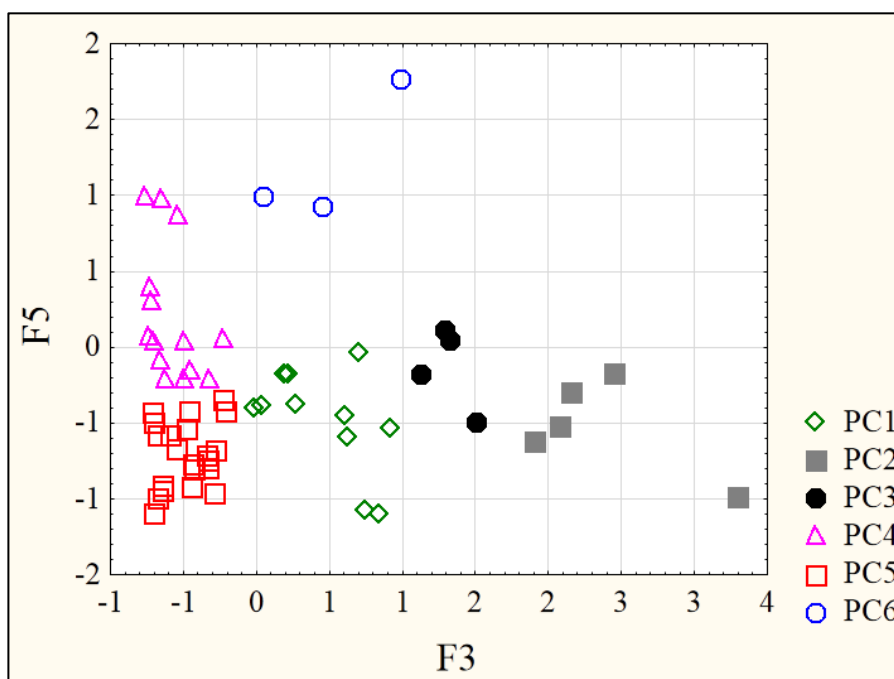
Рисунок 2.4– Горизонтальные дендрограммы 18 регионов ИСС

Примеры регионов, входящих в кластеры и кластерные средние приведены в таблице 2.4.

Таблица 2.4–Средние факторных показателей кластеров регионов

DM_PC+ИСС

Кластер	F3	F5	Кол-во регионов	Примеры регионов
PC1	0,467	-0,484	11	Ярославская обл., Респ. Татарстан, Ульяновская обл., Иркутская обл. Приморский край.
PC2	2,375	-0,517	5	Архангельская обл., Самарская обл., Нижегородская обл., г. Москва, Московская обл.
PC3	1,309	-0,129	4	Ростовская обл., Челябинская обл., г. Санкт-Петербург, Свердловская обл.
PC4	-0,583	0,206	14	Кемеровская область, Кировская область, Тамбовская обл., Пензенская обл., Орловская обл.
PC5	-0,486	-0,703	21	Курганская обл., Владимирская обл., Респ. Мордовия, Ставропольский край, Респ. Дагестан
PC6	0,495	1,231	3	Пермский край, Респ. Башкортостан, Респ. Бурятия
ИСС1	-0,172	2,733	4	г. Санкт-Петербург, Калужская обл., Ленинградская обл., Свердловская обл.
ИСС2	1,634	1,944	1	Респ. Татарстан
ИСС3	3,967	-0,554	1	Самарская обл.
ИСС4	-0,554	-0,783	2	Ярославская обл., Нижегородская обл.
ИСС5	-0,658	1,154	4	Красноярский край, Амурская обл., г. Москва, Хабаровский край
ИСС6	-0,716	0,424	6	Саратовская обл., Московская обл. Краснодарский край, Ивановская обл., Челябинская обл., Астраханская обл.



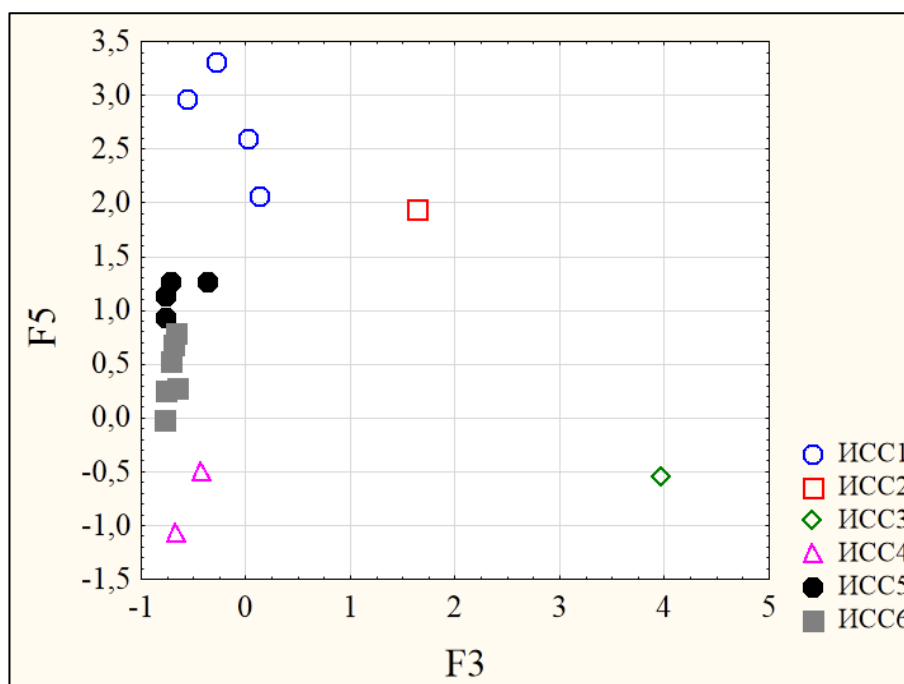


Рисунок 2.5– Диаграмма рассеяния регионов РС и ИСС по 2-м факторным показателям F3 и F5

При этом качество построенной модели оценивается как параметрическими, так и непараметрическими критериями. В рассматриваемом случае F-критерий показывает, что различия между кластерными средними значений регионов высоко значимо для всех факторов (уровень значимости $p_F < 0,0005$). Малость выборок кластеров предполагает контроль полученных результатов ранговым критерием Краскела-Уоллиса, который подтверждает выводы F-критерия для РС и сглаживает до статистически значимых (уровень значимости $0,005 < p_{K-y} < 0,05$) в случае ИСС (табл.2.5).

Таблица 2.5–Результаты дисперсионного анализа качества кластеризации регионов

Фактор	РС		ИСС	
	p_F	p_{K-y}	p_F	p_{K-y}
F3	0,0000	0,0000	0,0000	0,03
F5	0,0000	0,0000	0,0000	0,006

Различия кластеров и исследование их на однородность были оценены с помощью параметрического и непараметрического дисперсионного анализа отдельно по формам собственности и совместно.

а) кластеризация РС

С помощью параметрического F-критерия были выявлены однородные группы кластеров и распределены в порядке возрастания кластерных средних (табл. 2.6, табл. 2.7):

F3: {PC4, PC5}, {PC1, PC6}, {PC3}, {PC2};

F5: {PC5, PC2}, {PC2, PC1}, {PC3}, {PC4}, {PC6}.

Таблица 2.6 – Параметрический F-критерий для F3

F3	PC6	PC5	PC1	PC4	PC3	PC2
PC6		0.000	0.872	0.000	0.000	0.000
PC5	0.000		0.000	0.304	0.000	0.000
PC1	0.872	0.000		0.000	0.000	0.000
PC4	0.000	0.304	0.000		0.000	0.000
PC3	0.000	0.000	0.000	0.000		0.000
PC2	0.000	0.000	0.000	0.000	0.000	

Таблица 2.7 – Параметрический F-критерий для F5

F5	PC6	PC5	PC1	PC4	PC3	PC2
PC6		0.000	0.000	0.000	0.000	0.000
PC5	0.000		0.081	0.000	0.002	0.263
PC1	0.000	0.081		0.000	0.072	0.856
PC4	0.000	0.000	0.000		0.080	0.000
PC3	0.000	0.002	0.072	0.080		0.086
PC2	0.000	0.263	0.856	0.000	0.086	

Непараметрический критерий Краскела-Уоллиса подтверждает результаты параметрического F- критерия в случае F3 и сглаживает результаты до незначимых относительно различий между PC3 и PC4 в случае F5:

F3: {PC4, PC5}, {PC1, PC6}, {PC3}, {PC2};

F5: {PC5, PC2}, {PC2, PC1, PC3}, {PC3, PC4}, {PC6}.

Причем по фактору F3 группа однородных кластеров {PC4, PC5} высоко значимо (на уровне значимости $p < 0,0005$), отличается от PC1; {PC1, PC6} статистически значимо (на уровне значимости $0,005 < p \approx 0,01 < 0,05$) от PC3; различия между PC3 и PC2 статистически значимые (на уровне значимости $0,005 < p \approx 0,01 < 0,05$). По фактору F5 группа однородных кластеров {PC5, PC2} слабо значимо отличается от PC1 (на уровне значимости $0,05 < p \approx 0,07 < 0,1$); {PC2, PC1, PC3} отличается от PC4 статистически высоко значимо (на уровне значимости $p < 0,0005$); {PC3, PC4} отличается от PC6 статистически значимо (на уровне значимости $0,005 < p \approx 0,04 < 0,05$).

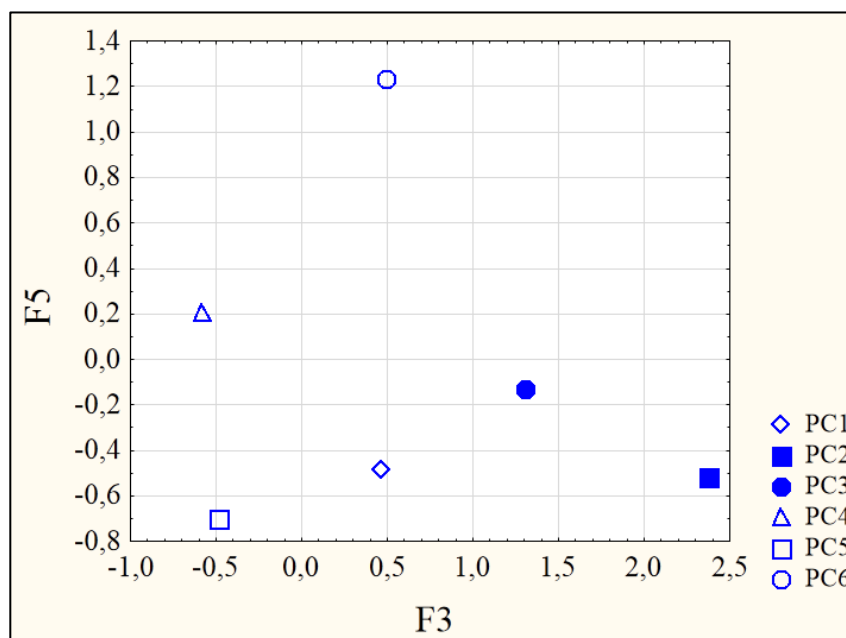


Рисунок 2.6– Диаграмма рассеяния кластерных средних для регионов PC по 2-м факторным показателям F3 и F5

Из рисунка 2.6 следует, что лидерами по количеству занятого населения и ФОТ являются регионы кластера PC2 (Архангельская область, Самарская область, Нижегородская область, г. Москва, Московская область). Несмотря на это, темпы роста занятого населения в этих регионах низкие, как и в большинстве регионах (кластер PC5: 21 регион). Регионы кластера PC6 (Пермский край, Республика Башкортостан, Республика Бурятия) показывают самые высокие темпы роста занятого населения.

б) кластеризация ИСС

С помощью параметрического F-критерия были выявлены однородные группы кластеров и распределены в порядке возрастания кластерных средних (табл. 2.8, табл. 2.9):

F3: {ИСС6, ИСС5, ИСС4}, {ИСС1}, {ИСС2}, {ИСС3};

F5: {ИСС4, ИСС3}, {ИСС6}, {ИСС5}, {ИСС2}, {ИСС1}.

Таблица 2.8 – Параметрический F-критерий для F3

F3	ИСС1	ИСС2	ИСС3	ИСС4	ИСС5	ИСС6
ИСС1		0.000	0.000	0.041	0.004	0.001
ИСС2	0.000		0.000	0.000	0.000	0.000
ИСС3	0.000	0.000		0.000	0.000	0.000
ИСС4	0.041	0.000	0.000		0.545	0.323
ИСС5	0.004	0.000	0.000	0.545		0.648
ИСС6	0.001	0.000	0.000	0.323	0.648	

Таблица 2.9 – Параметрический F-критерий для F5

F5	ИСС1	ИСС2	ИСС3	ИСС4	ИСС5	ИСС6
ИСС1		0.073	0.000	0.000	0.000	0.000
ИСС2	0.073		0.000	0.000	0.072	0.002
ИСС3	0.000	0.000		0.612	0.001	0.027
ИСС4	0.000	0.000	0.612		0.000	0.001
ИСС5	0.000	0.072	0.001	0.000		0.008
ИСС6	0.000	0.002	0.027	0.001	0.008	

Непараметрический критерий Краскела-Уоллиса сглаживает результаты параметрического F- критерия:

F3: {ИСС6, ИСС5, ИСС4}, {ИСС4, ИСС1, ИСС2, ИСС3};

F5: {ИСС4, ИСС3}, {ИСС3, ИСС6}, {ИСС5, ИСС2}, {ИСС2, ИСС1}.

Группа однородных кластеров по F3 {ИСС6, ИСС5, ИСС4} отличается от ИСС1 статистически значимо (на уровне значимости $0,005 < p \approx 0,04 < 0,05$). По фактору F5 выявлены слабо значимые различия между {ИСС4, ИСС3} и ИСС6 (на уровне значимости $0,05 < p \approx 0,07 < 0,1$); статистически значимые между {ИСС3, ИСС6} и ИСС5 (на уровне

значимости $0,005 < p \approx 0,01 < 0,05$); статистически значимые между {ИСС5, ИСС2} и ИСС1.

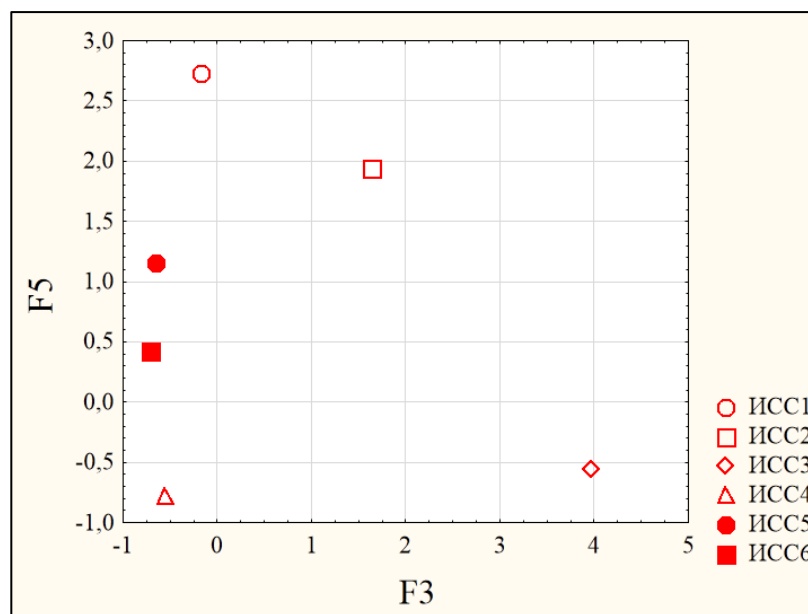


Рисунок 2.7– Диаграмма рассеяния кластерных средних для регионов ИСС по 2-м факторным показателям F3 и F5

Из рис. 2.7 можно сделать вывод, что самые крупные предприятия в Самарской области (ИСС3), но несмотря на это данный регион имеет низкие темпы роста занятого населения и ФОТ, как и регионы кластера ИСС4 (Нижегородская область и Ярославская область). Самые высокие темпы роста показывают регионы кластера ИСС1 (г. Санкт-Петербург, Калужская область, Ленинградская область, Свердловская область).

в) совместная кластеризация

Таким образом, распределение регионов в случае каждой формы собственности является значимо неоднородным по всем факторным показателям (табл.2.5, рис.2.8).

С помощью непараметрического критерия Краскела-Уоллиса были выявлены незначимые различия (на уровне значимости $p \approx 0,5 > 0,1$) между кластерами ИСС4 и ИСС5 по фактору F5 и слабо значимые различия (на уровне значимости $0,1 > p \approx 0,09 > 0,05$) по фактору F3. Между факторами ИСС6 и ИСС4 выявлены слабо значимые различия по фактору F3 (на

уровне значимости $0,1 > p \approx 0,08 > 0,05$) и статистически значимые различия по фактору F5 (на уровне значимости $0,005 > p \approx 0,03 > 0,005$).

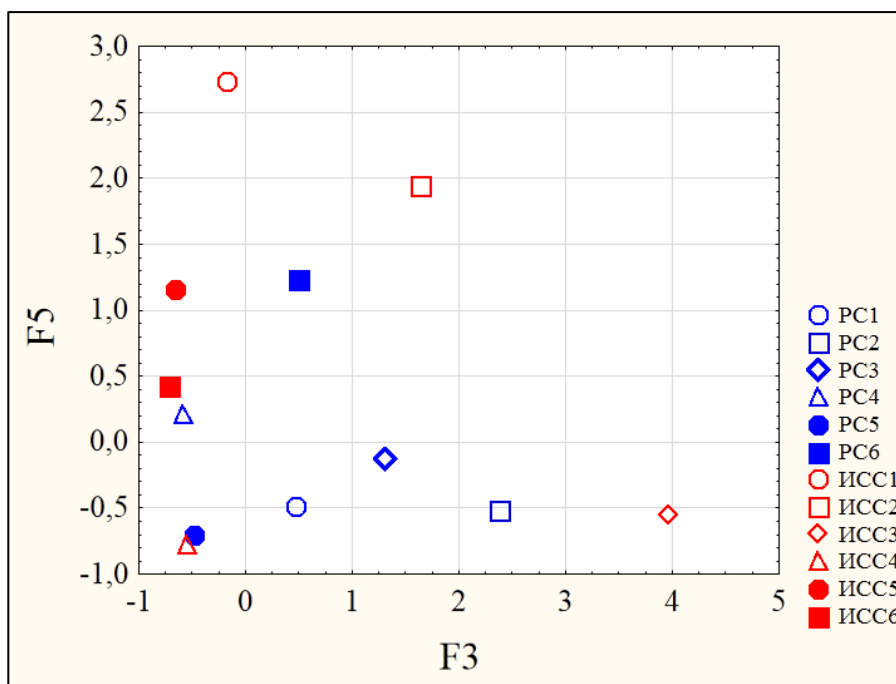


Рисунок 2.8–Диаграмма рассеяния кластерных средних регионов РС и ИСС по 2-м факторным показателям F3 и F5

Согласно проведенному дисперсионному анализу кластеров было выявлено, что самое большое количество занятого населения у предприятий монокластера ИСС3 (Самарская обл.), среди предприятий РС лидерами по количеству занятого населения является кластер РС2 (Архангельская обл., Самарская обл., Нижегородская обл., г. Москва, Московская обл.). Несмотря на лидерство в данном показателе, предприятия в регионах этих кластеров показывают низкий темп роста, в отличие от кластеров ИСС1 (г. Санкт-Петербург, Калужская обл., Ленинградская обл., Свердловская обл.), ИСС2 (Респ. Татарстан). Среди РС самым высоким темпом роста выделяются регионы кластера РС5. Аутсайдерами как по количеству занятого населения, так и по темпам роста являются кластеры РС5 и ИСС4 (Ярославская обл., Нижегородская обл.).

Аналогичным образом были проведены кластеризации регионов для F2 и F6, F1 и F3.

Кластеризация F1 и F3.

Проведена кластеризация регионов (проверка однородности) по каждой форме собственности методом К-средних и методами иерархической классификации. Разными методами кластеризации получены близкие результаты на уровне 5-ти кластерной модели 58-ми регионов РС и 6-ти кластерной модели 18-ти регионов ИСС (рис.2.9).

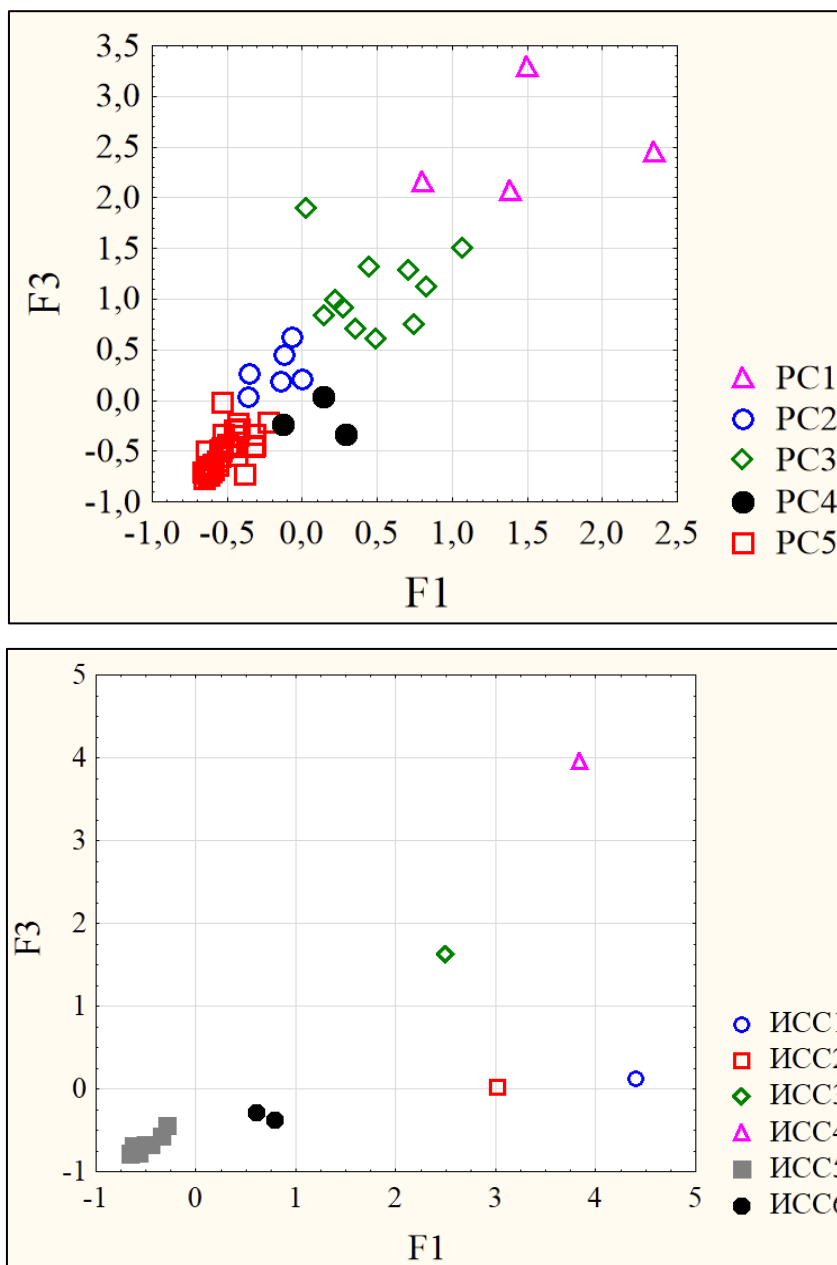


Рисунок 2.9 – Диаграмма рассеяния регионов РС и ИСС по 2-м факторным показателям F1 и F3

Примеры регионов, входящих в кластеры и кластерные средние приведены в таблице 2.10.

Таблица 2.10 – Средние факторных показателей кластеров регионов
DM_PC+ИСС

Кластер	F1	F3	Кол-во регионов	Примеры регионов
PC1	1,505	2,494	4	Самарская обл., Нижегородская обл., г. Москва, Московская обл.
PC2	-0,177	0,298	6	Пермский край, Тверская обл., Новосибирская обл., Воронежская обл., Красноярский край, Хабаровский край, Калужская обл.
PC3	0,484	1,085	11	Респ. Башкортостан, респ. Татарстан, Ярославская обл., Ульяновская обл., Иркутская обл., Приморский край, Ростовская обл., Челябинская обл., г. Санкт-Петербург, Свердловская обл., Архангельская обл.
PC4	0,099	-0,174	3	Респ. Бурятия, Калининградская обл., Брянская обл.
PC5	-0,525	-0,524	34	Саратовская обл., Ленинградская обл., Алтайский край, Тюменская обл., Омская обл., Кемеровская обл.
ИСС1	4,399	0,131	1	г. Санкт-Петербург
ИСС2	3,015	0,024	1	Калужская обл.
ИСС3	2,494	1,634	1	Респ. Татарстан
ИСС4	3,838	3,967	1	Самарская обл.
ИСС5	-0,553	-0,685	12	Свердловская обл., Нижегородская обл., Хабаровский край, Краснодарский край, Московская обл., Ярославская обл., Красноярский край, Амурская обл., Астраханская обл., Челябинская обл., Саратовская обл., Ивановская обл.
ИСС6	0,693	-0,328	2	г. Москва, Ленинградская обл.

Качество построенных кластерных моделей регионов оценено критериями дисперсионного анализа (параметрическим F-критерием и ранговым критерием Краскела-Уоллиса). Согласно F-критерию различия между кластерными средними значений регионов высоко значимы (на уровне $p_F < 0,0005$) для каждой формы собственности как в случае F1, так и F3 (табл.2.11). Малость выборок кластеров предполагает контроль полученных результатов ранговым критерием Краскела-Уоллиса, который смягчает их в случае PC и ИСС до статистически значимых (на уровне $0,05 > p_{K-Y} > 0,005$) (табл.2.11).

Таблица 2.11 – Результаты дисперсионного анализа качества кластеризации регионов

Фактор	PC		ИСС	
	p_F	p_{K-y}	p_F	p_{K-y}
F1	0,0000	0,0000	0,0000	0,04
F3	0,0000	0,0000	0,0000	0,04

Таким образом, распределение регионов в случае каждой формы собственности является значимо неоднородным по всем факторным показателям (табл.2.11, рис.2.10). При выделении групп однородных кластеров с помощью непараметрического критерия Краскела-Уоллиса были выявлены незначимые различия (на уровне значимости $p \approx 0,3 > 0,1$) между кластерами ИСС5 и PC5 по фактору F1 и слабо значимые различия (на уровне значимости $0,005 > p \approx 0,004 > 0,0005$) по фактору F3. Причем ближайшие к данной группе кластеры отличаются значимо: PC2 – сильно значимо по фактору F1 ($0,005 > p \approx 0,001 > 0,0005$) и высоко значимо по фактору F3 ($p < 0,0005$); PC4 – статистически значимо по фактору F1 ($0,05 > p \approx 0,01 > 0,005$) и сильно значимо по фактору F3 ($0,005 > p \approx 0,001 > 0,0005$), ИСС6 – статистически значимо по фактору F1 ($0,05 > p \approx 0,04 > 0,005$) и сильно значимо по фактору F3 ($0,005 > p \approx 0,004 > 0,0005$).

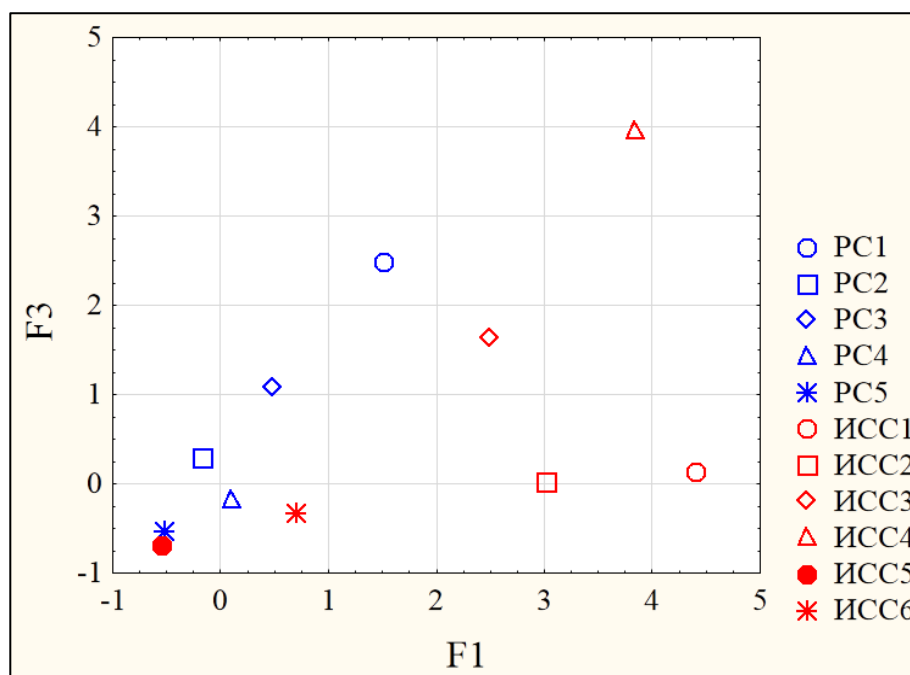


Рисунок 2.10 – Диаграмма рассеяния кластерных средних регионов РС и ИСС по 2-м факторным показателям F1 и F3

Согласно проведенному дисперсионному анализу кластеров предприятий РС и ИСС можно сделать выводы, что предприятия ИСС выпускают больше продукции, чем предприятия РС, несмотря на то, что количество занятого населения в основном больше на предприятиях РС. Лидерами по объёму отгруженной продукции являются предприятия РС регионов: Самарская область, Нижегородская область, Московская область и г. Москва (РС1). Предприятия данных регионов являются также лидерами по количеству занятого населения и соответственно ФОТ. Для предприятий ИСС лидерами по объёму отгруженной продукции являются: г. Санкт-Петербург (ИСС1), Калужская область (ИСС2), Республика Татарстан (ИСС3), Самарская область (ИСС4). Причем самые крупные предприятия ИСС находятся в Самарской области (ИСС4) и Республике Татарстан (ИСС3). Наименьшим объёмом отгруженной продукции, а также количеством занятых и ФОТом как предприятий РС, так и предприятий ИСС характеризуются большинство регионов (РС5, ИСС5).

Кластеризация F2 и F6.

Проведена кластеризация регионов (проверка однородности) по каждой форме собственности методом К-средних и методами иерархической классификации. Различными методами кластеризации получены близкие разбиения регионов на кластеры на уровне 10-ти кластерной модели 58-ти регионов РС и 9-ти кластерной модели 18-ти регионов ИСС (рис.2.11).

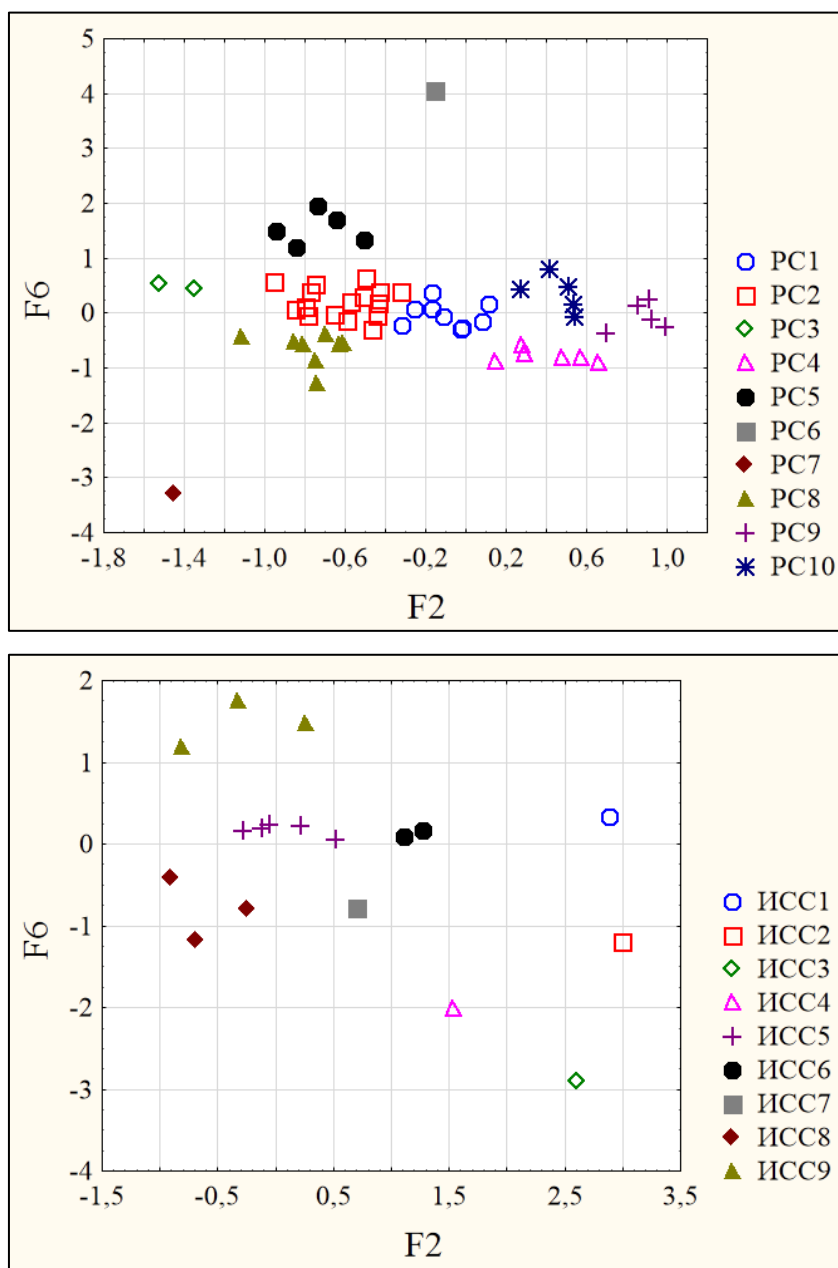


Рисунок 2.11 – Диаграмма рассеяния регионов РС и ИСС по 2-м факторным показателям F2 и F6

Примеры регионов, входящих в кластеры и кластерные средние приведены в таблице 2.12.

Таблица 2.12 – Средние факторных показателей кластеров регионов
DM_PC+ИСС

Кластер	F2	F6	Кол-во регионов	Примеры регионов
PC1	-0,097	-0,029	9	Кемеровская обл., Ярославская обл., Калужская обл., Краснодарский край, Ростовская обл., Челябинская обл., Пермский край, Новосибирская обл., Тверская обл.
PC2	-0,616	0,21	16	Вологодская обл., Саратовская обл., Курганская обл., Владимирская обл., Алтайский край, респ. Мордовия
PC3	-1,438	0,486	2	Ставропольский край, респ. Дагестан
PC4	0,403	-0,791	6	Ленинградская обл., Тюменская обл., Забайкальский край, Красноярский край, респ. Бурятия, Калининградская обл.
PC5	-0,734	1,542	5	Удмуртская респ., Ивановская обл., Брянская обл., Башкортостан респ., Смоленская обл.
PC6	-0,158	4,073	1	Новгородская обл.
PC7	-1,451	-3,297	1	Карачаево –Черкесская респ.
PC8	-0,777	-0,639	8	Оренбургская обл., Астраханская обл., Курская обл., Костромская обл., Волгоградская обл., Тульская обл., респ. Марий Эл, Ульяновская обл.
PC9	0,871	-0,064	5	Камчатский край, Иркутская обл., г. Санкт-Петербург, г. Москва, Московская обл.
PC10	0,450	0,373	5	Хабаровский край, респ. Хакасия, Приморский край, Свердловская обл., Архангельская обл.
ИСС1	2,884	0,335	1	г. Санкт-Петербург
ИСС2	2,994	-1,197	1	г. Москва
ИСС3	2,607	-2,902	1	Калужская обл.
ИСС4	1,530	-2,013	1	Ленинградская обл.
ИСС5	0,046	0,176	5	Краснодарский край, Красноярский край, Амурская обл., Респ. Татарстан, Самарская обл.
ИСС6	1,191	0,131	2	Хабаровский край, Московская обл.
ИСС7	0,693	-0,773	1	Челябинская обл.
ИСС8	-0,612	-0,789	3	Саратовская обл., Астраханская обл., Ярославская обл.
ИСС9	-0,301	1,475	3	Ивановская обл., Нижегородская обл., Свердловская обл.

Качество построенных кластерных моделей регионов оценено параметрическими и непараметрическими критериями дисперсионного анализа (параметрическим F-критерием и ранговым критерием Краскела-

Уоллиса). С помощью F-критерия выявлено, что различия между кластерными средними значений регионов высоко значимы (на уровне $p_F < 0,0005$) для каждой формы собственности как в случае F2, так и F6 (табл.2.13). Малость выборок кластеров предполагает контроль полученных результатов ранговым критерием Краскела-Уоллиса, который смягчает их до слабо значимого на уровне $0,1 > p_{K-Y} \approx 0,07 > 0,05$) в случае РС и до статистически значимого в случае ИСС (на уровне $0,05 > p_{K-Y} \approx 0,04 > 0,005$) (табл.2.13).

Таблица 2.13 – Результаты дисперсионного анализа качества кластеризации регионов

Фактор	РС		ИСС	
	p_F	p_{K-Y}	p_F	p_{K-Y}
F2	0,0000	0,0000	0,0000	0,07
F6	0,0004	0,0000	0,0000	0,04

Таким образом, распределение регионов в случае каждой формы собственности является значимо неоднородным по всем факторным показателям (табл.2.13, рис.2.12). Однако при выделении групп однородных кластеров с помощью непараметрического критерия Краскела-Уоллиса были выявлены незначимые различия (на уровне значимости $>0,1$) между кластерами РС8 и ИСС8 по обоим факторам, причем фактор РС2 отличается от группы высоко значимо по фактору F6. Также выявлены незначимые различия между РС1 и ИСС5 по фактору F2 (на уровне значимости $p \approx 0,46 > 0,1$) и слабо значимые различия по F6 (на уровне значимости $0,005 > p \approx 0,07 > 0,0005$). Еще одна группа однородных кластеров РС4 и ИСС7: незначимые различия (на уровне значимости $>0,1$) как для F2 так и для F6. Причем кластер РС9 отличается от группы статистически значимо как для F2 так и для F6.

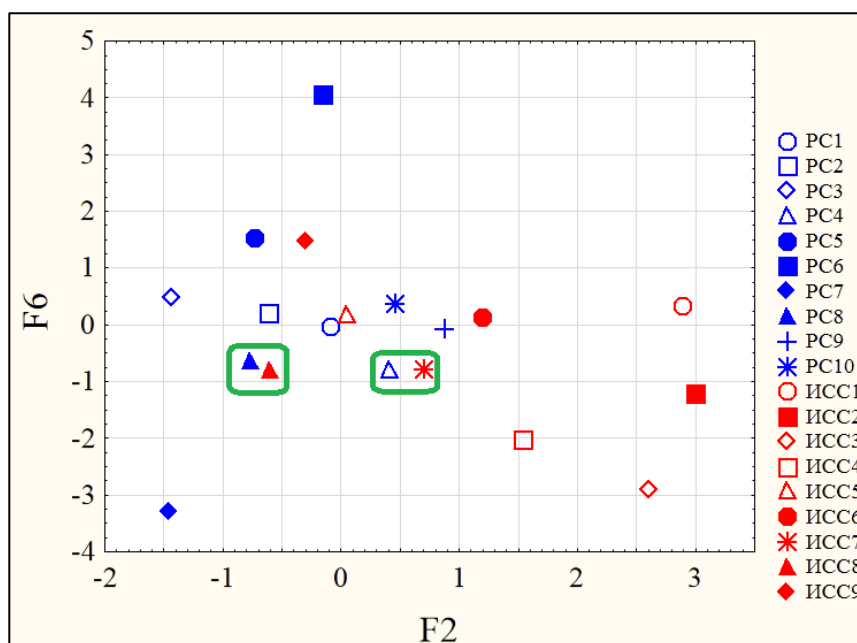


Рисунок 2.12 – Диаграмма рассеяния кластерных средних регионов РС и ИСС по 2-м факторным показателям F2 и F6

Согласно проведенному дисперсионному анализу выявлено, что средняя ЗП выше на предприятиях ИСС, чем на предприятиях РС. Однако предприятия РС показывают более высокие темпы роста, чем предприятия ИСС. Лидерами по средней ЗП предприятий РС являются регионы кластера РС9: Камчатский край, Иркутская область, г. Санкт-Петербург, г. Москва, Московская область. При этом рост ЗП для этих регионов невысокий. Лидерами по средней ЗП для предприятий ИСС являются г. Москва, г. Санкт-Петербург и Калужская область. Несмотря на это данные монокластеры имеют разные темпы роста ЗП, так самые высокие темпы роста зафиксированы в г. Санкт-Петербурге, а наименьшие - в Калужской обл.

2.4. Выводы

В ходе проведения многомерного статистического анализа СЭП предприятий подраздела ДМ в РС и ИСС на уровне регионов России были проведены: корреляционный и кластерный анализы первоначальных показателей, факторный анализ первоначальных

показателей, а также параметрический и непараметрический дисперсионный анализ кластеров.

1) Согласно корреляционному и кластерному анализу первоначальных показателей были выявлены высоко значимые (на уровне значимости $p < 0,0005$) положительные корреляционные связи между показателями №3 и №4 (коэффициент Пирсона $r = 0,96$, коэффициент Спирмена $R = 0,97$), показателями №5 и №6 (коэффициент Пирсона $r = 0,86$, коэффициент Спирмена $R = 0,83$) и показателями №8 и №9 (коэффициент Пирсона $r = 0,84$, коэффициент Спирмена $R = 0,92$), а также между показателями №1 и №3, №1 и №4.

2) Наличие корреляционных связей даёт основание для проведения факторного анализа. Так, методом главных компонент было построено 6-мерное экономическое факторное пространство.

3) Кластеризация регионов различных ФС по факторным показателям была проведена в 3-х основных направлениях:

Кластеризация F3 и F5: отражает размеры Занятого населения, ФОТа и их темпы роста. С помощью методов К-средних и методом иерархической кластеризации были получены высоко качественные разбиения на уровне 6-ти кластерной модели 58-ти регионов РС и 6-ти кластерной модели 18-ти регионов ИСС. При проведении Дисперсионного анализа различий средних кластеров однородных групп кластеров по обоим факторам выявлено не было.

Кластеризация F1 и F3: отражает размеры отгруженной продукции, ФОТа и количества занятого населения. С помощью методов К-средних и методом иерархической кластеризации были получены высоко качественные разбиения на уровне 5-ти кластерной модели 58-ми регионов РС и 6-ти кластерной модели 18-ти регионов ИСС. При проведении дисперсионного анализа различий средних кластеров однородных групп кластеров по обоим факторам выявлено не было.

Кластеризация F2 и F6: отражает размеры ЗП и её темпы роста. С помощью методов К-средних и методом иерархичной кластеризации были получены высоко качественные разбиения на уровне 10-ти кластерной модели 58-ти регионов РС и 9-ти кластерной модели 18-ти регионов ИСС. При проведении дисперсионного анализа средних кластеров были выявлены 2 группы однородных факторов РС8+ИСС8, РС4+ИСС7.

Проведение кластеризации в данных направлениях позволило выделить регионы, являющиеся лидерами/аутсайдерами по количеству отгруженной продукции, размеру средней заработной платы, количеству занятого населения; отследить темпы роста и изменения данных показателей.

3. ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ

3.1. Оценка коммерческого потенциала и перспективности проведения научных исследований с позиции ресурсоэффективности и ресурсосбережения

3.1.1. Потенциальные потребители результатов исследования

Для анализа потребителей результатов исследования необходимо рассмотреть целевой рынок и провести его сегментирование [31].

Целевой рынок – сегменты рынка, на котором будет продаваться в будущем разработка. В свою очередь, сегмент рынка – это особым образом выделенная часть рынка, группы потребителей, обладающих определенными общими признаками.

Сегментирование – это разделение покупателей на однородные группы, для каждой из которых может потребоваться определенный товар (услуга).

Потенциальные потребители результатов исследования:

Организации, использующие статистические данные в качестве составления рейтинга (регионов РФ, предприятий, ВЭДов и т.д.):

- российские государственные учреждения;
- российские частные компании;
- отделы развития экономики на различных уровнях управления.

Организации, использующие статистические данные для формирования, изменения и утверждения методов развития экономического роста предприятий:

- российские государственные учреждения;
- российские частные компании;
- отделы развития экономики на различных уровнях управления.

		Виды использования статистических данных	
Вид потребителя		Составление рейтинга	Формирование, изменение и утверждение методов развития экономического роста предприятий:
	Российские государственные учреждения		
	Российские частные компании		
	Отделы развития экономики на различных уровнях управления		
		Фирма А. 	Фирма Б.

Рисунок 3.1 – Карта сегментирования рынка услуг по разработке комплексной математической модели

3.1.2. Анализ конкурентных технических решений

Основными конкурентами созданной модели статистического многомерного анализа являются модели, использующие 6-факторную модель, а также модели, использующие другое количество разбиения на кластеры. Модель, построенная в 6-ти факторном экономическом пространстве, является высококачественной, и имеет логические обоснования. Модели, содержащие другое количество разбиения на кластеры не подтверждаются различными методами разбиения.

Ниже представлена оценочная карта для сравнения конкурентных программных разработок (Таблица 3.1):

Где B_{ϕ} – построенная модель в 5-ти факторном пространстве с разбиением на 6 кластеров;

B_{k1} - модель в 4х факторном пространстве с разбиением на 6 кластеров;

B_{k2} - в 7-ми факторном пространстве с разбиением на 4 кластера.

Таблица 3.1–Оценочная карта для сравнения конкурентных технических решений (разработок)

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		B_{ϕ}	B_{k1}	B_{k2}	K_{ϕ}	K_{k1}	K_{k2}
1	2	3	4	5	6	7	8
Технические критерии оценки ресурсоэффективности							
1. Повышение производительности труда пользователя	0,07	5	4	4	0,35	0,28	0,28
2. Удобство в эксплуатации (соответствует требованиям потребителей)	0,09	5	5	5	0,45	0,45	0,45
3. Помехоустойчивость	-						
4. Энергоэкономичность	-						
5. Надежность	0,09	5	5	4	0,45	0,45	0,36
6. Уровень шума	-						
7. Безопасность	0,06	5	5	5	0,3	0,3	0,3
8. Потребность в ресурсах памяти	0,06	5	5	5	0,3	0,3	0,3
9. Функциональная мощность (предоставляемые возможности)	-						
10. Простота эксплуатации	0,08	5	4	5	0,4	0,32	0,4
11. Качество интеллектуального интерфейса	0,08	5	5	5	0,4	0,4	0,4
12. Возможность подключения в сеть ЭВМ	-						
Экономические критерии оценки эффективности							

1. Конкуренентоспособность продукта	0,08	5	4	5	0,4	0,32	0,4	
2. Уровень проникновения на рынок	0,09	4	4	4	0,36	0,36	0,36	
3. Цена	0,08	4	4	4	0,4	0,4	0,4	
4. Предполагаемый срок эксплуатации	-							
5. Послепродажное обслуживание	0,08	4	4	4	0,32	0,32	0,32	
6. Финансирование научной разработки	0,09	5	4	4	0,45	0,36	0,36	
7. Срок выхода на рынок	0,05	4	4	4	0,2	0,2	0,2	
8. Наличие сертификации разработки	-							
ИТОГО	1				4,78	4,46	4,53	

3.1.3. SWOT- анализ

SWOT – Strengths (сильные стороны), Weaknesses (слабые стороны), Opportunities (возможности) и Threats (угрозы) – представляет собой комплексный анализ научно-исследовательского проекта. SWOT-анализ применяют для исследования внешней и внутренней среды проекта [32]. Разработанная для данного исследования матрица SWOT представлена в таблице 3.2.

Таблица 3.2–Матрица SWOT

Внешняя среда	<p>Сильные стороны научно-исследовательского проекта:</p> <p>С1. Использование корреляционного анализа</p> <p>С2. Использование факторного анализа</p> <p>С3. Использование кластерного анализа</p> <p>С4. Использование дисперсионного анализа</p> <p>С5. Использование современного инструментария компьютерного анализа на примере Statistica</p>	<p>Слабые стороны научно-исследовательского проекта:</p> <p>Сл1. Модель основывается только на статистических и математических данных;</p> <p>Сл2. Большая теоретическая база данных по использованию данной модели;</p> <p>Сл3. Возможность использования данной модели людей с математическим образованием.</p>
	<p>Внутренняя среда</p> <p>Возможности:</p> <p>В1. Использование модели для любых социально-</p>	<p>Использование различных методов многомерного статистического анализа позволяет точно определить</p>

<p>экономических баз данных;</p> <p>В2. Использование модели для любой другой аналогичной базы данных;</p> <p>В3. Появление дополнительного спроса на данную модель.</p>	<p>и выявить сходства и различия наблюдаемых выборок. Корреляционный, кластерный, дисперсионный анализы могут быть применены к любой аналогичной базе данных, что позволяет расширить круг потенциальных клиентов.</p>	<p>Результаты статистического анализа представляют собой подтверждение или опровержение статистических гипотез. Основываясь на результатах проверки данных гипотез, прикладной математик делает выводы и интерпретирует их на исходную базу данных.</p>
<p>Угрозы:</p> <p>У1. Отсутствие спроса на предлагаемые услуги;</p> <p>У2. Ухудшение экономического развития предприятий;</p> <p>У3. Появление конкуренции в данном виде услуг.</p>	<p>Данная модель включает в себя различные виды статистического анализа, что в целом повышает качество модели. Качественная модель анализа экономических данных может быть применена и учтена при формировании законопроектов, планов оздоровления компаний, систем развития экономики. Использование современного компьютерного инструментария, позволяет оформлять качественные и понятные результаты, что повышает спрос на данную модель и делает ее конкурентоспособной.</p>	<p>Сложность в использовании и интерпретации данной модели могут понизить на нее спрос. Некоторые организации основываются только на среднем значении исследуемых показателей, что очень сильно облегчает задачу статистика, но заметно снижает качество модели.</p>

Вывод по SWOT-анализу:

В ходе проведения SWOT-анализа были выявлены сильные, слабые стороны научно-исследовательского проекта, а также его угрозы и возможности. Проект имеет 5 сильных сторон, которые в комплексе дают высококачественную модель анализа данных. Эта модель может быть применена к любой социально-экономической базе данных или аналогичной базе данных. Основной угрозой для данного проекта является конкуренция, которая может существенно повлиять на спрос данной модели.

3.1.4. Оценка готовности проекта к коммерциализации

На какой бы стадии жизненного цикла не находилась научная разработка полезно оценить степень ее готовности к коммерциализации и выяснить уровень собственных знаний для ее проведения (или завершения). Для этого необходимо заполнить специальную форму, содержащую показатели о степени проработанности проекта с позиции коммерциализации и компетенциям разработчика научного проекта. Перечень вопросов приведен в таблице 3.3.

Таблица 3.3– Бланк оценки степени готовности научного проекта к коммерциализации

№ п/п	Наименование	Степень проработанности научного проекта	Уровень имеющихся знаний у разработчика
1.	Определен имеющийся научно-технический задел	5	5
2.	Определены перспективные направления коммерциализации научно-технического задела	5	4
3.	Определены отрасли и технологии (товары, услуги) для предложения на рынке	5	5
4.	Определена товарная форма научно-технического задела для представления на рынок	5	3
5.	Определены авторы и осуществлена охрана их прав	5	4
6.	Проведена оценка стоимости интеллектуальной собственности	4	4
7.	Проведены маркетинговые исследования рынков сбыта	4	3
8.	Разработан бизнес-план коммерциализации научной разработки	1	1
9.	Определены пути продвижения научной разработки на рынок	4	3
10.	Разработана стратегия (форма) реализации научной разработки	1	1
11.	Проработаны вопросы международного сотрудничества и выхода на зарубежный рынок	1	1
12.	Проработаны вопросы использования услуг инфраструктуры поддержки, получения льгот	1	1

13.	Проработаны вопросы финансирования коммерциализации научной разработки	5	4
14.	Имеется команда для коммерциализации научной разработки	5	5
15.	Проработан механизм реализации научного проекта	5	5
	ИТОГО БАЛЛОВ	56	49

Итого получилось суммарное количество баллов по каждому направлению: 56 баллов – по степени проработанности научного проекта; 49 балла – по уровню, имеющихся знаний у разработчика. Согласно этим баллам, можно сказать, что перспективность данной разработки выше среднего.

3.1.5. Методы коммерциализации результатов научно-технического исследования

При коммерциализации научно-технических разработок продавец (а это, как правило, владелец соответствующих объектов интеллектуальной собственности), преследует вполне определенную цель, которая во многом зависит от того, куда в последующем он намерен направить (использовать, вложить) полученный коммерческий эффект. Это может быть получение средств для продолжения своих научных исследований и разработок (получение финансирования, оборудования, уникальных материалов, других научно-технических разработок и т.д.), одноразовое получение финансовых ресурсов для каких-либо целей или для накопления, обеспечение постоянного притока финансовых средств, а также их различные сочетания.

Выделяют следующие методы коммерциализации научных разработок:

1. *Торговля патентными лицензиями*, т.е. передача третьим лицам права использования объектов интеллектуальной собственности на лицензионной основе.

2. *Передача ноу-хау*, т.е. предоставление владельцем ноу-хау возможности его использовать другим лицом, осуществляемое путем раскрытия ноу-хау.

3. *Инжиниринг* как самостоятельный вид коммерческих операций предполагает предоставление на основе договора инжиниринга одной стороной, именуемой консультантом, другой стороне, именуемой заказчиком, комплекса или отдельных видов инженерно-технических услуг, связанных с проектированием, строительством и вводом объекта в эксплуатацию, с разработкой новых технологических процессов на предприятии заказчика, усовершенствованием имеющихся производственных процессов вплоть до внедрения изделия в производство и даже сбыта продукции.

4. *Франчайзинг*, т.е. передача или переуступка (на коммерческих условиях) разрешения продавать чьи-либо товары или оказывать услуги в некоторых областях.

5. *Организация собственного предприятия*.

6. *Передача интеллектуальной собственности* в уставной капитал предприятия.

7. *Организация совместного предприятия*, т.е. объединение двух и более лиц для организации предприятия.

8. *Организация совместных предприятий*, работающих по схеме «российское производство – зарубежное распространение».

Данное научное исследование перспективно и находится на уровне выше среднего, поэтому не все аспекты рассмотрены и изучены. Основной научно-технический задел определен, этого достаточно для коммерциализации для следующих методов (пункты 1 - 2): торговля патентной лицензией и передача ноу-хау. Степени проработанности научного проекта и уровень знаний разработчика достаточно для реализации пунктов, которые были выбраны.

3.1.6. Инициация проекта

Группа процессов инициации состоит из процессов, которые выполняются для определения нового проекта или новой фазы существующего. В рамках процессов инициации определяются изначальные цели и содержание и фиксируются изначальные финансовые ресурсы. Определяются внутренние и внешние заинтересованные стороны проекта, которые будут взаимодействовать и влиять на общий результат научного проекта. Данная информация закрепляется в Уставе проекта.

Устав проекта документирует бизнес-потребности, текущее понимание потребностей заказчика проекта, а также новый продукт, услугу или результат, который планируется создать.

Таблица 3.4–Заинтересованные стороны проекта

Заинтересованные стороны проекта	Ожидания заинтересованных сторон
НИ ТПУ кафедра Высшей математики и математической физики	Создание математической модели оценивания различий предприятий российской, иностранной и совместной форм собственности регионов России

Таблица 3.5–Цели и результаты проекта

Цели проекта:	Исследование деятельности предприятий подразделения «Производство транспортных средств и оборудования» обрабатывающей промышленности России в разрезе форм собственности с использованием методов многомерного статистического анализа
Ожидаемые результаты проекта:	Данное исследование поможет выявить основные различия экономического развития предприятий подразделения «Производство транспортных средств и оборудования» обрабатывающей промышленности России в разрезе форм собственности
Критерии приемки результата проекта:	Адекватность результатов
Требования к результату проекта:	Требование:
	Получение высококачественной математической модели

3.1.7. Организационная структура проекта

На данном этапе работы необходимо решить следующие вопросы: кто будет входить в рабочую группу данного проекта, определить роль каждого участника в данном проекте, а также прописать функции, выполняемые каждым из участников и их трудозатраты в проекте.

Таблица 3.6–Рабочая группа проекта

№ п/п	ФИО, основное место работы, должность	Роль в проекте	Функции	Трудо-затраты, час.
1	Михальчук А.А. Доцент кафедры ВММФ	Руководитель	Координирует деятельность участников проекта	6
2	Новосельцева Д.А. Магистр	Исполнитель	Выполнение НИР	6
3	Спицын В.В. Кандидат экономических наук, доцент ТПУ	Эксперт проекта	Консультирование	2
4	Рыжакина Т.Г. Кандидат экономических наук, доцент ТПУ	Эксперт проекта	Консультирование	2
ИТОГО:				16

3.1.8. Ограничения и допущения проекта

Ограничения проекта – это все факторы, которые могут послужить ограничением степени свободы участников команды проекта, а также «границы проекта» - параметры проекта или его продукта, которые не будут реализованных в рамках данного проекта.

Таблица 3.7–Ограничения проекта

Фактор	Ограничения/ допущения
3.1.1. Источник финансирования	НИТПУ
3.2. Сроки проекта:	1.02.2017-31.05.2017
3.2.1. Дата утверждения плана	1.02.2017

управления проектом	
3.2.2. Дата завершения проекта	31.05.2017

3.2. Планирование научно - исследовательских работ

3.2.1. Иерархическая структура работ проекта

Иерархическая структура работ (ИСР) – детализация укрупненной структуры работ. В процессе создания ИСР структурируется и определяется содержание всего проекта. На рисунке 3.2 представлена иерархическая структура работ для выполнения дипломной работы.



Рисунок 3.2 – Иерархическая структура работ проекта

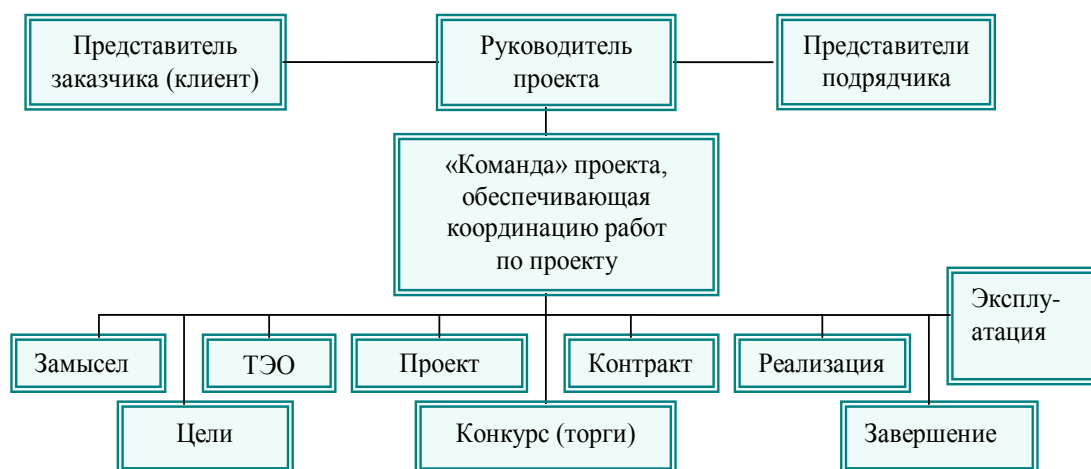


Рисунок 3.3 – Проектная структура проекта

3.2.2. Структура работ в рамках научного исследования

Планирование комплекса предполагаемых работ осуществляется в следующем порядке [33]:

- иерархическая структура работ проекта;
- контрольные события проекта;
- план проекта;
- бюджет научного исследования.

Таблица 3.8—Перечень этапов, работ и распределение исполнителей

Основные этапы	№ раб	Содержание работ	Должность исполнителя
Разработка научного задания	1	Составление и утверждение научного задания	Руководитель
Выбор направления	2	Подбор и изучение материалов по теме	Студент

исследований	3	Проведение патентных исследований	Студент
	4	Выбор направления исследований	Руководитель, Студент
	5	Календарное планирование работ по теме	Руководитель, Студент
Практическое исследование	6	Выбор экспериментальных данных	Студент
	7	Проведение практических расчетов обоснований	Студент
	8	Построение макетов (моделей) и проведение экспериментов	Руководитель, Студент
	9	Сопоставление результатов экспериментов с проведенными исследованиями	Студент
Обобщение и оценка результатов	10	Оценка эффективности полученных результатов	Студент
Оформление отчета НИР	11	Составление пояснительной записки	Студент
	12	Оформление пояснительной записки по ГОСТу	Студент

3.2.3. Определение трудоемкости выполнения работ и разработка графика проведения научного исследования

Диаграмма Ганта – горизонтальный ленточный график, на котором работы по теме представляются протяженными во времени отрезками, характеризующимися датами начала и окончания выполнения данных работ. Для определения трудоемкости работ $t_{ожі}$ используются следующие формулы:

$$t_{ожі} = \frac{3t_{\min i} - 2t_{\max i}}{5} \quad (3.1)$$

Где $t_{ожі}$ – ожидаемая трудоемкость выполнения i -ой работы чел.-дн.;

t_{mini} – минимально возможная трудоемкость выполнения заданной i -ой работы (оптимистическая оценка: в предположении наиболее благоприятного стечения обстоятельств), чел.-дн.;

t_{maxi} – максимально возможная трудоемкость выполнения заданной i -ой работы (пессимистическая оценка: в предположении наиболее неблагоприятного стечения обстоятельств), чел.-дн.

$$T_{pi} = \frac{t_{ожi}}{Ч_i} \quad (3.2)$$

Где T_{pi} – продолжительность одной работы, раб. дн.;

$t_{ожi}$ – ожидаемая трудоемкость выполнения i -ой работы чел.-дн.;

$Ч_i$ – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел.

$$T_{ki} = T_{pi} * k_{\text{кал}} \quad (3.3)$$

где T_{ki} – продолжительность выполнения i -й работы в календарных днях;

T_{pi} – продолжительность выполнения i -й работы в рабочих днях;
кал

$k_{\text{кал}}$ – коэффициент календарности.

Коэффициент календарности определяется по следующей формуле:

$$k_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}} = 1,478 \quad (3.4)$$

Где $T_{\text{кал}}$ – количество календарных дней в году;

$T_{\text{вых}}$ – количество выходных дней в году;

$T_{\text{пр}}$ – количество праздничных дней в году.



Рассчитанные значения в календарных днях по каждой работе T_{ki} необходимо округлить до целого числа.

Таблица 3.9 – Временные показатели проведения научного исследования

№ раб.	Трудоёмкость работ			Исполнители	Длительность работ в рабочих днях T_{pi}	Длительность работ в календарных днях T_{ki}
	t_{min} чел.-дни	t_{max} чел.-дни	$t_{ож}$ чел.-дни			
1.	5	7	5,8	Руководитель	5,8	9
2.	7	10	8,2	Студент	8,2	12
3.	7	10	8,2	Студент	8,2	12
4.	7	10	8,2	Руководитель, Студент	4,1	6
5.	7	10	8,2	Руководитель, Студент	4,1	6
6.	3	5	3,8	Студент	3,8	6
7.	7	14	9,8	Студент	9,8	14
8.	7	10	8,2	Руководитель, Студент	4,1	6
9.	3	5	3,8	Студент	3,8	6
10.	7	10	8,2	Студент	8,2	12
11.	7	10	8,2	Студент	8,2	12
12	7	10	8,2	Студент	8,2	12

Таблица 3.10– Календарный план-график проведения НИР по теме

	Содержание работ	Исполн и-тели	Т _{кi} , кал дн	Февраль		Март			Апрель			Май		
				2	3	1	2	3	1	2	3	1	2	3
				1	Составление и утверждение научного задания	Р	9							
2	Подбор и изучение материалов по теме	И	12											
3	Проведение патентных исследований	И	12											
4	Выбор направления исследований	Р, И	6											
5	Календарное планирование работ по теме	Р, И	6											
6	Выбор экспериментальных данных	И	6											
7	Проведение практических расчетов и обоснований	И	14											
8	Построение макетов (моделей) и проведение экспериментов	Р, И	6											
9	Сопоставление результатов экспериментов с проведенными исследованиями	И	6											
10	Оценка эффективности полученных результатов	И	12											
11	Составление пояснительной записки к НИР	И	12											
12	Оформление пояснительной записки к НИР по ГОСТу	И	12											

Руководитель  Студент 

3.3. Бюджет научно - технического исследования (НТИ)

3.3.1. Расчет материальных затрат НТИ

Данная статья отражает стоимость всех материалов, используемых при разработке проекта, включая расходы на их приобретение и, при необходимости доставку. Транспортные расходы (если таковые имеются) принимаются в пределах 3-5% от стоимости материалов.

Расчет затрат на материалы производится в таблице 3.11.

Таблица 3.11– Затраты на материалы

Наименование материалов и покупных изделий	Единица измерения	Кол-во	Цена за ед., руб.	Сумма, руб.
Бумага, формат А4	Пачка	1	175	175
Flashcard, 2 Гб	Штук	1	500	500
Ручка	Штук	1	15	15
Блокнот	Штук	1	50	50
Итого				740

3.3.2. Основная заработная плата

В этой статье расходов планируется и учитывается основная заработная плата исполнителей, непосредственно участвующих в проектировании разработки:

$$C_{осн/зн} = \sum t_i \cdot C_{zni} \quad (3.5)$$

где t_i – затраты труда, необходимые для выполнения i -го вида работ, в рабочих днях;

C_{zni} – среднедневная заработная плата работника, выполняющего i -ый вид работ, (руб./день).

Среднедневная заработная плата определяется по формуле:

$$C_{zni} = \frac{D \cdot K \cdot M_p}{F_o} \quad (3.6)$$

где D – месячный должностной оклад работника; для руководителя $D=14814$ руб., для исполнителя $D=1470$ руб.;

K – коэффициент, учитывающий коэффициент по премиям (40%) и районный коэффициент (для Томска – 30%); для руководителя $K=1,7$, для исполнителя $K=1,3$;

M_p – количество месяцев работы без отпуска в течение года (при отпуске 48 дней $M_p = 10,4$; при отпуске 24 дней $M_p = 11,2$);

F_o – действительный годовой фонд рабочего времени работника, в днях.

Расчет годового фонда рабочего времени приведен в таблице 7.

Таблица 3.12– Расчет годового фонда рабочего времени

Показатели рабочего времени	Количество дней	
	Руководитель на кафедре	Бакалавр
Календарное число дней в году	365	365
Количество нерабочих дней		
Выходные	118	118
Праздники		
Планируемые потери отпуска	24	60
Действительный годовой фонд	223	187

Месячный должностной оклад работника:

$$Z_m = Z_{тс} * k_p \quad (3.7)$$

где $Z_{тс}$ – заработная плата по тарифной ставке, руб.;

k_p – районный коэффициент, равный 1,3 (для Томска).

Таблица 3.13– Расчёт основной заработной платы

Исполнители	$Z_{тс}$, руб	k_p	Z_m ,руб	$Z_{дн}$,руб	T_p ,дни	$Z_{осн}$,руб
Руководитель (доцент)	23000	1,3	29900	1423	18,1	25756,3

Бакалавр	1710	-	1710	81,42	70,7	5756,4
ИТОГО $Z_{осн}$						31512,7

3.3.3. Дополнительная заработная плата исполнителей темы

Расчет дополнительной заработной платы ведется по следующей формуле:

$$Z_{доп} = k_{доп} * Z_{осн} \quad (3.8)$$

где $k_{доп}$ – коэффициент дополнительной заработной платы (на стадии проектирования принимается равным 0,12 – 0,15).

Таблица 3.14 – Расчет дополнительной заработной платы

Исполнители	Основная ЗП,руб.	Дополнительная ЗП,руб.
Руководитель (доцент)	25756,3	3836,4
Бакалавр	5756,4	773,5
ИТОГО		4609,9

3.3.4. Отчисления во внебюджетные фонды (страховые отчисления)

Величина отчислений во внебюджетные фонды определяется исходя из следующей формулы:

$$Z_{внеб} = k_{внеб} * (Z_{осн} + Z_{доп}) \quad (3.9)$$

где $k_{внеб}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и пр.).

Таблица 3.15 – Отчисления во внебюджетные фонды

Исполнители	Основная ЗП, руб.	Дополнительная ЗП, руб.
Руководитель (доцент)	25 756,3	3 836,4
Бакалавр	5 756,4	773,5
Коэффициент отчислений во	0,271	-

внебюджетные фонды		
	ИТОГО	8 020

3.3.5. Расчет затрат на научные и производственные командировки

Затраты на научные и производственные командировки исполнителей определяются в соответствии с планом выполнения темы и с учетом действующих норм командировочных расходов различного вида и транспортных тарифов. В данном дипломном проекте таких затрат нет.

3.3.6. Контрагентные расходы

Контрагентные расходы включают в себя затраты, связанные с выполнением каких-либо работ по теме сторонними организациями.

Расчет величины этой группы расходов зависит от планируемого объема работа и определяется из условий договоров с контрагентами или субподрядчиками. Контрагентные расходы составляют 10% от основной и дополнительной заработной платы. В данном дипломном проекте таких затрат нет.

3.3.7. Накладные расходы

Накладные расходы учитывают прочие затраты организации, не попавшие в предыдущие статьи расходов: печать и ксерокопирование материалов исследования, оплата услуг связи, электроэнергии, почтовые и телеграфные расходы, размножение материалов и т.д. Их величина определяется по следующей формуле:

$$Z_{\text{накл}} = (Z_{\text{осн}} + Z_{\text{доп}}) \cdot k_{\text{нр}},$$

где $k_{\text{нр}}$ – коэффициент, учитывающий накладные расходы.

$$Z_{\text{накл}} = (31\,512,7 + 4\,609,9) \cdot 0,6 = 21\,673,6$$

3.3.8. Формирование бюджета затрат НИИ

Рассчитанная величина затрат научно-исследовательской работы(темы) является основой для формирования бюджета затрат проекта, который при формировании договора с заказчиком защищается научной организацией в качестве нижнего предела затрат на разработку научно-технической продукции. Определение бюджета затрат на научно-исследовательский проект приведен в Таблице 3.16.

Таблица 3.16 – Расчет бюджета затрат НИИ

Наименование статьи	Сумма, руб.	Примечание
1. Материальные затраты НИИ	740	Пункт 3.3.1
2. Затраты по основной заработной плате исполнителей темы	31 512,7	Пункт 3.3.2
3. Затраты по дополнительной заработной плате исполнителей темы	4 609,9	Пункт 3.3.3
4. Отчисления во внебюджетные фонды	8 020	Пункт 3.3.4
5. Расчет затрат на научные и производственные командировки	0	Пункт 3.3.5
6. Контрагентные расходы	0	Пункт 3.3.6
7. Накладные расходы	21 673,6	Пункт 3.3.7
Бюджет затрат НИИ	55 556,2	Сумма ст.1-7
Аналог 1	70 556,2	Использование дополнительного оборудования, стоимостью 15 000 руб.
Аналог 2	74 556,2	Использование дополнительного оборудования, стоимостью 19 000 руб.

3.4. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Эффективность научного ресурсосберегающего проекта включает в себя:

- социальную эффективность;

Показатели общественной эффективности учитывают социально-экономические последствия осуществления инвестиционного проекта как для общества в целом, так и результаты в смежных секторах экономики, экологические, социальные и иные внеэкономические эффекты.

-экономическую эффективность;

Показатели экономической эффективности проекта учитывают финансовые последствия его осуществления для предприятия, реализующего данный проект.

-бюджетную эффективность.

Бюджетная эффективность характеризуется участием государства в проекте с точки зрения расходов и доходов бюджетов всех уровней.

3.4.1. Динамические методы экономической оценки инвестиций

Динамические методы оценки инвестиций базируются на применении показателей:

- чистая текущая стоимость (NPV);

- срок окупаемости (DPP);

- внутренняя ставка доходности (IRR);

- индекс доходности (PI).

Все перечисленные показатели основываются на сопоставлении чистых денежных поступлений от операционной и инвестиционной деятельности, и их приведении к определенному моменту времени. Теоретически чистые денежные поступления можно приводить к любому моменту времени (к будущему либо текущему периоду). Но для практических целей оценку

инвестиции удобнее осуществлять на момент принятия решений об инвестировании средств.

3.4.2. Чистая текущая стоимость (NPV)

Данный метод основан на сопоставлении дисконтированных чистых денежных поступлений от операционной и инвестиционной деятельности.

Чистая текущая стоимость является абсолютным показателем. Условием экономичности инвестиционного проекта по данному показателю является выполнение следующего неравенства: $NPV > 0$.

Чем больше NPV, тем больше влияние инвестиционного проекта на экономический потенциал предприятия, реализующего данный проект, и на экономическую ценность этого предприятия. Таким образом, инвестиционный проект считается выгодным, если NPV является положительной.

Таблица 3.17 – Расчет чистой текущей стоимости по проекту в целом

№	Наименование показателей	Шаг расчета				
		0	1	2	3	4
1	Выручка от реализации, тыс.руб	0,00	66,72	66,72	66,72	66,72
2	Итого приток	0,00	66,72	66,72	66,72	66,72
3	Инвестиционные издержки, тыс.руб.	-55,60	0,00	0,00	0,00	0,00
4	Операционные затраты, тыс. руб	0,00	30,00	30,00	30,00	30,00
4.1	Налогооб. прибыль		36,72	36,72	36,72	36,72
5	Налоги Выр-опер=донал.приб*20%	0,00	7,34	7,34	7,34	7,34
6	Итого отток	-55,60	37,34	37,34	37,34	37,34
7	Чистый денежный поток	-55,60	29,38	29,38	29,38	29,38
8	Коэффициент дисконтирования (приведения при $I = 0,20$)	1,00	0,83	0,69	0,58	0,48
9	Дисконтированный чистый денежный поток	-55,60	24,47	20,39	16,98	14,16
10	То же нарастающим итогом (NPV =20,40 тыс. руб.)	-55,60	-31,13	-10,74	6,24	20,40

Таким образом, чистая текущая стоимость по проекту в целом составляет 20,40 тыс. руб., что позволяет судить о его эффективности.

3.4.3. Дисконтированный срок окупаемости

Одним из недостатков показателя простого срока окупаемости является игнорирование в процессе его расчета разной ценности денег во времени.

Этот недостаток устраняется путем определения дисконтированного срока окупаемости.

Наиболее приемлемым методом установления дисконтированного срока окупаемости является расчет кумулятивного (нарастающим итогом) денежного потока (таблица 3.18).

Таблица 3.18 – Дисконтированный срок окупаемости

№	Наименование показателя	Шаг расчета				
		0	1	2	3	4
1.	Дисконтированный чистый денежный поток	-55,60	24,47	20,39	16,98	14,16
2.	То же нарастающим итогом	-55,60	-31,13	-10,74	6,24	20,40
3.	Дисконтированный срок окупаемости	$PP_{дск} = 1 + 31,13 / 20,39 = 2,53$ месяцев				

3.4.4. Внутренняя ставка доходности (IRR)

Для установления показателя чистой текущей стоимости (NPV) необходимо располагать информацией о ставке дисконтирования, определение которой является проблемой, поскольку зависит от оценки экспертов. Поэтому, чтобы уменьшить субъективизм в оценке эффективности инвестиций на практике широкое распространение получил метод, основанный на расчете внутренней ставки доходности (IRR).

Между чистой текущей стоимостью (NPV) и ставкой дисконтирования существует обратная зависимость. Эта зависимость следует из таблицы 3.18 и графика, представленного на рисунке 3.4.

Таблица 3.19– Зависимость NPV от ставки дисконтирования

Наименование показателя	0	1	2	3	4	NPV
Чистые денежные потоки	-55.6	29.376	29.376	29.376	29.376	
коэффициент дисконтирования						

i=0,1	1	0.909	0.826	0.751	0.683	
i=0,2	1	0.833	0.694	0.578	0.482	
i=0,3	1	0.714	0.51	0.364	0.26	
i=0,4	1	0.667	0.444	0.295	0.198	
i=0,5	1	0.625	0.39	0.244	0.095	
i=0,6	1	0.588	0.335	0.203	0.07	
Дисконтированный денежный поток						
i=0,1	-55.6	26.7	24.3	22.1	20.1	37.5
i=0,2	-55.6	24.5	20.4	17.0	14.2	20.4
i=0,3	-55.6	21.0	15.0	10.7	7.6	-1.3
i=0,4	-55.6	19.6	13.0	8.7	5.8	-8.5
i=0,5	-55.6	18.4	11.5	7.2	2.8	-15.8
i=0,6	-55.6	17.3	9.8	6.0	2.1	-20.5

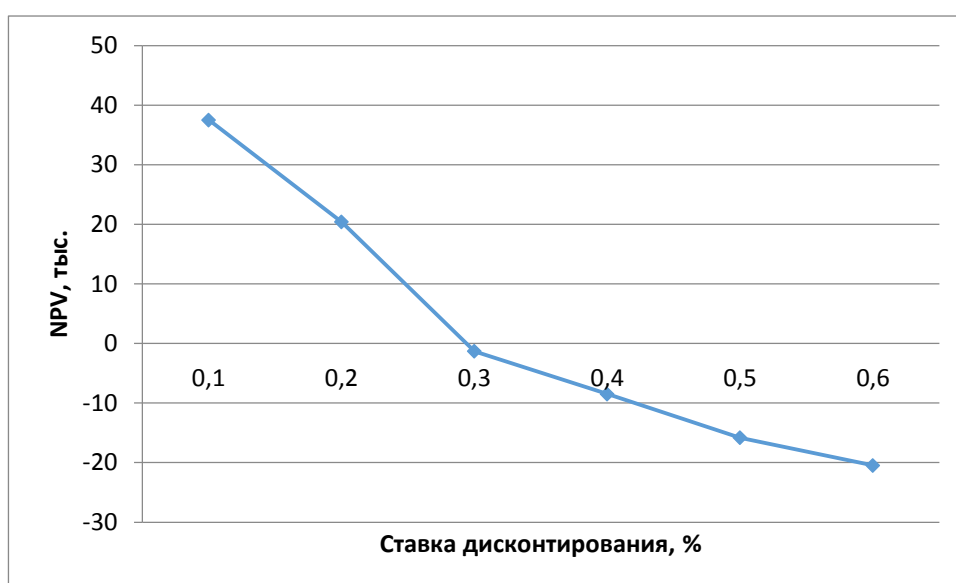


Рисунок 3.4– Зависимость NPV от ставки дисконтирования

Из таблицы и графика следует, что по мере роста ставки дисконтирования чистая текущая стоимость уменьшается, становясь отрицательной. Значение ставки, при которой NPV обращается в нуль, носит название «внутренней ставки доходности» или «внутренней нормы прибыли». Из графика получаем, что IRR составляет 0,27.

3.4.5. Индекс доходности (рентабельности) инвестиций (PI)

Индекс доходности показывает, сколько приходится дисконтированных денежных поступлений на рубль инвестиций.

Расчет этого показателя осуществляется по формуле

$$PI = \sum_{t=1}^n \frac{ЧПД_t}{(1+i)^t} / I_0,$$

где I_0 – первоначальные инвестиции.

$$PI = \frac{76,44}{55,6} = 1,37$$

$PI = 1,37 > 1$, следовательно, проект эффективен при $i=0,2$; $NPV=20,4$ тыс.

Определение эффективности происходит на основе расчета интегрального показателя эффективности научного исследования. Его нахождение связано с определением двух средневзвешенных величин: финансовой эффективности и ресурсоэффективности.

Интегральный показатель финансовой эффективности научного исследования получают в ходе оценки бюджета затрат трех (или более) вариантов исполнения научного исследования. Для этого наибольший интегральный показатель реализации технической задачи принимается за базу расчета (как знаменатель), с которым соотносятся финансовые значения по всем вариантам исполнения.

Интегральный финансовый показатель разработки определяется как:

$$I_{\text{финр}}^{\text{исп},i} = \frac{\Phi_{pi}}{\Phi_{\text{max}}}, \quad (3.10)$$

где $I_{\text{финр}}^{\text{исп},i}$ – интегральный финансовый показатель разработки;

Φ_{pi} – стоимость i -го варианта исполнения;

Φ_{max} – максимальная стоимость исполнения научно – технического проекта.

$$I_{\text{финр}}^{\text{исп},1} = \frac{55\ 556}{55\ 556} = 1$$

Для аналогов (с использованием дополнительного оборудования, стоимость которого 15000 руб. и 19000 руб.) соответственно:

$$I_{\text{финр}}^{\text{исп},2} = \frac{70\ 556}{55\ 556} = 1,27 \quad I_{\text{финр}}^{\text{исп},2} = \frac{74\ 556}{55\ 556} = 1,47$$

Таблица 3.20– Сравнительная оценка характеристик вариантов исполнения проекта

Критерии	Весовой коэффициент параметра	Исп. 1	Исп.2	Исп.3
1. Способствует росту производительности труда пользователя	0,1	5	3	4
2. Удобство в эксплуатации	0,15	4	2	3
3. Помехоустойчивость	0,15	5	3	3
4. Энергосбережение	0,2	4	3	3
5. Надежность	0,25	4	4	4
6. Материалоемкость	0,15	4	4	4
Итого	1	4,25	3,25	3,5

Интегральный показатель эффективности вариантов исполнения разработки ($I_{испi}$) определяется на основании интегрального показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{исп1} = \frac{4,25}{1} = 4,25 I_{исп2} = \frac{3,25}{1,27} = 2,56 I_{исп3} = \frac{3,5}{1,47} = 2,38$$

Сравнение интегрального показателя эффективности Исполнения 1 и Исполнений 2 и 3 позволит определить сравнительную эффективность проекта.

Сравнительная эффективность проекта:

$$\mathcal{E}_{ср} = \frac{I_{исп1}}{I_{исп2}} = \frac{4,25}{2,56} = 1,66 \mathcal{E}_{ср} = \frac{I_{исп1}}{I_{исп3}} = \frac{4,25}{2,38} = 1,79$$

Таблица 3.21– Сравнительная оценка характеристик вариантов исполнения проекта

№ п/п	Показатели	Исп. 1	Исп.2	Исп.3
1	Интегральный финансовый показатель разработки	1	1,27	1,47
2	Интегральный показатель ресурсоэффективности разработки	4,25	3,25	3,5
3	Интегральный показатель	4,25	2,56	2,38

	эффективности			
4	Сравнительная эффективность вариантов исполнения		1,66	1,79

3.5. Вывод

В процессе выполнения части работы по финансовому менеджменту, ресурсоэффективности и ресурсосбережению был проведен тщательный анализ разрабатываемого исследования. Во-первых, оценен коммерческий потенциал и перспективность проведения исследования. Полученные результаты говорят о потенциале и перспективности на уровне выше среднего. Во-вторых, проведено планирование НИР, а именно: определена структура и календарный план работы, трудоемкость и бюджет НТИ по трем исполнениям для сравнения. Результаты соответствуют требованиям к Дипломной работе по срокам и иным параметрам. В-третьих, определена эффективность исследования в разрезах ресурсной, финансовой, бюджетной, социальной и экономической эффективности. Полученные результаты и общая сравнительная эффективность позволяют сделать вывод, что исполнение является оптимальным с точки зрения использования ресурсов, затрат, бюджета и экономики.

4. СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ

В настоящее время охрана труда, окружающей среды и обеспечение безопасности трудящихся в чрезвычайных ситуациях являются важными компонентами организации рабочего процесса.

Использование компьютерных технологий сильно изменило характер труда и требования к организации и охране труда. Компьютеры уже заполнили все отрасли деятельности на предприятиях, в организациях, офисах, а также в домашних условиях. Персональные ЭВМ (ПЭВМ) как в производстве, так и в научно-исследовательских работах. При этом компьютер является источником вредного воздействия на организм человека, а, следовательно, и источником профессиональных заболеваний.

Целью данного раздела является: выявление и изучение вредных и опасных производственных факторов при работе с ПЭВМ и определение способов их устранения; оценка условия труда; рассмотрение вопросов техники безопасности, безопасности в чрезвычайных ситуациях и охраны окружающей среды.

Объектом исследования является рабочее место и помещение, в котором оно находится.

Характеристика помещения, где был разработан дипломный проект: ширина, составляет $b = 3$ м, длина комнаты $a = 5$ м, высота $h = 2,8$ м. Площадь помещения составляет $S = ab = 15 \text{ м}^2$, объем $V = abh = 42 \text{ м}^3$. В помещении одно окно, с параметрами: ширина 1,5 м, высота 1,5 м. Количество рабочих мест $n = 1$. Помещение содержит только естественную вентиляцию, т.е. воздух поступает и удаляется через дверь и окно. В зимнее время помещение отапливается, что обеспечивает достаточное, постоянное и равномерное нагревание воздуха. В помещении используется комбинированное освещение — искусственное и естественное. Искусственное освещение создается люминесцентными лампами типа ЛБ. Рабочая поверхность имеет высоту 0,75 м. Конструкция стола соответствует нормам СН 245-78, который оборудуется специальными ящиками для необходимых для работы предметов.

Электроснабжение сети переменного напряжения 220В. Помещение без повышенной опасности в отношении поражения человека электрическим током по ГОСТ 12.1.013-78.

На столе стоит компьютер, который обладает следующими характеристиками: Intel Core i7-4710HQ, оперативная память 6144 МБ, система Microsoft Windows 10, частота процессора – 2.5 ГГц, 15.6-и дюймовый монитор с разрешением 1920 на 1080 точек и дискретной видеокартой Nvidia Geforce 840m.

4.1. Анализ опасных и вредных факторов

Вредный фактор - это производственный фактор, воздействие которого на работающего в определенных условиях человека приводит к заболеванию или снижению работоспособности. В зависимости от уровня и продолжительности воздействия вредный производственный фактор может стать опасным.

Опасный фактор - это производственный фактор, воздействие которого на работающего в определенных условиях человека приводит к травме, а также к внезапно резкому ухудшению здоровья.

При работе с ПЭВМ пользователь (оператор, программист) подвергается воздействию следующих вредных и опасных производственных факторов:

- шуму;
- электростатических полей;
- электромагнитных полей;
- интенсивной напряженности трудового процесса.

Последствия действия этих факторов могут привести к ухудшению здоровья пользователя, а также к профессиональным заболеваниям. Помимо этого, неудобная рабочая поза, длительное сосредоточенное наблюдение, вызывают повышенное напряжение мышц зрительного аппарата. Совместное влияние неблагоприятных производственных факторов способствуют развитию общего утомления и снижению работоспособности.

Комбинации вредных и опасных производственных факторов, наблюдаемых при работе с ПЭВМ:

1. Монитор компьютера является источником: электромагнитного поля (ЭМП); электростатического поля; рентгеновского излучения; вредного действия светового потока и отраженного света.

2. В результате несовершенства создания изображения на экране зрительный аппарат подвергается значительной нагрузке.

3. Работа компьютера сопровождается акустическими шумами.

4. Несоблюдение эргономических параметров, обеспечивающих безопасность приёмов работы пользователя ПЭВМ: гигиенических и психофизиологических; антропометрических; эстетических может повлечь снижение эффективности действий человека.

Анализ санитарно-технических условий помещения (площадь и объём производственных помещений, освещённости, микроклимата, вентиляции, шумов, излучения), оценка вредных факторов и приведение их к требованиям нормативных документов могут обеспечить безопасные условия труда.

4.2. Производственная безопасность

4.2.1. Характеристика помещения, микроклимат

Для людей, работающих непосредственно с ЭВМ, необходимо, чтобы условия труда на их рабочем месте соответствовали требованиям безопасности труда, а также психофизиологическим и санитарно-гигиеническим нормам.

Согласно санитарным нормам СанПин 2.2.2.542-96, на каждого работника должно быть выделено не менее 6 м^2 площади и не менее 20 м^3 объема, при высоте помещения – 2,8 м, расстояние от глаз до клавиатуры, экрана и документов было 500 – 600мм, клавиатура находилась на высоте 740 – 790мм от уровня пола, верхний край дисплея находился на уровне прямого взгляда, а нижний – не ниже 40° от прямого взгляда.

Площадь помещения $S = 15\text{ м}^2$.

Количество рабочих мест $n = 1$. Тогда на одного работающего приходится

$$\frac{S}{n} = \frac{15}{1} = 15 \text{ м}^2.$$

Объем помещения на одного человека составит:

$$V_1 = \frac{abh}{n} = \frac{3*5*2,8}{1} = 42 \text{ м}^3.$$

Согласно полученным результатам, можно сделать вывод, что по занимаемой площади и по занимаемому объему помещение полностью удовлетворяет нормативным требованиям.

4.2.2. Характеристика помещения, микроклимат

Микроклимат производственных помещений – это климат внутренней среды помещений, который определяется действующими на организм человека сочетаниями температур воздуха и поверхностей, относительной влажности воздуха, скорости движения воздуха и интенсивности теплового излучения. Показатели микроклимата должны обеспечивать сохранение теплового баланса человека с окружающей средой и поддержание оптимального или допустимого теплового состояния организма. Вычислительная техника является источником существенных тепловыделений, что может привести к повышению температуры и снижению относительной влажности в помещении. В помещениях, с установленными компьютерами, должны соблюдаться определенные параметры микроклимата. В санитарных нормах СН-245-71 установлены величины параметров микроклимата, которые создают комфортные условия. Эти нормы зависят от времени года, характера трудового процесса и характера производственного помещения. Данные показатели оказывают большое влияние на функциональную деятельность человека, его самочувствие, здоровье, надежность работы.

Допустимые условия микроклимата согласно СанПиН 2.2.2.542–96 «Оптимальные нормы микроклимата для помещений с ВТД и ПЭВМ»

представлены в таблице 4.1. Нормы подачи свежего воздуха в помещения, где расположены компьютер, приведены в таблице 4.2.

Таблица 4.1 - Оптимальные параметры микроклимата для помещений, где установлены компьютеры

<i>Период года</i>	<i>Категория работы</i>	<i>Температура, °C</i>	<i>Относительная влажность, %</i>	<i>Скорость движения воздуха, м/с</i>
<i>Холодный</i>	<i>средняя</i>	<i>22–24</i>	<i>40–60</i>	<i>0.1</i>
<i>Теплый</i>	<i>средняя</i>	<i>23–25</i>	<i>40–60</i>	<i>0.2</i>

Таблица 4.2 - Нормы подачи свежего воздуха в помещения, где расположены компьютеры

Характеристика помещения	Объемный расход воздуха подаваемого в помещение свежего воздуха, м ³ /на одного
Объем до 20 м ³ на человека	Не менее 30 м ³
20...40 м ³ на человека	Не менее 20 м ³
Более 40 м ³ на человека	Естественная вентиляция

Анализ микроклимата в помещении, где был разработан проект выпускной квалификационной работы, показал, что помещение соответствует допустимым условиям. В данном помещении присутствует только естественная вентиляция, т.е. воздух поступает и удаляется через дверь и окно. Согласно расчетам естественная вентиляция на одного работающего в данном помещении составляет 42 м³, что полностью соответствует установленным нормам (табл. 4.2). В зимнее время помещение отапливается, что обеспечивает достаточное, постоянное и равномерное нагревание воздуха.

4.2.3. Производственный шум

Шум превышающий пределы громкости и частоты звуковых колебаний, представляют собой профессиональную вредность. Шум — это сочетание звуков различной интенсивности и частоты, которое оказывает раздражающее и

вредное действие на организм человека. Под влиянием шума у человека может изменяться кровяное давление, работа желудочно-кишечного тракта, а длительное его действие в ряде случаев приводит к частичной или полной потере слуха. Шум влияет на производительность труда рабочих, ослабляет внимание, вызывает тугоухость и глухоту, раздражает нервную систему, в результате чего снижается восприимчивость к сигналам опасности, что может привести к несчастному случаю. Шумы на рабочих местах нормируются по ГОСТ 12.1.003-99.

Уровень шума на рабочем месте математиков-программистов и операторов видеоматериалов не должен превышать 50 дБА, а в залах обработки информации на вычислительных машинах - 65 дБА. Защита от шумов – заключение вентиляторов в защитный кожух и установление их внутри корпуса ЭВМ. Для снижения уровня шума стены и потолок помещений, где установлены компьютеры, могут быть облицованы звукопоглощающими материалами с максимальными коэффициентами звукопоглощения в области частот 63 - 8000 Гц.

Согласно ГОСТ 12.1.003-83 при разработке технологических процессов, проектировании, изготовлении и эксплуатации машин, производственных зданий и сооружений, а также при организации рабочих мест следует принимать все необходимые меры по снижению шума, воздействующего на человека, до значений, не превышающих допустимые. Защита от шума должна обеспечиваться разработкой шумобезопасной техники, применением средств и методов коллективной защиты, в том числе строительно-акустических, применением средств индивидуальной защиты. В первую очередь следует использовать средства коллективной защиты. По отношению к источнику возбуждения шума коллективные средства защиты подразделяются на средства, снижающие шум в источнике его возникновения, и средства, снижающие шум на пути его распространения от источника до защищаемого объекта.

Средства индивидуальной защиты (СИЗ) применяются в том случае, если другими способами обеспечить допустимый уровень шума на рабочем месте не

удается. Принцип действия СИЗ – защитить наиболее чувствительный канал воздействия шума на организм человека – ухо. Применение СИЗ позволяет предупредить расстройство не только органов слуха, но и нервной системы от действия чрезмерного раздражителя.

Наиболее эффективны СИЗ, как правило, в области высоких частот. СИЗ включают в себя противозумные вкладыши (беруши), наушники, шлемы и каски, специальные костюмы.

4.2.4. Освещенность рабочей зоны

Освещение – использование световой энергии солнца и искусственных источников света для обеспечения зрительного восприятия окружающего мира.

От степени освещенности напрямую зависит не только здоровье глаз и работоспособность человека, но еще и его физическое и психоэмоциональное состояние. Причем в помещениях различного назначения требования по освещенности должны различаться.

Различают следующие виды освещения:

естественное - обуславливается световым потоком прямых солнечных лучей и диффузионным светом неба, т. е. многократным отражением солнечных лучей от мельчайших взвешенных в атмосфере частиц пыли и воды. Величина такого освещения изменяется как в течение года, так и в течение суток. Изменения в течение суток зависят от географических координат, расположения других зданий, месторасположения объекта, времени суток, прозрачности воздуха, облачности и других характеристик окружающей среды;

искусственное - обеспечивается электрическими источниками света. Оно применяется в темное время суток, а также днем, при недостаточном естественном освещении;

комбинированное - включает в себя как естественное, так и искусственное освещение.

В данном помещении используется комбинированное освещение — искусственное и естественное. Естественное освещение проникает в помещение

через окно, но коэффициент естественного освещения не соответствует норме, поэтому применяется искусственное освещение. Искусственное освещение устанавливается во всех основных и вспомогательных помещениях производственных зданий в соответствии со СНиП 23-05-95 (табл.4.3).

Для того чтобы естественное освещение удовлетворяло СанПиН 2.2.2.542 – 96, достаточно, чтобы площадь световых проемов помещения соответствовало $1/6 – 1/8$ от площади пола.

Помещение имеет следующие размеры: длина комнаты $a = 5$ м, ширина $b = 3$ м, высота $h = 2,8$ м, количество окон – 1. Размеры светопроема в данном помещении: ширина $b_o = 1,5$ м, высота $h_o = 1,5$ м.

Площадь пола $S_n = ab = 15 \text{ м}^2$.

Площадь светопроемов $S_{ok} = 0,9b_o h_o = 2,025 \text{ м}^2$.

Соотношение площади светопроемов к площади пола

$$\frac{S_{ok}}{S_n} = \frac{2,025}{15} = 0,135,$$

что соответствует санитарным нормам.

Искусственное освещение создается люминесцентными лампами типа **ЛБ-40 (белый свет)**, длина лампы – 1200 мм, ширина лампы – 40 мм в количестве 10 штук (Приложение Б).

Проверочный расчет существующей искусственной освещенности проведем методом коэффициента использования. Сначала определяется индекс помещения (i).

$$i = \frac{S}{(a+b)h_1},$$

где S – площадь помещения;

h_1 – расчетная высота подвеса светильника, м;

a и b – длина и ширина помещения, м.

Высота подвеса светильника h_1 определяется как расстояние от светильника до рабочей поверхности. Рабочая поверхность имеет высоту 0,75м.

Соответственно $h_1 = h - 0,75 = 2,05$ м.

Следовательно $i = \frac{15}{8 * 2,05} = 0,91$.

Световой поток лампы определяется по формуле:

$$\Phi = \frac{E_n \cdot S \cdot K_z \cdot Z}{N \cdot \eta},$$

где E_n – нормируемая минимальная освещённость по СНиП 23-05-95, лк; S – площадь освещаемого помещения, m^2 ; K_z – коэффициент запаса, учитывающий загрязнение светильника (источника света, светотехнической арматуры, стен и пр., т. е. отражающих поверхностей), наличие в атмосфере цеха дыма, пыли; Z – коэффициент неравномерности освещения, отношение E_{cp}/E_{min} . Для люминесцентных ламп при расчётах берётся равным 1,1; N – число ламп в помещении; η – коэффициент использования светового потока.

Коэффициент использования светового потока показывает, какая часть светового потока ламп попадает на рабочую поверхность. Он зависит от индекса помещения i , типа светильника, высоты светильников над рабочей поверхностью h и коэффициентов отражения стен ρ_c и потолка ρ_n .

Данное помещение со средним выделением пыли, поэтому $K_z = 1,8$; состояние потолка – свежепобеленный, поэтому $\rho_n = 70$; состояние стен – побеленные бетонные стены, поэтому $\rho_c = 50$. Коэффициент использования светового потока, учитывая индекс помещения $\eta = 33$. Нормируемая минимальная освещенность, согласно табл.3 при использовании ЭВМ и одновременной работе с документами должна быть равна 300лк.

$$\Phi = \frac{300 * 15 * 1,8 * 1,1}{0,33 * 10} = 2700 \text{ Лм}$$

Для люминесцентных ламп с мощностью 40Вт и напряжением сети 220В, со стандартный световой поток ЛТБ равен 2850 Лм.

$$-10\% \leq \frac{\Phi_{л.станд} - \Phi_{л.расч}}{\Phi_{л.станд}} 100\% \leq +20\%$$

$$-10\% < 5,26\% < 20\%$$

Согласно полученным результатам анализа освещенности помещения, данная система освещения удовлетворяет нормативным требованиям.

Таблица 4.3 - Нормирование значения освещенности на рабочих местах производственных помещений при искусственном освещении, согласно СНиП 23-05-95

Характеристика зрительной работы	Наименьший Размер объекта, мм	Подразряд зрительной работы	Контраст объекта различения с фоном	Характеристика фона	Освещенность, лк		
					Комбинированное освещение	Общее освещение	
средняя точность (4 разряд зрительной работы)	0,5-1,0	а	Малый	Темный	750	300	
			Малый	Средний	500	200	
		б	Средний	Темный	500	200	
			в	Малый	Светлый	400	200
				Средний	Светлый	400	200
		г	Большой	Темный	400	200	
			Средний	Светлый	300	150	
				Большой	Светлый	300	150
			Большой	Средний	300	150	

4.3. Рабочее место, оценка условий труда на рабочем месте

Рабочее место должно удовлетворять следующим требованиям:

- Рабочий стол должен регулироваться по высоте в пределах 680–760 мм, при отсутствии регулировки, его высота должна быть 720 мм. Под столешницей стола должно быть свободное пространство для ног, с размером по высоте не менее 600 мм. Рабочий стул должен быть снабжен подъемным поворотным устройством, обеспечивающим регулировку высоты, сидения и спинки. Поверхность сидения и спинки должна быть полумягкой, с

нескользящим, не электризующимся и воздухопроницаемым покрытием.

- Яркость свечения экрана монитора должна быть не менее 100кд/м², контрастность изображения знака “-” не менее 0,8, частота регенерации не менее 72Гц, низкочастотное дрожание изображения в диапазоне 0,05 - 1,0Гц должно находиться в пределах 0,1мм. Экран должен иметь антибликовое покрытие, а расстояние от глаза оператора до экрана должно быть в пределах 40 - 80см. Клавиатура не должна быть жестко связана с монитором, ее наклон должен находиться в пределах 10 - 15градусов. Видеомонитор должен быть оборудован поворотной площадкой.
- При работе с текстовой информацией для комфортного состояния глаз наиболее применимым является предъявление черных знаков на светлом фоне (черных на белом).

Условия труда характеризуются тяжестью труда, которая представляет собой совокупное воздействие всех элементов, составляющих условия труда, на работоспособность человека, его здоровье, жизнедеятельность и восстановление работоспособности.

При работе с ПЭВМ виды трудовой деятельности разделяются на 3 группы, согласно СанПиН 2.2.2.542-96:

А – работа по считыванию информации с экрана ПЭВМ с предварительным запросом;

В – работа по вводу информации;

С – творческая работа в режиме диалога с ПЭВМ.

Для видов трудовой деятельности устанавливается 3 категории тяжести и напряженности работы с ПЭВМ, которые приведены в **таблице 4.4.**

Таблица 4.4 - Категории тяжести и напряженности работ с ПЭВМ

Категория работы	Уровень нагрузки за рабочую смену
------------------	-----------------------------------

с ПЭВМ	Группа А, количество знаков	Группа В, количество знаков	Группа С, час
1	До 20 000	До 15 000	До 2
2	До 40 000	До 30 000	До 4
3	До 60 000	До 40 000	До 6

Характеристики рабочего места: высота рабочего стола - 750мм; рабочий стул снабжен поворотным и подъемным устройствами; монитор полностью удовлетворяет требованиям.

Характеристики рабочей смены: работа выполнялась каждый день, категория работы с ПЭВМ - 1. Таким образом, рабочее место и условия труда полностью удовлетворяют установленным требованиям.

4.3.1. Психофизиологические факторы

Умственные нагрузки и интеллектуальная работа могут вызвать переутомление функционального состояния центральной нервной системы, нервно-мышечного аппарата рук, а некомфортная поза вызывает повышенное позвоночное напряжение мышц и обуславливает развитие общего утомления и снижение работоспособности.

Установленный режим труда и отдыха, основанный на соблюдении регламентируемых перерывов, позволяет снизить психофизиологическую напряженность, и организовать динамику функционального состояния систем организма.

Режим труда и отдыха работника: при вводе данных, редактировании программ, чтении информации с экрана непрерывная продолжительность работы не должна превышать 4-х часов при 8-часовом рабочем дне. Через каждый час работы необходимо делать перерыв на 5-10 минут, а через два часа на 15 минут.

Также, для снижения и устранения нервно-психологического, зрительного и мышечного напряжения, необходимо проводить комплекс физических

упражнений и сеансы психофизической разгрузки и снятия усталости во время перерывов, и после окончания рабочего дня (занятие спортом или смена деятельности).

4.3.2. Анализ опасных производственных факторов и обоснование мероприятий по их устранению (техника безопасности)

Основные опасные факторы, которые могут возникнуть при работе с компьютером представлены в **таблице 4.5**.

Таблица 4.5 - Основные элементы производственного процесса, формирующие опасные факторы

Наименование видов работ и параметров производственного процесс	Факторы (ГОСТ 12.0.003-74 ССБТ)	Нормативные документы
Оператор ПК	Электрический ток	ГОСТ 12.1.002–84

4.4. Электробезопасность

Действие электрического тока на живую ткань носит разносторонний и своеобразный характер. Проходя через организм человека, электроток производит термическое, электролитическое, механическое и биологическое действия.

Термическое действие тока проявляется ожогами отдельных участков тела, функциональными расстройствами органов. Электролитическое действие выражается в разложении органической жидкости, в том числе крови, в нарушении ее физико-химического состава. Механическое действие приводит к расслоению, разрыву тканей организма, мгновенного взрывоподобного образования пара из тканевой жидкости и крови. Биологическое действие

проявляется раздражением и возбуждением живых тканей организма, нарушением внутренних биологических процессов.

Наиболее опасным для человека является переменный ток с частотой 20 – 100 Гц. Опасной величиной тока является ток, равный 0,001 А, а смертельной 0,1 А. Также исход электропоражения зависит от состояния внешней среды.

В отношении опасности поражения людей электрическим током согласно правилам устройства электроустановок (ПУЭ) различаются:

1) помещения без повышенной опасности, в которых отсутствуют условия, создающие повышенную или особую опасность;

2) помещения с повышенной опасностью;

3) особо опасные помещения;

4) территория открытых электроустановок в отношении опасности поражения людей электрическим током приравнивается к особо опасным помещениям.

Помещение, в котором выполнялась выпускная квалификационная работа, относится к категории помещений без повышенной опасности, поскольку помещение является сухим и хорошо отапливаемым, полы токонепроводящие, средняя температура от 18 до 20°, влажность от 40 до 50%. Однако, возможность поражения электрическим током все-таки существует, т.к. ЭВМ работает от источника тока. Для исключения опасности поражения электрическим током важно соблюдать правила электрической безопасности:

- Перед включением компьютера в сеть необходимо визуально проверить электропроводку на отсутствие видимых нарушений изоляции, а также на отсутствие замыкания токопроводящих частей на корпус компьютера;
- При появлении признаков замыкания необходимо немедленно отключить от электрической сети ЭВМ и устранить неисправность;
- Запрещается при включенной ЭВМ одновременно прикасаться к приборам, имеющим естественное заземление.

Для защиты от поражения электрическим током в случае повреждения изоляции должны быть применены по отдельности или в сочетании следующие меры защиты при косвенном прикосновении:

- защитное заземление;
- автоматическое отключение питания;
- уравнивание потенциалов;
- выравнивание потенциалов;
- двойная или усиленная изоляция;
- сверхнизкое (малое) напряжение;
- защитное электрическое разделение цепей;
- изолирующие (непроводящие) помещения, зоны, площадки.

Требования к устройству защитного заземления и зануления определены Правилами устройства электроустановок (ПУЭ), в соответствии с которыми защитному заземлению или занулению подлежат все металлические и другие токопроводящие части электроустановок и оборудования, которые случайно в аварийном режиме могут оказаться под напряжением (ССБТ ГОСТ 12.1.030-81):

- при номинальном напряжении 380 В и выше переменного тока, 440 В и выше постоянного тока – во всех электроустановках;
- при номинальном напряжении выше 42 В, но ниже 380 В переменного тока и выше 110 В, но ниже 440 В постоянного тока – только в помещениях с повышенной опасностью, особо опасных помещениях и в наружных электроустановках;
- во взрывоопасных помещениях необходимо заземлять всеоборудование независимо от напряжения.

Наибольшие допустимые значения сопротивления заземления растеканию тока, установленные правилами устройства электроустановок, составляют для установок до 1000В 4 Ом – при суммарной мощности генераторов или трансформаторов, питающих данную сеть, не более 100 кВА;

Поэтому к работе на ПК допускаются лица, прошедшие обучение безопасным методам труда, вводный инструктаж, первичный инструктаж на рабочем месте. Каждый работник должен знать правила первой медицинской помощи при поражении электрическим током, для того, чтобы быть готовым оказать помощь другим работникам.

4.5. Экологическая безопасность

Функционирование любого производства (промышленного, сельскохозяйственного и т. п.) сопровождается образованием отходов. Они поступают в окружающую среду в виде выбросов в атмосферу, сбросов в водоемы, твердых промышленных и бытовых отходов и мусора на поверхность и в недра Земли. В наше время обеспечение безопасности жизнедеятельности человека неразрывно связано с решением задач по охране природной среды.

Экологическая безопасность на основе изучения закономерностей взаимодействия природы и человека дает научно обоснованные рекомендации по охране природы, природопользованию и воспроизведению природных ресурсов. Охрану природы можно - это комплекс государственных, международных и общественных мероприятий, направленных на рациональное использование природы, восстановление, улучшение и охрану природных ресурсов.

Для уменьшения количества выбросов и отходов предприятия внедряют новейшие технологии, электрооборудование, ПЭВМ в процесс эксплуатации, отчистки и утилизации отходов производства. Однако все большее внедрение и применение ПЭВМ приводит к увеличению затрат электроэнергии, количества электростанций и их мощностей. Это может привести к различным экологическим нарушениям: глобальное потепление, загрязнение атмосферы, воды и почвы вредными веществами и т.д. При написании данной выпускной квалификационной работы были использованы следующие ресурсы:

1. бумага;

2. канцелярские принадлежности, а именно ручки, сделанные из пластмассы;

3. компьютер;

При написании ВКР негативного воздействия на окружающую среду не оказывалось, т.к. работа проводилась за ПЭВМ. Никаких выбросов и отходов в окружающую среду, чрезвычайных ситуаций при работе с компьютером не было.

Среди основных методов утилизации бумажных отходов, которые применяются на сегодняшний день следует отметить следующие:

1. Сжигание документов.
2. Шредирование.
3. Закапывание.
4. Химическая обработка.

Переработка ресурсов, использованных при выполнении ВКР, включает в себя несколько этапов:

1. Удаление всех опасных компонентов.

2. Удаление всех крупных пластиковые части. В большинстве случаев операция 2 также осуществляется вручную. Пластик сортируется в зависимости от типа и измельчается для того, чтобы в дальнейшем его можно было использовать повторно. Оставшиеся после разборки части отправляют в большой измельчитель, и все дальнейшие операции автоматизированы.

3. Измельченные в гранулы остатки компьютеров подвергаются сортировке. Сначала с помощью магнитов извлекаются все железные части. Затем приступают к выделению цветных металлов, которых в ПК значительно больше. В сухом остатке получается смесь пластика и меди. Медь выделяют способом флотации – гранулы помещают в специальную жидкость, пластик всплывает, а медь остается на дне.

4.6. Пожарная безопасность

Пожары представляют особую опасность, так как несут за собой большие материальные потери, а также огромную угрозу здоровью и жизни людей.

Возникновение пожара при работе с электронной аппаратурой может быть по причинам электрического и неэлектрического характера.

Причины возникновения пожара неэлектрического характера:

а) халатное неосторожное обращение с огнем (курение, оставленные без присмотра нагревательные приборы, использование открытого огня);

б) самовоспламенение и самовозгорание веществ.

Причины возникновения пожара электрического характера:

а) короткое замыкание;

б) перегрузки по току;

в) искрение и электрические дуги;

г) статическое электричество и т. п.

Существуют 4 класса пожароопасности помещений (А, Б, В, Г, Д). Данное помещение относится к классу В - помещения, в которых находятся (обращаются) горючие и трудногорючие жидкости, твердые горючие и трудногорючие вещества и материалы (в том числе пыли и волокна), вещества и материалы, способные при взаимодействии с водой, кислородом воздуха или друг с другом только гореть, т.к. в помещении находились деревянные столы, стулья и бумага.

Меры пожарной профилактики:

- **Строительно-планировочные.** Строительно-планировочные меры определяются огнестойкостью зданий и сооружений (выбор материалов конструкций по степени огнестойкости). В зависимости от степени огнестойкости определяются наибольшие дополнительные расстояния от выходов для эвакуации при пожарах.

- **Технические.** Использование разнообразных защитных систем; соблюдение параметров технологических процессов и режимов работы оборудования;
- **Организационные.** Представляют собой проведение инструктажа персонала по пожарной безопасности, соблюдение мер пожарной безопасности. Для предупреждения возникновения пожара необходимо соблюдать следующие правила пожарной безопасности: правильная эксплуатация оборудования, правильное содержание зданий и территорий, противопожарный инструктаж рабочих и служащих, обучение производственного персонала правилам противопожарной безопасности, издание инструкций, плакатов, наличие плана эвакуации, соблюдение противопожарных правил, норм при проектировании зданий, при устройстве электропроводов и оборудования, отопления, освещения, правильное размещение оборудования, своевременный профилактический осмотр, ремонт и испытание оборудования (Приложение В).

Помещение, где была сделана выпускная квалификационная работа, снабжено противопожарной защитой (огнетушители порошковые ОП-3,), направленной на предотвращение воздействия на людей опасных факторов пожара и ограничения материального ущерба от него.

Порошковый огнетушитель предназначен для применения в бытовых условиях в качестве первичного средства тушения пожаров класса А (твердых веществ), С (горючих газов), В (горючих жидкостей или плавящихся твердых тел) и электроустановок, находящихся под напряжением до 1000 В (в зависимости от марки применяемого огнетушащего порошка) и имеет следующие характеристики:

Длина выброса - 3,8 м.

Огнетушащая способность - 2А;89В, (2.80).

Габариты - 450*320*300 мм.

Масса с зарядом - 5 кг

Основы противопожарной защиты предприятий определены в стандартах ГОСТ 12.1.004-76 и ГОСТ 12.1.010-76.

4.7. Чрезвычайные ситуации

Чрезвычайная ситуация - это обстановка, сложившаяся на определенной территории или акватории в результате аварии, опасного природного явления, катастрофы, стихийного или иного бедствия, которые могут повлечь или повлекли за собой человеческие жертвы, ущерб здоровью людей или окружающей природной среде, значительные материальные потери и нарушение условий жизнедеятельности людей.

Чрезвычайные ситуации делятся на:

- природные
- техногенные
- социальные
- экологические

Для Томской области характерны **сильные штормы**, поэтому с поступлением штормового предупреждения, которое подает подразделение Гидрометеослужбы за несколько часов до наступления опасного метеоявления, проводятся оперативные защитные мероприятия:

1) широкое оповещение населения о пути следования и времени подхода к различным районам опасного метеорологического явления, о возможном характере его воздействия, мерах безопасности и правилах поведения людей, оптимальных для складывающейся ситуации;

2) переход к безопасным режимам работы производств, прекращение строительно-монтажных работ с применением подъемных механизмов (башенных кранов), прекращение погрузоразгрузочных работ с применением подъемных механизмов (автокранов, порталных кранов, козловых и башенных);

3) перевод и перемещение в прочные или защищенные помещения уникального и особо ценного оборудования, в сельской местности – подвоз запаса кормов к фермам, создание запаса воды и т.д.

Для Томской области также характерны низкие среднемесячные (от минус 50 до минус 20 °С) температуры в зимние месяцы. Холодная среда определяется условиями, которые вызывают большие, чем обычно потери тепла организмом.

При работе во время такой природной чрезвычайной ситуации, как сильные морозы, предприятие должно проводить специальное мероприятие, что работники работали в комфортных для себя условиях. Для этого необходимо проводить следующие мероприятия:

- 1) оборудовать все рабочие помещения на предприятии специальными обогревателями;
- 2) предоставить перевозку сотрудников от дома до места работы и обратно;
- 3) хранить в аптечке специальные медицинские препараты, если в случае данной природной ЧС состояние сотрудников ухудшится;
- 4) проведение семинаров, на которых сотрудники будут ознакомлены с тем, как вести себя в случае возникновения сильных морозов;
- 5) установка бензоэлектростанции на предприятии;
- 6) сделать утепление стен с помощью теплоизолирующих материалов.

Попытки несанкционированного проникновения посторонних.

Для предотвращения несанкционированного проникновения посторонних лиц на предприятие необходимо осуществление следующих мероприятий:

- 1) организация пропускной системы попадания на предприятие;
- 2) организация системы видеонаблюдения на предприятии;
- 3) соблюдение тайны о работниках предприятия;
- 4) издание (тиражирование) инструкций по охране труда;
- 5) наем охранной организации для контроля порядка на предприятии;
- 6) проведение семинаров, на которых сотрудники будут ознакомлены с тем, как вести себя в случае проникновения несанкционированных лиц на предприятие;
- 7) составление правил поведения на рабочем месте.

4.8. Выводы и рекомендации

Проанализировав условия труда на рабочем месте, где была разработана выпускная квалификационная работа, можно сделать вывод, что помещение удовлетворяет необходимым нормам и в случае соблюдения техники безопасности и правил пользования компьютером работа в данном помещении не приведет к ухудшению здоровья работника.

Также, действие вредных и опасных факторов сведено к минимуму, т.е. микроклимат, шум, электробезопасность и пожаробезопасность соответствуют требованиям, предъявленным в соответствующих нормативных документах. При этом, система освещения в помещении соответствует норме и создает нормальные условия для работы.

Анализ экологической безопасности показал, что деятельность помещения не приносит вред окружающей среде.

Важно добавить, что монитор компьютера служит источником ЭПМ – вредного фактора, который отрицательно влияет на здоровье работника при продолжительной непрерывной работе и приводит к снижению работоспособности. Поэтому во избежание негативного влияния на здоровье необходимо делать перерывы при работе с ЭВМ и проводить специализированные комплексы упражнений для глаз.

4.9. Законодательная и нормативная база

Для разработки комплекса решений для обеспечения безопасности пользователя рабочего места были использованы следующие нормативно-правовые акты:

- ГОСТ 12.0.003-74 (1999) ССБТ. Опасные и вредные производственные факторы классификации;
- ГОСТ 12.1.003-83 (1991) ССБТ. Шум. Общие требования безопасности;
- ГОСТ 12.1.006-84 (1999) ССБТ. Электромагнитные поля радиочастот. Допустимые уровни на рабочих местах и требования к проведению контроля;
- ГОСТ 12.1.009-76 (1999) ССБТ. Электробезопасность. Термины и

определения;

- ГОСТ 12.1.019-79 (1996) ССБТ. Электробезопасность. Общие требования и номенклатура видов защиты;
- ГОСТ 12.1.036-81 (1996) ССБТ. Шум. Допустимые уровни в жилых и общественных зданиях;
- СН 2.2.4/2.1.8.562-96 Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки;
- СНиП 23-05-95 Естественное и искусственное освещение;
- СН-3223-85. Санитарные нормы допустимых уровней шума на рабочих местах;
- СН 2.2.4/2.1.8.562-96 Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки;
- СНиП 23-03-03 Защита от шума.

ЗАКЛЮЧЕНИЕ

В работе проведен многомерный статистический анализ СЭП предприятий регионов России подраздела промышленности DM в РС (58 регионов) и ИСС (18 регионов).

В ходе данного исследования проведены корреляционный, факторный, кластерный и дисперсионный анализы. На основании данных анализов можно сделать следующие выводы:

1) Построено 6-мерное факторное экономическое пространство регионов России на основании корреляционного и факторного анализов социально-экономических показателей развития промышленности предприятий подраздела DM регионов России;

2) В результате проведения статистического сравнительного анализа результативности развития предприятий промышленности регионов в РС и ИСС в рамках дисперсионного анализа соответствующих факторных экономических пространств, было выявлено, что предприятия ИСС выпускают больше продукции, чем предприятия РС, несмотря на то, что количество занятого населения в основном больше на предприятиях РС; средняя ЗП выше на предприятиях ИСС, чем на предприятиях РС.

3) Созданные иностранные предприятия, ориентированные на российский рынок, сократили занятость предприятий РС в подразделе DM. Модели создания крупных зарубежных предприятий (Санкт-Петербург и Калужская область) были неэффективны для России с точки зрения социальных результатов. Тем не менее, они могут генерировать другие положительные эффекты (инвестиции, высококачественная продукция, налоги и т. д.).

Данная модель оценивания различий предприятий регионов в РС и ИСС может быть применена к любой базе экономических показателей, а также учтена при разработке мер государственного стимулирования.

СПИСОК ПУБЛИКАЦИЙ СТУДЕНТА

1. Ivankina L. I. , Mikhalchuk A. A. , Arefjev V. P. , Filipenko N. M. , Novoseltseva D. A. Multidimensional Statistical Analysis of Life Satisfaction of Older People // The European Proceedings of Social and Behavioural Sciences. - 2017 - Vol. 19. - p. 470-475
2. Spitsyn V. V. , Mikhalchuk A. A. , Novoseltseva D. A. , Boznyakov A. V. , Spitsyna L. Y. , Antonova I. S. Domestic and foreign firms in Russian food industry for the period of 2005-2014 // International Journal of Economics and Financial Issues. - 2016 - Vol. 6 - №. 4. - p. 1413-1418
3. Spitsyn V. V. , Pokrovskaya N. V. , Mikhalchuk A. A. , Novoseltseva D. A. , Bannova K. A. , Gumennikov I. V. , Spitsyna L. Y. Comparative Analysis of the Tax Burden on Innovative and Non-Innovative Sectors of Manufacturing in Russia // Advances in Computer Science Research. - 2016 - Vol. 51. - p. 349-353
4. Spitsyn V. V. , Mikhalchuk A. A. , Spitsyna L. Y. , Shabaldina N. V. , Novoseltseva D. A. , Shinkeev M. L. Product Innovation Efficiency of Russian Electronic Industry: DEA Approach and Cluster Analysis // Advances in Social Science, Education and Humanities Research. - 2016 - Vol. 53. - p. 607-610
5. Spitsyn V. V. , Tyuleneva N. A. , Mikhalchuk A. A. , Novoseltseva D. A. , Spitsyna L. Y. , Dyrina E. N. Comparison of social results produced by Russian and foreign companies: case Russian electronic industry // Advances in Social Science, Education and Humanities Research. - 2015 - Vol. 28. - p. 105-109
6. Spitsyn V. V. , Mikhalchuk A. A. , Spitsyna L. Y. , Novoseltseva D. A. , Shabaldina N. V. Clustering of Russian Regions on Indicators of the Product Innovation Efficiency: Case Subsection DL "Manufacture of Electrical and Optical Equipment" // The 26th International Business Information Management Association Conference: Proceedings IBIMA, Madrid, November 11-12, 2015. - СОЕДИНЕННЫЕ ШТАТЫ АМЕРИКИ: IBIMA, 2015 - p. 589-596
7. Spitsyn V. V. , Mikhalchuk A. A. , Spitsyna L. Y. , Shabaldina N. V. , Novoseltseva D. A. Comparative Analysis of Economic and Social Results of Foreign and Domestic Firms: Case Russian Electronic Industry // The 26th International Business Information Management Association Conference: Proceedings IBIMA, Madrid, November 11-12, 2015. - СОЕДИНЕННЫЕ ШТАТЫ АМЕРИКИ: IBIMA, 2015 - p. 669-675
8. Spitsyn V. V. , Mikhalchuk A. A. , Spitsyna L. Y. , Tyuleneva N. A. , Novoseltseva D. A. Social Results of Domestic and Foreign Firms: Case Manufacture of Transport Equipment in Russia // International Journal of Economics and Financial Issues. - 2016 - Vol. 6 - №. 1. - p. 147-153
9. Spitsyn V. V. , Mikhalchuk A. A. , Spitsyna L. Y. , Akerman E. N. , Tyuleneva N. A. , Semes A. V. , Novoseltseva D. A. Comparative analysis of salary, labor intensity and payroll-output ratio of foreign and domestic firms: case Russian vehicle industry // The 2nd International Conference on Education Reform and Modern Management (ERMM2015) : Proceedings. Part of the series ASSEHR.. - Париж : Atlantis Press. - 2015 - Vol. 15. - p. 371-374

10. Барышева Г. А., Михальчук А. А., Бурец Ю. С., Новосельцева Д. А. Факторная динамическая модель инновационного процесса в регионах АИРР // Инновационное развитие экономики. - 2016 - №. 4. - С. 7-16
11. Трифонов А. Ю., Спицын В. В., Михальчук А. А., Рыжкова М. В., Новосельцева Д. А., Гуменников И. В. Развитие пищевой промышленности России в разрезе форм собственности и регионов // Фундаментальные исследования. - 2016 - №. 9-2. - С. 429-435
12. Трифонов А. Ю., Михальчук А. А., Спицын В. В., Новосельцева Д. А. Сравнительный анализ экономических и социальных результатов российских и иностранных предприятий по производству машин и оборудования // Фундаментальные исследования. - 2015 - №. 11-3. - С. 611-616
13. Трифонов А. Ю., Михальчук А. А., Спицын В. В., Новосельцева Д. А., Гуменников И. В. Анализ инвестиционной активности российских и иностранных предприятий по производству машин и оборудования // Фундаментальные исследования. - 2015 - №. 12-5. - С. 1059-1066
14. Арефьев В. П., Михальчук А. А., Филипенко Н. М., Новосельцева Д. А. Кластеризация направлений подготовки российского вуза в факторном пространстве вступительных испытаний // Открытое и дистанционное образование. - 2015 - №. 3. - С. 69-76
15. Спицын В. В., Михальчук А. А., Спицына Л. Ю., Новосельцева Д. А. Сравнительный анализ показателей инвестиционной деятельности российских и иностранных предприятий по производству транспортных средств // Современные проблемы науки и образования. - 2015 - №. 1. - С. 1
16. Новосельцева Д. А., Крицкий О. Л. Использование соотношения call-put для расчета стохастической процентной ставки и нахождения улыбки волатильности // Экономика и предпринимательство. - 2014 - №. 5-2. - С. 87-89
17. Новосельцева Д. А. Статистическая модель кластеризации российских вузов в факторном пространстве вступительных испытаний по направлению подготовки ядерные физика и технологии // Молодежь и современные информационные технологии: сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых: в 2 т., Томск, 7-11 Ноября 2016. - Томск: ТПУ, 2017 - Т. 1 - С. 123-125
18. Новосельцева Д. А., Михальчук А. А., Спицын В. В. Статистический анализ социальных показателей химической промышленности России в разрезе форм собственности // Молодежь и современные информационные технологии: сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых ученых. В 2 т., Томск, 7-11 Ноября 2016. - Томск: ТПУ, 2017 - Т. 2 - С. 290-292
19. Новосельцева Д. А., Михальчук А. А., Спицын В. В. Дисперсионный анализ динамики отгруженной продукции российских и иностранных предприятий по производству пищевых продуктов // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов III Международной конференции: в 2 т., Томск, 23-26 Мая 2016. - Томск: ТПУ, 2016 - Т. 2 - С. 299-301

20. Новосельцева Д. А. Кластеризация российских вузов в факторном пространстве вступительных испытаний по направлению подготовки Математика [Электронный ресурс] // Перспективы развития фундаментальных наук: сборник научных трудов XIII Международной конференции студентов, аспирантов и молодых ученых, Томск, 26-29 Апреля 2016. - Томск: ТПУ, 2016 - Т. 3. Математика - С. 81-83 - http://science-persp.tpu.ru/Arch/Proceedings_2016_vol_3.pdf
21. Новосельцева Д. А. Статистическая модель оценивания различий экономических показателей предприятий России по производству электроники и электрооборудования // Молодежь и современные информационные технологии: сборник трудов XIII Международной научно-практической конференции студентов, аспирантов и молодых ученых, Томск, 9-13 Ноября 2015. - Томск: ТПУ, 2016 - Т. 1 - С. 160-161
22. Новосельцева Д. А. , Михальчук А. А. , Спицын В. В. Кластеризация регионов по социальным показателям деятельности российских и иностранных предприятий по производству машин и оборудования // Молодежь и современные информационные технологии: сборник трудов XIII Международной научно-практической конференции студентов, аспирантов и молодых ученых , Томск, 9-13 Ноября 2015. - Томск: ТПУ, 2016 - Т. 2 - С. 238-239
23. Novoseltseva D. A. Computer anova of quality of results of higher mathematics E-learning [Electronicresources] // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов II Международной конференции, Томск, 19-22 Мая 2015. - Томск: ТПУ, 2015 - С. 581-582 - <http://www.lib.tpu.ru/fulltext/c/2015/C24/C24.pdf>
24. Новосельцева Д. А. Многомерный статистический анализ качества успеваемости студентов по высшей математике [Электронный ресурс] // Перспективы развития фундаментальных наук: сборник трудов XII Международной конференции студентов и молодых ученых, Томск, 21-24 Апреля 2015. - Томск: ТПУ, 2015 - С. 642-644 - http://science-persp.tpu.ru/Previous%20Materials/Konf_2015.pdf
25. Новосельцева Д. А., Михальчук А. А., Спицын В. В. Статистический анализ показателей инвестиционной деятельности российских и иностранных предприятий по производству электроники и электрооборудования [Электронный ресурс] // Информационные технологии в науке, управлении, социальной сфере и медицине: сборник научных трудов II Международной конференции, Томск, 19-22 Мая 2015. - Томск: ТПУ, 2015 - С. 351-354 - <http://www.lib.tpu.ru/fulltext/c/2015/C24/C24.pdf>
26. Новосельцева Д. А. Динамическая дисперсионная модель мониторинга образовательных показателей студентов во Втузе [Электронный ресурс] // Молодежь и современные информационные технологии: сборник трудов XII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых: в 2 т., Томск, 12-14 Ноября 2014. - Томск: ТПУ, 2014 - Т. 1 -

- С. 178-179. - Режим доступа: http://www.lib.tpu.ru/fulltext/c/2014/C04/V1/C04_V1.pdf
27. Новосельцева Д. А. Дисперсионная модель оценивания результатов входного контроля математических знаний студентов во Втузе [Электронный ресурс] // Молодежь и современные информационные технологии: сборник трудов XII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых: в 2 т., Томск, 12-14 Ноября 2014. - Томск: ТПУ, 2014 - Т. 1 - С. 176-177. - Режим доступа: http://www.lib.tpu.ru/fulltext/c/2014/C04/V1/C04_V1.pdf
28. Новосельцева Д. А., Михальчук А. А., Спицын В. В. Многомерный анализ социально-экономических показателей российских и иностранных предприятий подразделов DL и DM // Экономика России в XXI веке: сборник научных трудов XII Всероссийской научно-практической конференции "Экономические науки и прикладные исследования": в 2 т., Томск, 17-21 Ноября 2015. - Томск: Изд-во ТПУ, 2015 - Т. 2 - С. 404-410
29. Новосельцева Д. А., Шапарев В. Э., Саклаков В. М. Многомерный анализ показателей инвестиционной деятельности российских и иностранных предприятий по производству электроники и электрооборудования // Современные технологии поддержки принятия решений в экономике: сборник трудов Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых, Юрга, 19-20 Ноября 2015. - Томск: ТПУ, 2015 - С. 314-317

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. *Прикладная статистика: Исследование зависимостей* – М.: Финансы и статистика, 1985. – 487 с.
2. Андерсон Т. *Введение в многомерный статистический анализ*. – М.: Изд. иностр. литер., 1963.– 500 с.
3. Афифи А., Эйзен С. *Статистический анализ. Подход с ЭВМ*. – М.: Мир, 1982. – 488 с.
4. Большев Л.Н., Смирнов Н.В. *Таблицы математической статистики*. – М.: Наука, 1983. – 416 с.
5. Холлендер М., Вульф Д.А. *Непараметрические методы статистики*, пер. с англ. – М.: Мир, 1983. – 518 с.
6. Боровиков В.П. *Программа Statistica для студентов и инженеров*. – М.: КомпьютерПресс, 2001. – 301с.
7. Боровиков В.П., *Statistica. Искусство анализа данных на компьютере*. – СПб.: Питер, 2003. – 688с.
8. *Вероятность и математическая статистика: Энциклопедия/ Гл. ред. Ю.В. Прохоров*. – М.: Большая Российская энциклопедия, 1999. – 910с.
9. Вуколов Э. А. *Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: учебное пособие*. – М.: Форум, 2011. – 464 с.
10. Рао С.Р. *Линейные статистические методы и их применение*. – М.: Наука, 1968. – 548 с.
11. Дубров А.М. и Мхитарян В.С., Трошин Л.И. *Многомерные статистические методы: Учебник*. – М.: Финансы и статистика, 2003.– 352 с.
12. Дюран Б., Оделл П. *Кластерный анализ*. – М.: Статистика, 1977. – 128с.
13. Иберла К. *Факторный анализ: пер. с нем.* – М.: Статистика, 1980. – 398с.
14. Крамер Г. *Математические методы статистики*. – М.: Мир, 1975.– 648с.

15. Кремер Н.Ш. *Теория вероятностей и математическая статистика: Ученик для вузов.* – М: ЮНИТИ-ДАНА, 2007 – 573 с.
16. Мандель И. Д. *Кластерный анализ.* – М.: Финансы и статистика. 1988. – 176 с.
17. *Факторный, дискриминантный и кластерный анализ / сборник работ под ред. И. С. Енюкова.* – М.: Финансы и статистика, 1989. – 215 с.
18. Поллард Дж. *Справочник по вычислительным методам статистики.* – М.: Финансы и статистика, 1982. – 344 с.
19. Симчера В. М. *Методы многомерного анализа статистических данных: учебное пособие.* – М.: Финансы и статистика, 2008. – 389 с.
20. Шеффе Г. *Дисперсионный анализ.* – М.: Наука, 1980. – 512 с.
21. Материалы сайта Росстата — URL: <http://www.gks.ru/>
22. Buys, A., Ownership and innovative behavior. The case of South African automotive component manufacturing industry. *Proceeding Technology Management for Global Economic Growth (PICMET)*, 2010, 1-9 p.
23. Antonescu, D. , Empirical analysis of Foreign direct investments at NUTS 2 region, in European Union and Romania. *Procedia Economics and Finance*, 22, 2015, 681-689 p.
24. Bellak, C., How domestic and Foreign firms differ and why does it matter? *Journal of Economic Surveys*, 18(4), 2004, 483-514 p.
25. Barnes, J., Kaplinsky, R., Morris, M., Industrial policy in developing economies: Developing dynamic comparative advantage in the South African automobile sector. *Competition & Change*, 8(2), 2004,153-172 p.
26. Javorcik, B., Does Foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American Economic Review*, 94(3), 2004, 605-627 p.
27. Myachin, D., Royzen, A., Pershikov, A., Regional features of attracting Foreign investments into the Russian economy. *Procedia - Social and Behavioral Sciences*, 166, 2015,131-134 p.


28. Wang, J., Wang, X., Benefits of Foreign ownership: Evidence from Foreign direct investment in China. *Journal of International Economics*, 97(2), 2015, 325-338 p. DOI: 10.1016/j.jinteco.2015.07.006.
29. Temouri, Y., Driffield, N., Añón Higón, D., Analysis of productivity differences among Foreign and domestic firms: Evidence from Germany. *Review of World Economics*, 144(1), 2008, 32-54 p.
30. Chen, Z., Ge, Y., Lai, H., Foreign direct investment and wage inequality: Evidence from China. *World Development*, 39(8), 2011, 1322-1332.
31. Крейнина М.Н. Финансовый менеджмент: учеб. пособие. - М.: "Дело и Сервис", 2001. - 400 с.
32. Черняк, В.З. Бизнес планирование: Учебник для вузов.- М. ЮНИТИ-ДАНА, 2010. – 470 с.
33. Головань С.И Бизнес-планирование: Учебное пособие. – М.: Феникс, 2009. – 320с.
34. Основы функционально-стоимостного анализа: Учебное пособие / Под ред. М.Г. Карпунина и Б.И. Майданчика. - М.: Энергия, 1980. - 175 с.
35. Долин П.А. Справочник по технике безопасности. – М.: Энергоатомиздат, 1982. – 800с.
36. Белов С.В. Безопасность жизнедеятельности и защита окружающей среды: учебник для вузов. – М.: Изд- во Юрайт, 2013. – 671

ПРИЛОЖЕНИЕ А


Раздел 1

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА

Студент:

Группа	ФИО	Подпись	Дата
0BM51	Новосельцева Дарья Андреевна		10.03.17

Консультант – лингвист кафедры ИЯ:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель ТПУ	Зяблова Наталья Николаевна	Кандидат филологических наук		10.03.17

1. THEORETICAL BASIS OF MULTIVARIATE STATISTICAL ANALYSIS

1.1. Analysis of variations

Analysis of variations (ANOVA) is statistical method, that allows to analyze the influence of various factors (categorical, grouping, independent variables) on the results of the experiment (dependent variables) [1, 3]. There is the need for ANOVA: independent variable must be categorical, dependent variable must be metric. The essence of ANOVA is the decomposition of the variance of the measured characteristic into independent summands that characterizes the influence of a factor or their interaction. The analysis is based on the calculation of F-test (Fisher test), which is the ratio of two variances: intergroup and intragroup. The F-test in a one-way ANOVA determines significance of the differences of means of several independent samples. It replaces the T-test for independent samples in the presence of more than two samples and gives the same result in the case of two samples. The mathematical model (or mathematical relation) for ANOVA. It is built and represents each dependent variable as a sum of mean and error. In its turn, means of each variable are represented like a sum of the general mean and the "effect" of each factor. The statistical problems that arise in this action are related to the evaluation of these effects and the verification of statistical hypotheses about them. The problem of testing the hypotheses can be solved only by introducing additional assumptions about the probabilistic structure of the errors in the observations. It is usually assumed that they are independent and obey a normal law with zero mean and constant variance. It allows us to use the developed theory of the method of least squares (MLS). Less stringent assumptions require a sufficiently large number of observations, in which it becomes justified to resort to results of an asymptotic theory.

The classification of ANOVA models primarily is based on the nature of the analyzing factors. There are models with fixed factors (model I) and models with random factors (model II).

The technique of ANOVA varies depending on the number of studied independent factors.

Parametric one-way ANOVA.

Parametric one-way ANOVA is used in cases, that have three or more independent samples (groups) obtained from one general complex by changing any factor A having, respectively, three or more levels $A_i, i = 1, \dots, k$. It is assumed that these samples are distributed according to the normal law and have different sample mean and identical sample variances σ^2 . It is necessary to answer the question whether this factor had a significant influence on the spread of the sample mean group or the spread is a consequence of randomness caused, for example, by a small sample size.

Let Y be a random variable defined on this general complex, and μ is its mean.

In the representation $\mu_i = \mu + \alpha_i$ for the average value of Y on the i-th sample of the volume n , the value α_i is the differential effect corresponding to the level A_i . y_{ij} is denoted by the value of Y determined by the jth observation ($j = 1, \dots, n_i$) of the i-th sample. This ANOVA model describes the situation under the assumption, that the distribution of Y at each sample is normal with the same variance σ^2 , i.e. with the law $N(\mu_i, \sigma^2)$.

Each value of y_{ij} is equal to the sum of the general mean μ (uniform for all k factor levels), the differential effect α_i , which is determined by the level A_i , and the random error ε_{ij} caused by the influence of uncontrolled factors:

$$y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij}; \quad (1.1)$$

$$j = 1, \dots, n_i; \quad i = 1, \dots, k,$$

$\{\varepsilon_{ij}\}$ are independent and distributed according to the normal law $N(0, \sigma^2)$.

The factor will be interpreted according to model I, that is, the experimenter is interested only in samples that correspond precisely to these levels of the factor. The model of one-way ANOVA with fixed effects (model I) is given by the relation (1.1).

The hypothesis H_0 : all $\alpha_i = 0$ is checked.

From the representation of the model in the form (1.1) one can derive the MLS $\hat{\mu}_i$ estimates for μ_i , $i = 1, \dots, k$. MLS of μ и α_i follow from them. To ensure the uniqueness of MLS estimates, we will have to impose an additional restriction on the parameters $\alpha_1, \dots, \alpha_k$.

The usual requirement is that the weighted sum of effects is zero:

$$\sum_{i=1}^k n_i \alpha_i = 0.$$

Thus, the problem reduces to minimizing the sum of squares

$$S = \sum_{i=1}^k \sum_{j=1}^{n_i} (\varepsilon_{ij})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [y_{ij} - \mu_i]^2$$

on variables μ_i , $i = 1, \dots, p$.

The main idea of ANOVA lies in the expansion (1.2). The total variation of the variable, that generated by the influence of the factor and measured by the sum of SS_T , is made up of two components: SS_B и SS_R . These components characterize the variability of this variable between the factor levels (SS_B) and the variability within the levels (SS_R).

$$SS_T = SS_B + SS_R, \quad (1.2)$$

where

$$SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} [y_{ij} - \bar{y}_{**}]^2$$

Is the total sum of squares of deviations;

$$SS_B = \sum_{i=1}^k \sum_{j=1}^{n_i} [\bar{y}_{i*} - \bar{y}_{**}]^2$$

– the sum of squares of deviations of group averages from the total average, or the intergroup (inter-level factorial) sum of squares of deviations;

$$SS_R = \sum_{i=1}^k \sum_{j=1}^{n_i} [y_{ij} - \bar{y}_{i*}]^2$$

– the sum of squares of deviations of observations from group averages, or the intragroup (residual) sum of squares of deviations.

In ANOVA, not the sums of the squares of deviations are analyzed, but the so-called mean squares, which are obtained by dividing the sums of the squares of deviations by the corresponding number of degrees of freedom (MS_B , MS_R). For the mean square MS_B , which is an unbiased estimate of the intergroup dispersion, the number of degrees of freedom $\nu_B = k - 1$, since in its calculation k group mean values are used that are related by one equation. For the mean square, which is an unbiased estimate of the intergroup dispersion, the number of degrees of freedom, since in its calculation of k average mean values are used that are related by a single equation. For the mean square MS_R , which is an unbiased estimate of the intragroup dispersion, the number of degrees of freedom $\nu_R = n - k$, because when calculating it, all n observations connected by k equations are used. Thus, we have:

$$MS_B = SS_B / (k - 1); MS_R = SS_R / (n - k).$$

To test the hypothesis $H_0: \alpha_1 = \dots = \alpha_k = 0$ that all differential effects are zero, calculate the observed F-ratio:

$$F_H = MS_B / MS_R,$$

In order to draw a final conclusion, it is necessary to find the critical value F_{kp} from the Fisher distribution table taking into account the number of degrees of freedom (ν_B и ν_R) and the corresponding significance level. If the observed value of the criterion F_H , calculated from the sample data, belongs to the critical region ($F_H > F_{kp}$), then the hypothesis is rejected; If the observed value does not belong to the critical region, then there is no reason to reject the hypothesis. Adoption (not rejection) of the hypothesis $H_0: \alpha_i = 0$ means the validity of the equivalent hypothesis, i.e. $H_0: \mu_1 = \dots = \mu_k = \mu$, that all k group averages are equal to the general average, which practically means an insignificant difference between the group and the general average. The result is significant if the hypothesis is rejected, since this indicates that there are significant differences between the mean values for the groups (factor levels).

In the case of several independent groups, the Kruskal-Wallis rank test of ANOVA is a nonparametric analogy (alternative) of a one-way ANOVA. Kruskal-Wallis-test is based on ranks, and not on initial observations and assumes that the variable under consideration is continuous and measured, at least, in the ordinal scale. The K-W tests the hypothesis: the compared samples have the same distribution.

Consider the theoretical basis of the Kruskal-Wallis test on the example of the following problem [5].

The data consist of $N = \sum_{j=1}^k n_j$ observations, from n_j observations to the j -th sample (group), $j = 1, \dots, k$.

The following initial mathematical model of nonparametric one-factor analysis of variance is proposed:

$$x_{ij} = \mu + \tau_j + \varepsilon_{ij}, \quad i=1, \dots, n_j, \quad j=1, \dots, k,$$

μ – is the unknown total mean; τ_j – unknown effect of j -th sample ($\sum_{j=1}^k \tau_j = 0$).

All ε_{ij} (random errors) are mutually independent and extracted from the same continuous collection.

According to the available assumptions, it is required to test the hypothesis: $H_0: \tau_1 = \tau_2 = \dots = \tau_k$. For doing this, we must perform the following actions:

- 1) To rank all N observations together from the smallest to the largest. Let r_{ij} denote the rank of x_{ij} in this joint ranking;
- 2) For $j = 1, \dots, k$

$$R_j = \sum_{i=1}^{n_j} r_{ij}, \quad R_{*j} = \frac{R_j}{n_j}, \quad R_{**} = \frac{N+1}{2}$$

For example, R_1 is the sum of the ranks assigned to sample 1, and R_{*1} is the average rank that this sample received.

- 3) To calculate statistics

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k n_j (R_{*j} - R_{**})^2 = \left[\frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} \right] - 3(N+1)$$

4) At the significance level α , reject H_0 , if $H \geq h(\alpha, k, (n_1, \dots, n_k))$, take H_0 if $H < h(\alpha, k, (n_1, \dots, n_k))$, where the constant $h(\alpha, k, (n_1, \dots, n_k))$ satisfies the condition $P\{H \geq h(\alpha, k, (n_1, \dots, n_k))\} = \alpha$. If H_0 is true, then the statistic H has an asymptotic distribution (with $\min(n_1, \dots, n_k) \rightarrow \infty$) of χ^2 with $k-1$ degrees of freedom. The approximate criterion for the level α is as follows: reject H_0 if $H \geq \chi^2(k-1, \alpha)$, take H_0 if $H < \chi^2(k-1, \alpha)$, where $\chi^2(k-1, \alpha)$ is the upper $\alpha\%$ χ^2 -distribution with $k-1$ degrees of freedom.

One-factor multidimensional ANOVA.

Let n_r ($n = \sum_{r=1}^k n_r$) be the number of observations obtained for the r -th sample of k samples (groups) for each of the m variables Y_1, Y_2, \dots, Y_m . Denote by $y_{ij}^{(r)}$ ($i = \overline{1, m}, j = \overline{1, n_r}, r = \overline{1, k}$) the value of the variable Y_i , determined by the j -th observation of the r -th sample. Then

$$\bar{Y}_{ir} = \frac{1}{n_r} \sum_{j=1}^{n_r} y_{ij}^{(r)} \quad (i = \overline{1, m}, r = \overline{1, k})$$

- selective (group) mean value of the variable Y_i for the r -th sample, and

$$\bar{Y}_i = \frac{1}{n} \sum_{r=1}^k \sum_{j=1}^{n_r} y_{ij}^{(r)} \quad (i = \overline{1, m})$$

- the total average value for the combined sample for each Y_i of m variables. In addition,

$$S_{ij}^{(r)} = \sum_{l=1}^{n_r} \sum_{q=1}^{n_r} [y_{il}^{(r)} - \bar{Y}_{ir}] [y_{jq}^{(r)} - \bar{Y}_{jr}] \quad (i, j = \overline{1, m})$$

- the value of the residual sum of squares and products for the r -th sample with n_r-1 degrees of freedom.

Similarly, SS_B in a one-dimensional ANOVA is defined in the multidimensional case

$$B_{ij} = \sum_{r=1}^k n_r \bar{Y}_{ir} \bar{Y}_{jr} - n \bar{Y}_i \bar{Y}_j \quad (i, j = \overline{1, m})$$

As the inter-group sum of squares of deviations, and similarly to SS_R

$$W_{ij} = \sum_{r=1}^k S_{ij}^{(r)} \quad (i, j = \overline{1, m})$$

Intragroup (residual) sum of squares and products of deviations.

The matrices (B_{ij}) and (W_{ij}) form the corresponding (intergroup and intragroup) sources of the variance of the system m of random variables $\{Y_1, Y_2, \dots, Y_m\}$ with respectively $k-1$ and $n-k$ numbers of degrees of freedom. These values form the basis of the multivariate ANOVA table. Then the multidimensional Wilkes Λ -test (Wilks - lambda test) for testing the hypothesis of equality of the mean values for k samples with respect to the set of m variables has the form

$$\Lambda = |W|/|W + B|,$$

Where $|W|$ and $|B + W|$ are determinants of the matrices (W_{ij}) and $(B_{ij}) + (W_{ij})$, respectively.

The statistics Λ has a U-distribution with m , $k-1$ and $n-k$ degrees of freedom [2]. With the exception of special cases, percentiles of U-distributions can be difficult to compute, and therefore in practice one of the two approximations is usually used. Thus, the question of whether the hypothesis to be tested should be rejected can be solved by comparing the Bartlett test

$$\chi^2 = -[n - 1 - (m+k)/2] \ln \Lambda$$

With percentiles of the χ^2 -distribution with $m(k-1)$ degrees of freedom. On the other hand, we can use the F -approximation U proposed by Rao [10]:

$$F = \frac{[n - 1 - (m + k) / 2] \cdot s - m(k - 1) / 2 + 1}{m(k - 1)} \cdot \frac{1 - \Lambda^{1/s}}{\Lambda^{1/s}}, \quad s = \sqrt{\frac{m^2(k - 1)^2 - 4}{m^2 + (k - 1)^2 - 5}}$$

The hypothesis is rejected if $F > F_{1-\alpha}$ with degrees of freedom determined by the numerator and denominator of the first fraction in the expression F .

Thus, if there are several dependent variables at the same time, only the complexity of the calculations increases and the content and basic principles do not

change. Instead of a one-dimensional F -test, a multidimensional F -test (Wilks-lambda test) is used, based on a comparison of the covariance error matrix and the inter-group covariance matrix.

If the general multidimensional test is significant, we can conclude that the corresponding effect is significant in the totality of all variables. In other words, if the hypothesis of the equality of the averages over the set of all variables is rejected on the basis of the Λ -test, then the hypothesis on the equality of the means for each variable on the basis of the one-dimensional F -test is checked. If it is rejected, then we conclude that some μ_i do not coincide. Since the F -test does not provide information about which of the averages are not equal, it is necessary to conduct a **multiple ANOVA**.

Multiple ANOVA.

There are three procedures for multiple comparison:

1) the Scheffe method - to test hypothesis $H_0: \sum_{r=1}^k c_r \mu_r = 0$ against alternative $H_1: \sum_{r=1}^k c_r \mu_r \neq 0$ with the significance level α , the following $100(1 - \alpha)\%$ confidence interval is formed:

$$\sum_{r=1}^k c_r \bar{y}_{r*} \pm \sqrt{k \cdot MS_R \cdot F_{1-\alpha}(k, n-k) \sum_{r=1}^k \frac{c_r^2}{n_r}},$$

MS_R is the mean square is an unbiased estimate of the intragroup dispersion with the number of degrees of freedom $\nu_R = n - k$ and is taken from the table of one-way variance analysis, and $F_{1-\alpha}(k, n - k)$ is $100(1 - \alpha)\%$ -th percentile Distribution $F(k, n - k)$. If this interval does not contain zero, then H_0 is rejected with the level α . This process is repeated for each linear combination of interest to us, and the level of significance common to all criteria remains α .

2) the Tukey method, which is applicable only to contrasts and only in the case of equal sample volumes, i.e. for $n_1 = n_2 = \dots = n_k = n/k$. To test hypothesis H_0 :

$\sum_{r=1}^k \lambda_r \mu_r = 0$ against alternative $H_1: \sum_{r=1}^k \lambda_r \mu_r \neq 0$ form the following confidence interval:

$$\sum_{r=1}^k \lambda_r \bar{y}_{r*} \pm \frac{1}{2} \sqrt{\frac{k \cdot MS_R}{n}} \cdot q_{1-\alpha} \sum_{r=1}^k |\lambda_r|,$$

And $q_{1-\alpha}$ is 100 (1 - α)% -th percentile of the distribution of the standardized range with k and $\nu = n - k$ degrees of freedom. If this interval does not contain zero, then H_0 is rejected with a significance level α .

3) t-method. Let k_0 be the number of preselected contrasts. Then to test hypothesis $H_0: \sum_{r=1}^k \lambda_r \mu_r = 0$ against alternative $H_1: \sum_{r=1}^k \lambda_r \mu_r \neq 0$ it is necessary to construct a confidence interval:

$$\sum_{r=1}^k \lambda_r \bar{y}_{r*} \pm t_{1-\alpha/2k_0} (v_R) \sqrt{MS_R \sum_{r=1}^k \frac{\lambda_r^2}{n_r}},$$

And $t_{1-\alpha/2k_0}$ is 100 (1 - $\alpha / 2k_0$)% -th percentile of t-distribution v_R by degrees of freedom. If this interval does not contain zero, then we must reject H_0 .

1.2. Factor analysis

If an object is described using n traits (trace elements), then as a result of the action of the method, a factor model is obtained, depending on a smaller number of variables [13, 14, 18]. It is assumed that the initial measured data $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$ is affected by a small number of latent (latent) signs (factors) $\{U_1, \dots, U_m\}$, $m < n$.

The purpose of factor analysis is to identify these hidden characteristics (factors) and to estimate their number.

Factor analysis is used either as a method of data reduction (reduction) or as a method of classification:

Factor analysis as a method of data reduction is based on the use of dependence between variables: a new variable is introduced based on regression that is, including the most significant features of the original variables, so that its use

(replacing several old correlated variables with one new one) leads to a reduction in the number of variables. The new factor (variable) is represented by a linear combination of the initial variables. After the first factor is selected, that is, the construction of the first regression line for which the variance is maximal, the next line maximizing the residual variation (the spread of data around the first straight line) is determined, that is, the second factor is allocated, etc. Factors are allocated one after another to maximize the variability remaining from the previous ones, that is, they are independent of each other. In other words, uncorrelated or orthogonal.

The decision about when to stop the procedure for the allocation of factors mainly depends on the point of view of what is considered a small "random" variability. This decision is quite arbitrary, but there are some general recommendations that allow rational choice of the number of factors: the criterion of accumulated or cumulative variance, the criterion of a rocky scree. Therefore, usually several solutions with a greater or lesser number of factors is investigated, and then one of the most "meaningful" ones is chosen.

Factor analysis as a classification method is based on estimations of correlations (factor loads) between initial variables and factors within the chosen factor model and allows to learn the significance of factors, that is, whether it is possible to interpret them in a reasonable way and how to do it. You can rotate the axes in any direction without changing the relative position of the points; however, the actual coordinates of the points, that is, the factor loads, must undoubtedly change. The purpose of rotation is to obtain an understandable (interpreted) load matrix, that is, factors that are clearly marked by high loads for some variables and low for others, which allows classification of variables.

Factor analysis, as a section of multidimensional statistical analysis, combines mathematical and statistical methods for solving problems associated with constructing a linear model:

$$\mathbf{X} = \mathbf{AU} + \boldsymbol{\varepsilon}, \quad (1.3)$$

$\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}^T$ – is - the initial observable n-dimensional random variable with zero mathematical expectations $M[\mathbf{X}] = 0$ and the correlation matrix $C = M[\mathbf{X}\mathbf{X}^T]$;

$A=(a_{ij})$ is the unknown $(n \times m)$ -matrix of loads of common factors on the observed quantities;

$U=\{U_1, \dots, U_m\}^T$ is an unobservable multidimensional random variable of common factors with zero mathematical expectations $M[U] = 0$ and a unit correlation matrix $M[UU^T]=E$;

$\varepsilon =\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}^T$ is the n -dimensional random error value $M[\varepsilon] = 0$, $M[\varepsilon U^T]=0$ and the unknown diagonal covariance $(D[\varepsilon_i] = \sigma_i^2)$ matrix $\Psi, = M[\varepsilon\varepsilon^T]$.

From the model of factor analysis it follows that

$$C=M[XX^T] =M[(AU +\varepsilon)(AU +\varepsilon)^T] =AM[UU^T] A^T+M[\varepsilon\varepsilon^T]=AA^T+ \Psi,$$

That is, the correlations c_{ij} between the initial random variables are related to the factor loads a_{ij} and the variances of the errors σ_i^2 with the help of the relations:

$$c_{ij} = \sum_{l=1}^m a_{il}a_{jl} \quad , \quad 1 = c_{ii} = \sum_{l=1}^m a_{il}^2 + \sigma_i^2 = h_i^2 + \sigma_i^2, \quad (1.4)$$

$h_i^2 = \sum_{l=1}^m a_{il}^2$ is called the commonness of the initial random variable X_i and is

determined by the sum of the squares of the factor loads, that is, the contribution of all m common factors to the variance of the quantity X_i .

Consideration was continued in the framework of the method of principal factors, by which we mean the application of the method of principal components to the reduced correlation matrix C_h in the case of a linear model:

$$X = AU \quad \text{или} \quad X_i = \sum_{j=1}^m a_{ij}U_j \quad (i = \overline{1, n}); \quad m < n, \quad (1.5)$$

$X=\{X_1, X_2, \dots, X_n\}^T$ с $M[X] = 0$ and the correlation matrix C_h .

the matrix of loads of common factors

$$A=(a_{ij})= \{\alpha_1, \alpha_2, \dots, \alpha_k, \dots, \alpha_m\},$$

$\alpha_k = \{a_{1k}, a_{2k}, \dots, a_{nk}\}^T$ – k -th column of the factor load matrix A .

On the other hand,

$$A=(a_{ij})= \{\beta_1, \beta_2, \dots, \beta_k, \dots, \beta_n\}^T,$$

$\beta_k = \{\mathbf{a}_{k1}, \mathbf{a}_{k2}, \dots, \mathbf{a}_{km}\}$ – k -th row of the matrix A . Then the linear model (1.5) can be represented as:

$$X_k = \beta_k U \quad (k = \overline{1, n}) \quad (1.6)$$

According to the method of principal components, the general factor U_k should contribute the maximum to the total community:

$$\sum_{l=1}^n a_{lk}^2 = \alpha_k^T \alpha_k \rightarrow \max$$

Under the condition $AA^T = C_h$, expressed in terms of β_k :

$$c_{ij} = \sum_{l=1}^m a_{il} a_{jl} = \beta_i \beta_j^T, \quad (1.7)$$

$$c_{ij} = c_{ji} \text{ and } c_{ii} = h_i^2$$

This problem of the conditional extremum is solved using the Lagrange uncertain multipliers method:

$$L = \alpha_k^T \alpha_k - \sum_{i,j=1}^n \mu_{ij} (\beta_i \beta_j^T - c_{ij}),$$

$\mu_{ij} = \mu_{ji}$ – are Lagrange multipliers.

Finding the partial derivatives of the Lagrange function with respect to the components of the vector α_k^T and equating them to zero, we obtain a system of n linear homogeneous algebraic equations, which, taking into account the notation

$$\sum_{l=1}^n a_{lk}^2 = \alpha_k^T \alpha_k = \lambda_k$$

after a number of algebraic transformations it is possible to

represent independently of k in the matrix form:

$$(C_h - \lambda_k E) \alpha_k = 0.$$

The last equality defines the eigenvalue problem (numbers) λ_k and the eigenvectors α_k of the reduced correlation matrix C_h .

The nontriviality condition for the solution of the eigenvector problem (the degeneracy of the corresponding matrix)

$$|C_h - \lambda_k E| = 0$$

leads to an algebraic equation of order n with respect to the eigenvalue. Thus, we obtain n ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq \dots \geq \lambda_n$.

In this case, each eigenvalue (number) $\lambda_k (k = \overline{1, n})$ corresponds to the eigenvector $\alpha_k - k$ -th column of the factor load matrix A , corresponding to the factor U_k , that is, the column of coefficients under the factor U_k in (1.5).

Taking into account the ordering of the eigenvalues and the normalization of the corresponding eigenvectors ($\alpha_k^T \alpha_k = \lambda_k$), it is possible to rank the factors by their contribution to the total community.

Based on the initial observed value X and the calculated load matrix of common factors A , it is possible to obtain estimates of the common factors U . For example, in the framework of the main component method ($m = n$), the matrix A will be quadratic and the solution of the matrix equation (1.5) with respect to U does not cause difficulties if rank A is n , that is, there is an A^{-1} - inverse to A matrix:

$$U = A^{-1}X.$$

In the framework of the method of principal factors ($m < n$), the rectangular matrix A does not have an inverse and the procedure for estimating U is more complicated:

$$U = (A^T A)^{-1} A^T X,$$

the matrix $A^T A$ is already square of order m .

According to the estimates of U , it is possible to judge each object of observation with respect to m common factors.

To test the significance of the constructed model of factor analysis, one can use the Bartlett criterion to test the null hypothesis H_0 that m common factors are sufficient to explain selective correlation coefficients. The observed statistical significance of the criterion is calculated from the formula

$$\chi_H^2 = \left(N - \frac{2n+5}{6} - \frac{2}{3}m \right) \ln \frac{|AA^T|}{|C|},$$

N – Number of observations (objects of research),

$|AA^T|$ – The determinant of the matrix of correlations reproduced from the model,

$|C|$ – Determinant of the original correlation matrix

For a given level of significance, α H_0 is rejected, $\chi_H^2 > \chi_{kp}^2$,, where χ_{kp}^2 it is determined using the probability distribution calculator according to the law χ^2 with $\nu = \frac{(n-m)^2 - n - m}{2}$ degrees of freedom so that $P(\chi_H^2 > \chi_{kp}^2) = \alpha$. This means that m of common factors is not enough and we should allocate at least $m + 1$ factor. After that, the verification procedure is repeated.

Further, based on the results of factor analysis, we can, for example, classify observational objects by common factors, the number of which is much smaller than the number of initial indicators.

Based on the initial observed value X and the calculated matrix of loads of common factors A , it is possible to obtain estimates of the common factors U .

According to the estimates of U , it is possible to judge each object of observation with respect to m common factors.

1.3. Cluster analysis

Cluster analysis is a statistical analysis that allows one to obtain a partition of a large volume of data into classes or groups by some criterion or a combination of them [12, 17].

To carry out the classification of data X_1, \dots, X_n use the notion of metric or distance. In the theory of cluster analysis, different metrics are used to measure the distance between individual points (vectors), here are some of them:

1) Euclidean distance

$$\rho^2(X^{(i)}, X^{(j)}) = \sum_{k=1}^m (X_k^{(i)} - X_k^{(j)})^2 ;$$

w_k – is weight, proportional to the importance of the characteristic in the classification problem. Weights are asked after additional studies and believe that

$$\sum_{k=1}^m w_k = 1;$$

2) City-block (Manhattan) distance – is distance between cities in the town

$$\rho(X^{(i)}, X^{(j)}) = \sum_{k=1}^m |X_k^{(i)} - X_k^{(j)}|;$$

3) Chebychev distance metric

$$\rho(X^{(i)}, X^{(j)}) = \max \left\{ |X_k^{(i)} - X_k^{(j)}| \right\}$$

4) Correlation distance (1 - r Pearson). The correlation distance is calculated from the formula (1 - r), where r Pearson correlation coefficient, it is also called linear correlation, since it measures the degree of linear relationships between the variables. This type of distance differs from all other types of distances, because it uses correlation analysis techniques to estimate distances between clusters.

In a number of algorithms, along with the distances between vectors, distances between clusters and clusters are used. The methods of cluster analysis are divided into agglomerate (unifying), divisible (separating) and iterative.

The first consistently combine individual objects into clusters, the latter, on the contrary, dismember clusters into objects. Still others unite the first two. Their peculiarity is the formation of clusters, based on the partitioning conditions (the so-called parameters), which can be changed while the algorithm is working to improve the quality of the partition. Iterative methods are usually used to classify large amounts of information.

Agglomerate methods are the simplest and most common among the algorithms of cluster analysis. In the first step, each vector or object X_1, \dots, X_n of the source data is treated as a separate cluster or class. Based on the calculated distance matrix, the closest to each other are selected and combined. Obviously, the process will end through $(n - 1)$ step, when as a result all objects will be merged into one cluster.

The sequence of associations can be represented in the form of a dendrogram, or tree (Fig.1.1).

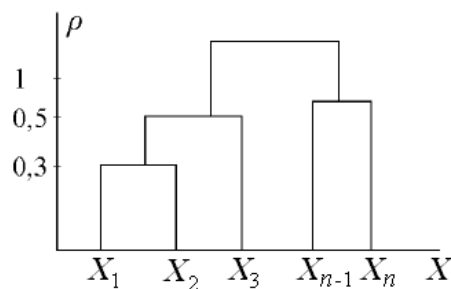


Figure 1.1 – The dendrogram

Methods of Ward, Single Linkage and Complete Linkage are agglomerate methods.

5) **Single Linkage.** Let ξ_1, \dots, ξ_n be data, with each vector forming one cluster. Then the matrix of distances between these clusters is calculated. Using this matrix, two closest vectors are selected, which form the first cluster S1. In the next step, the following distance matrix is calculated between S1 and the remaining vectors (which we consider to be clusters). The cluster closest to the previously obtained class S1 combines with it, forming S2, etc.

2) **Complete Linkage.** The distance is calculated only on the basis of the principle of a "distant neighbor" (the same can be said about the distance between classes united in classes - only the principle of a distant neighbor at $\alpha = \beta = \delta = 1/2, \gamma = 0$).

3) **Average linkage.** As with the full link method, all distances between clusters are compared to a user-defined number. And if it is less than a given number, the new object is included in the old class. Thus, the method of mean communication differs from the method of complete communication only by the method of calculating the distance between clusters.

4) **Ward's method.** This method differs from all other methods, because it uses the methods of dispersion analysis to estimate the distances between clusters. The method minimizes the sum of squares (SS) for any two (hypothetical) clusters

that can be generated at each step. In general, the method seeks to create small clusters.

The iterative methods of clustering include the k-means method.

1) **Method of k-means.** Suppose that there are vectors $\xi_1, \dots, \xi_n \in \mathbb{R}^p$ and they must be divided into k clusters. At the zero step of n vectors, randomly select k of them, assuming that each forms one cluster. We obtain the set of standard clusters $e_1^{(0)}, \dots, e_k^{(0)}$ with weights $w_1^{(0)}, \dots, w_k^{(0)}$, which determine the number of elements in them. Zero above denotes the iteration number. Zero above denotes the iteration number. At this stage, all weights are equal to one. In the next step, we select a vector from the remaining data set, for example ξ_i , and calculate the distance matrix between ξ_i and the standards $e_1^{(0)}, \dots, e_k^{(0)}$ in some metric, for example Euclidean:

$$\rho^2(\xi^{(i)}, e_m^{(0)}) = \sum_{j=1}^p (\xi_j^{(i)} - \xi_{m,j}^{(0)})^2, m = \overline{1, k}$$

Based on the knowledge of the calculated distance matrix, the vector ξ_i is placed in that standard, the distance to which is minimal. Let us assume for definiteness that this is $e_m^{(0)}$. It is replaced by a new one, recalculated taking into account the attached point, according to the formula

$$e_m^{(1)} = \begin{cases} \frac{w_m^{(0)} e_m^{(0)} + \xi_i}{w_m^{(0)} + 1}, & \xi_i \text{ ВКЛЮЧЕН} \\ e_m^{(0)}, & \xi_i \text{ НЕ ВКЛЮЧЕН} \end{cases}$$

In addition, the weight is also recalculated:

$$w_m^{(1)} = \begin{cases} w_m^{(0)} + 1, & \xi_i \text{ ВКЛЮЧЕН} \\ w_m^{(0)}, & \xi_i \text{ НЕ ВКЛЮЧЕН} \end{cases}$$

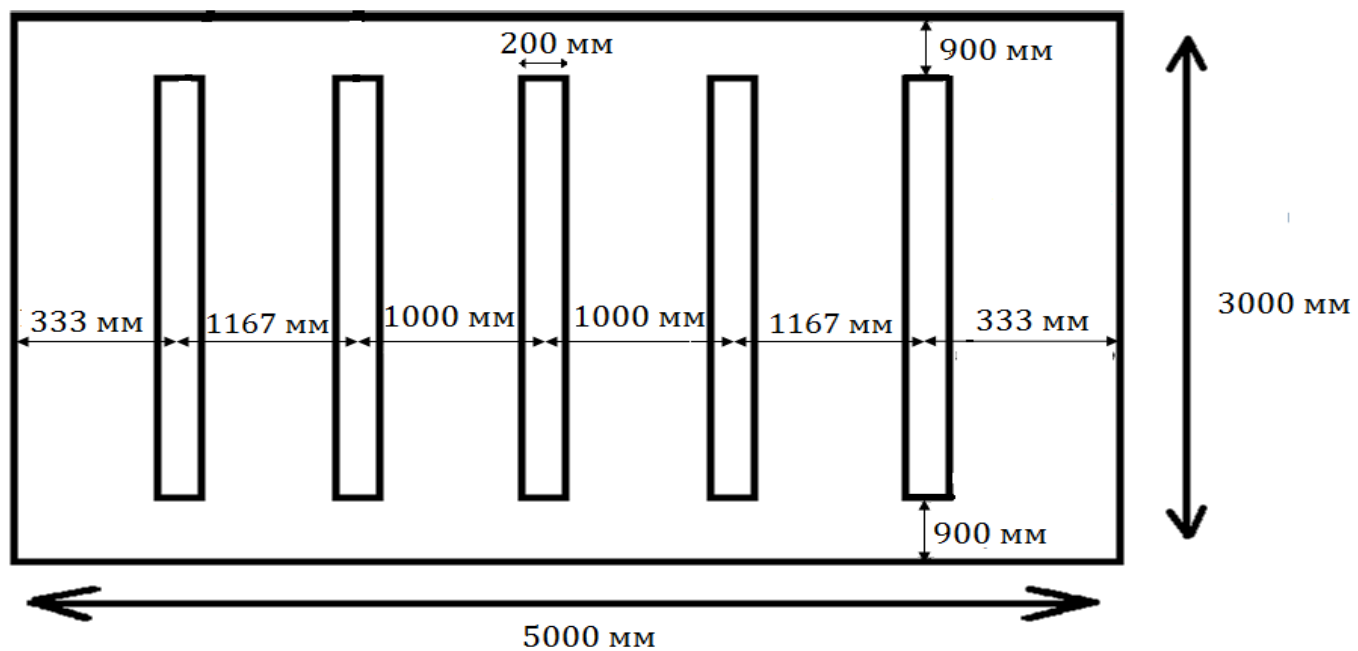
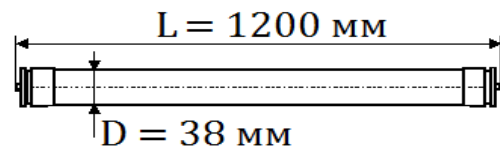
If two or more minimum distances occur in the matrix, then ξ_i is included in the cluster with the smallest sequence number. In the next step, the next vector from the remaining ones is selected, and the procedure is repeated. Thus, through $(n - k)$ steps, each weight $e_m^{(n-k)}$ corresponds to the weight $w_m^{(n-k)}$ and the clustering procedure is completed.

Note that the clustering problem does not have a unique solution. As a consequence, it is quite difficult to go through all the possible partitions of data into classes and it is not always possible. In order to evaluate the quality of various methods of clustering, the concept of a partitioning quality functional is introduced, which assumes a minimum value at the best (from the point of view of the researcher) partition.

ПРИЛОЖЕНИЕ Б

План помещения и размещения светильников с люминесцентными лампами типа ЛБ-40

Схема люминесцентной лампы типа ЛБ-40



ПРИЛОЖЕНИЕ В

План эвакуации людей.

