

**МИНОБРНАУКИ РОССИИ**

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

**«Российский государственный гуманитарный университет»**

**(РГГУ)**

**ИНСТИТУТ ИНФОРМАЦИОННЫХ НАУК И ТЕХНОЛОГИЙ БЕЗОПАСНОСТИ**

Факультет информационных систем и безопасности

Кафедра фундаментальной и прикладной математики

**СТАТИСТИЧЕСКИЕ ПАКЕТЫ ПРИКЛАДНЫХ ПРОГРАММ**

Рабочая программа дисциплины

Направление подготовки 01.03.04 Прикладная математика

Направленность (профиль) Прикладная математика

Уровень квалификации выпускника - бакалавр

Форма обучения - очная

РПД адаптирована для лиц  
с ограниченными возможностями  
здоровья и инвалидов

Москва 2017

СТАТИСТИЧЕСКИЕ ПАКЕТЫ ПРИКЛАДНЫХ ПРОГРАММ  
Рабочая программа дисциплины

Составитель(и):

кандидат физ.-мат. наук, доц., доцент кафедры фундаментальной и прикладной математики *Синицын В.Ю.*,  
доктор физ.-мат. наук, проф. *Пресман Э.Л.*

Ответственный редактор

доктор пед. наук, проф., заведующий кафедрой фундаментальной и прикладной математики *Жаров В.К.*

УТВЕРЖДЕНО

Протокол заседания кафедры  
фундаментальной и прикладной математики  
№ 14 от 20.06.2017

## **ОГЛАВЛЕНИЕ**

### **1. Пояснительная записка**

#### 1.1 Цель и задачи дисциплины

#### 1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине

#### 1.3. Место дисциплины в структуре образовательной программы

### **2. Структура дисциплины**

### **3. Содержание дисциплины**

### **4. Образовательные технологии**

### **5. Оценка планируемых результатов обучения**

#### 5.1. Система оценивания

#### 5.2. Критерии выставления оценки по дисциплине

#### 5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

### **6. Учебно-методическое и информационное обеспечение дисциплины**

#### 6.1. Список источников и литературы

#### 6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

### **7. Материально-техническое обеспечение дисциплины**

### **8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов**

### **9. Методические материалы**

#### 9.1. Планы лабораторных занятий

## **Приложения**

### Приложение 1. Аннотация дисциплины

### Приложение 2. Лист изменений

## 1. Пояснительная записка

### 1.1. Цель и задачи дисциплины

*Цель дисциплины:* формирование у студентов адекватных представлений о современных программных средствах прикладной статистики для их эффективного практического применения в различных областях научных исследований и инженерной практики.

*Задачи:* познакомить студентов с профессиональными статистическими пакетами (R, SPSS, Statistica и другими) для разработки алгоритмов и инструментальных средств обработки данных, а также для создания пользовательских интерфейсов к вычислительным процедурам; обсудить примеры применения этих пакетов для решения прикладных задач.

### 1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине

Коды компетенций	Содержание компетенций	Перечень планируемых результатов обучения по дисциплине
ПК-1	способность использовать стандартные пакеты прикладных программ для решения практических задач на электронных вычислительных машинах, отлаживать, тестировать прикладное программное обеспечение	<p><i>Знать:</i> основные современные статистические пакеты прикладных программ и технологии их использования</p> <p><i>Уметь:</i> применять статистические пакеты для решения прикладных задач, разрабатывать алгоритмы и статистические инструментальные средства для анализа и обработки данных, а также создавать интерфейсы к вычислительным процедурам</p> <p><i>Владеть:</i> навыками работы с программными системами для статистического анализа данных и разработки новых инструментальных средств при решении прикладных задач</p>

### 1.3. Место дисциплины в структуре образовательной программы

Дисциплина «Статистические пакеты прикладных программ» относится к вариативной части блока дисциплин учебного плана.

Для освоения дисциплины необходимы компетенции, сформированные в ходе изучения следующих дисциплин (модулей): «Математический анализ», «Линейная алгебра», «Аналитическая геометрия», «Дискретная математика», «Теория вероятностей», «Математическая статистика и теория случайных процессов», «Прикладная статистика», «Численные методы», «Иностранный язык».

В результате освоения дисциплины формируются компетенции, необходимые для изучения следующих дисциплин и прохождения практик: «Математическое моделирование», «Имитационное моделирование случайных процессов», «Методы принятия решений», Производственная практика (Научно-исследовательская работа), преддипломная практика.

## 2. Структура дисциплины

Общая трудоёмкость дисциплины составляет 3 з.е., 108 ч., в том числе контактная работа обучающихся с преподавателем 42 ч., самостоятельная работа обучающихся 66 ч.

№ п/п	Раздел дисциплины/темы	Семестр	Виды учебной работы (в часах)				Формы текущего контроля успеваемости, форма промежуточной аттестации (по семестрам)
			контактная		Промежуточная аттестация	Самостоятельная работа	
			Лекции	Лабораторные работы			
1	Обзор современных программных средств статистического анализа данных	7	2	4		10	Опрос
2	Пакет Statistica	7	4	4		10	Опрос
3	Пакет SPSS	7	4	4		10	Опрос Тестирование № 1
4	Вычислительная среда и язык программирования R	7	4	4		10	Опрос Тестирование № 2
5	Статистические средства универсальных математических пакетов	7	2	4		10	Опрос Тестирование № 3
6	Статистические средства офисных пакетов	7	2	2		10	Опрос Контрольная работа
	Зачёт с оценкой	7		2		6	Ответы на теоретические вопросы, итоговая контрольная работа
	Итого:		18	24		66	

## 3. Содержание дисциплины

### Тема 1. Обзор современных программных средств статистического анализа данных.

Классификация задач прикладной статистики и методов их решения. Выбор соответствующего задаче метода обработки данных. Виды статистических пакетов. Требования к статистическим пакетам общего назначения. Сравнительный анализ возможностей по обработке данных, которые предоставляют системы Statistica, SPSS, R, SAS, Stata, Minitab, Statgraphics, Microsoft Excel Analysis ToolPak, STADIA, МЕЗОЗАВР, ЭВРИСТА, САНИ и др. Особенности российского рынка программного обеспечения прикладной статистики.

### Тема 2. Пакет Statistica.

Архитектура пакета Statistica. Интерфейс пользователя. Управление данными. Дизайн и сопровождение статистических баз данных. Встроенный язык программирования STATISTICA Visual Basic. Использование внешних языков программирования. Сетевые возможности пакета. Обзор статистических методов, реализованных в пакете.

Графические инструменты анализа данных. Технология обработки данных и подготовки отчетов. Многомерные статистические методы в пакете Statistica. Корреляционный и регрессионный анализ, анализ таблиц сопряженности, кластерный и дискриминантный анализ, факторный анализ, дисперсионный анализ, многомерное шкалирование и анализ надёжности, методы статистического контроля качества продукции, анализ выживаемости. Анализ временных рядов и прогнозирование. Моделирование структурными уравнениями (модуль SEPATH). Нейросетевой пакет STATISTICA Neural Networks и его использование для решения прикладных задач.

### **Тема 3. Пакет SPSS.**

Архитектура пакета SPSS. Специфика оконного интерфейса. Редактор данных и вывод результатов вычислений. Дизайн и сопровождение статистических баз данных. Собственные средства программирования системы SPSS. Пакеты «R Essentials» и «Python Essentials». Интеграция SPSS с другими средствами статистических вычислений и языками программирования. Сетевые возможности SPSS. Обзор статистических методов, реализованных в пакете. Графические инструменты анализа данных. Технология обработки данных и подготовки отчетов. Многомерные статистические методы в пакете SPSS. Корреляционный и регрессионный анализ, анализ таблиц сопряженности, кластерный и дискриминантный анализ, факторный анализ, дисперсионный анализ, многомерное шкалирование и анализ надёжности, методы статистического контроля качества продукции, анализ выживаемости. Анализ временных рядов и прогнозирование. Моделирование структурными уравнениями. Нейросетевой модуль Neural Networks и его использование для решения прикладных задач.

### **Тема 4. Вычислительная среда и язык программирования R.**

Исторические сведения о среде статистических вычислений и языке программирования R. Установка R в различных операционных системах. Режим командной строки, скрипты, базовые и рекомендованные пакеты. Сообщество разработчиков, техническая поддержка пользователей, документация, книги, журналы, регулярные международные конференции по языку R и его приложениям. Типы данных в R и принципы работы с ними. Числовые векторы, факторы, пропущенные данные, матрицы, списки. Таблицы данных. Векторизованные вычисления. Графические средства языка R. Два типа графических команд. Графические устройства и графические опции. Сохранение результатов работы. Статистическая обработка данных. Описательная статистика. Одномерные статистические тесты. Создание своих функций. Параметрические и непараметрические критерии проверки однородности выборок. Проверка гипотез нормальности распределения. Корреляционный анализ и анализ таблиц сопряженности. Графические интерфейсы к R: RCommander, RKWard, JGR, SciViews-K, Rattle, PMG, RPMG, RWeb, gnumeric, Emacs и др. Поддержка работы с языком R в текстовых редакторах и средах разработки. Интеграция R с системами SPSS и Statistica. Интеллектуальный анализ данных (Data Mining) с помощью R. Графический анализ многих переменных. Сокращение размерности. Классификация без обучения. Кластерный анализ. Классификация с обучением.

### **Тема 5. Статистические средства универсальных математических пакетов.**

Статистические средства и инструментальные средства разработки универсальных математических пакетов MathCAD, Mathematica, MatLab, Maple, Maxima и др. Создание пользовательских интерфейсов к вычислительным процедурам.

### Тема 6. Статистические средства офисных пакетов.

Microsoft Excel Analysis ToolPak. Установка пакета и специфика интерфейса. Использование диалоговых окон. Подготовка данных. Создание, редактирование и печать диаграмм. Инструментарий статистического анализа данных и его использование. Описательная статистика. Генераторы случайных чисел. Создание выборки. Корреляции и ковариации. Двухвыборочный F-тест. Т-тест двухвыборочный с одинаковыми и неодинаковыми дисперсиями. Т-тест парный двухвыборочный для средних. Однофакторный и двухфакторный дисперсионный анализ с повторением и без повторения. Экспоненциальное сглаживание. Скользящее среднее. Парная и множественная линейная регрессия. Другие статистические средства офисных пакетов.

#### 4. Образовательные технологии

№ п/п	Наименование раздела	Виды учебных занятий	Образовательные технологии
1	2	3	4
1	Обзор современных программных средств статистического анализа данных	Лекции 1-2  Лабораторные занятия 1-2  Самостоятельная работа	Вводная лекция с использованием видеоматериалов и ресурсов сети Интернет, лекция с применением проектора  Решение типовых задач для закрепления и формирования знаний, умений, навыков  Дистанционный курс в сети Интернет, решение задач
2	Пакет Statistica	Лекции 3-4  Лабораторные занятия 3-5  Самостоятельная работа	Лекция-беседа с применением ИКТ  Решение типовых задач для закрепления и формирования знаний, умений, навыков  Дистанционный курс в сети Интернет, решение задач
3	Пакет SPSS	Лекции 5-6  Лабораторные занятия 6-8  Самостоятельная работа	Лекция-беседа с применением ИКТ  Решение типовых задач для закрепления и формирования знаний, умений, навыков  Дистанционный курс в сети Интернет, решение задач
4	Вычислительная среда и язык программирования R	Лекции 7-8  Лабораторные занятия 9-11  Самостоятельная работа	Лекция-беседа с применением ИКТ  Решение типовых задач для закрепления и формирования знаний, умений, навыков  Дистанционный курс в сети Интернет, решение задач
5	Статистические средства универсальных математических пакетов	Лекции 9-10  Лабораторные занятия 12-13  Самостоятельная работа	Лекция-беседа с применением ИКТ  Решение типовых задач для закрепления и формирования знаний, умений, навыков  Дистанционный курс в сети Интернет,

		работа	решение задач
6	Статистические средства офисных пакетов	Лекции 11-12  Лабораторные занятия 14-15  Самостоятельная работа	Лекция-беседа с применением ИКТ  Решение типовых задач для закрепления и формирования знаний, умений, навыков  Дистанционный курс в сети Интернет, решение задач

## 5. Оценка планируемых результатов обучения

### 5.1. Система оценивания

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль: - опрос - тестирование - контрольная работа	5 баллов 10 баллов 10 баллов	20 баллов 30 баллов 10 баллов
Промежуточная аттестация: - ответы на теоретические вопросы - итоговая контрольная работа		20 баллов 20 баллов
<b>Итого за семестр (дисциплину)</b> Экзамен		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

### 5.2. Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A, B	«зачтено (отлично)»	Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации. Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.



Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		<p>Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	«зачтено (хорошо)»	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D, E	«зачтено (удовлетворительно)»	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p>
49-0/ F, FX	«незачтено»	<p>Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами.</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		<p>Демонстрирует фрагментарные знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.</p>

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

### Текущий контроль

**Вопросы для опроса** см. п.9.1 РПД, контрольные вопросы

### **Примерные задания для тестирования № 1 по теме «Pакem Statistica»:**

Решите задачи, используя систему Statistica и файл с данными NEO PI-R.sta, который содержит результаты социологического опроса и личностные психологические показатели студентов РГГУ.

Задача 1.

Для девушек, степень религиозности которых сильная, среднее значение переменной E1\_Доброжелательность (с точностью до 0,01) равно

Ответ 30,23

Задача 2.

С помощью критерия Стьюдента (Т-критерия) выясните, на каком уровне значимости (с точностью до 0,001) различаются генеральные средние показатели N3\_Депрессивность для юношей и девушек.

Ответ 0,004

Задача 3.

С помощью критерия Колмогорова-Смирнова выясните, какие из приведенных ниже психологических показателей статистически значимо различаются для студентов факультета информатики (ФИ) и историко-филологического факультета (ИФФ).

Ответ 1. N3\_Депрессивность

Ответ 2. N4\_Застенчивость

Ответ 3. E3\_Настойчивость

Ответ 4. +E4\_Активность

Ответ 5. +O3\_Чувства

Ответ 6. +O4\_Действия

Ответ 7. A3\_Альтруизм

Ответ 8. A4\_Уступчивость

Ответ 9. C3\_Ответственность

Ответ 10. C4\_Целеустремленность

Задача 4.

Коэффициент корреляции Спирмена пунктов 131 и 161 опросника NEO PI-R (с точностью до 0,001) равен  
 Ответ -0,393

Задача 5.

Для респондентов юношей постройте линейную регрессионную модель для психологического показателя **N2\_Враждебность** методом пошагового исключения независимых переменных, в качестве которых рассматривайте все остальные подшкалы теста NEO PI-R.

Коэффициент детерминации для полученной оптимальной модели с точностью до 0,001 равен

Ответ 0,610

Задача 6.

С помощью кластерного анализа методом К средних классифицируйте юношей с низким личным доходом на четыре класса, используя утверждения теста NEO PI-R от 121 до 1120

Для полученной классификации расстояние от респондента с номером 176 до центра кластера, в котором он находится, (с точностью до 0,001) равно

Ответ 0,882

Задача 7.

Постройте наилучшую теоретическую классификацию респондентов на две группы, соответствующих степени согласия с утверждением 122 : 2 - “не согласен”, 4 - “согласен”. При построении классификации используйте метод пошагового дискриминантного анализа с включением независимых переменных, в качестве которых рассматривайте все тридцать подшкал теста NEO PI-R.

Для построенной классификации процент правильно теоретически распознанных ответов респондентов 2 - “не согласен” с точностью до 0,1% равен

Ответ 91,2%

Задача 8.

Для респондентов юношей с помощью критерия Шапиро-Уилка выясните, какие из них выполните факторный анализ для респондентов девушек, используя данные по всем тридцати подшкалам теста NEO PI-R. Для выделения факторов примените метод Главных компонент с последующим Варимакс вращением. Классифицируйте подшкалы теста NEO PI-R, включив каждую из них в свою группу, соответствующую фактору, с которым у этой подшкалы наибольший (по абсолютной величине) коэффициент корреляции.

Используя построенную классификацию, укажите шкалы теста NEO PI-R из приведенного ниже списка, которые не пригодны для интерпретации фактора 2

Ответ 1. +N\_Нейротизм

Ответ 2. E\_Экстраверсия

Ответ 3. O\_Открытость опыту

Ответ 4. +A\_Согласие

Ответ 5. +C\_Сознательность

Задача 9.

Психометрическая подшкала **N3\_Депрессивность** теста NEO PI-R равна сумме восьми переменных (пунктов подшкалы) inv\_111, 141, inv\_171, 1101, 1131, 1161, 1191, 1221. Выполните анализ пригодности этой подшкалы.

Наибольший из коэффициентов корреляции подшкалы со своими пунктами с точностью до 0,001 равен

Ответ 0,569

Задача 10.

С помощью многомерного шкалирования (процедура ALSCAL) постройте двумерную модель множества всех подшкал теста NEO PI-R, используя данные только для респондентов с 51 до 350. При этом учитывайте, что шкала измерения данных Интервальная, а расстояние вычисляйте по формуле Расстояние Евклида.

Из приведенных ниже психологических показателей укажите три подшкалы, которые в построенной модели находятся дальше остальных (из этого списка) от подшкалы O1\_Фантазия

Ответ 1. N1\_Тревожность

Ответ 2. +N2\_Враждебность

Ответ 3. N3\_Депрессивность

Ответ 4. N4\_Застенчивость

Ответ 5. N5\_Импульсивность

Ответ 6. +N6\_Уязвимость

Ответ 7. E1\_Доброжелательность

Ответ 8. E2\_Общительность

Ответ 9. +E3\_Настойчивость

Ответ 10. E4\_Активность

### ***Примерные задания для тестирования № 2 по теме «Пакет SPSS»:***

Решите задачи, используя систему SPSS и файл с данными NEO PI-R SPSS.sav, который содержит результаты социологического опроса и личностные психологические показатели студентов РГГУ.

Задача 1.

Для девушек, степень религиозности которых сильная, среднее значение переменной N4\_Застенчивость (с точностью до 0,01) равно

Ответ 25,77

Задача 2.

С помощью критерия Стьюдента (Т-критерия) выясните, на каком уровне значимости различаются генеральные средние показателя N4\_Застенчивость для юношей и девушек.

Ответ 0,023

Задача 3.

С помощью критерия Манна-Уитни выясните, какие из приведенных ниже психологических показателей статистически значимо различаются для студентов факультета информатики (ФИ) и историко-филологического факультета (ИФФ).

Ответ 1. +N1\_Тревожность

Ответ 2. +N2\_Враждебность

Ответ 3. E1\_Доброжелательность

Ответ 4. E2\_Общительность

Ответ 5. +O1\_Фантазия

- Ответ 6. +O2\_Эстетичность  
 Ответ 7. A1\_Доверие  
 Ответ 8. A2\_Прямота  
 Ответ 9. +C1\_Компетентность  
 Ответ 10. +C2\_Организованность

Задача 4.

Выясните, для каких из приведенных ниже пар психологических показателей коэффициент корреляции Пирсона является статистически значимым.

- Ответ 1. N1\_Тревожность и E5\_Непоседливость  
 Ответ 2. +N2\_Враждебность и C3\_Ответственность  
 Ответ 3. C2\_Организованность и E5\_Непоседливость  
 Ответ 4. +N3\_Депрессивность и C1\_Компетентность  
 Ответ 5. +C4\_Целеустремленность и N1\_Тревожность

Задача 5.

Постройте линейную регрессионную модель для психологического показателя **A2\_Прямота** методом пошагового включения независимых переменных, в качестве которых рассматривайте все остальные подшкалы теста NEO PI-R.

Коэффициент детерминации для модели, содержащей 5 самых важных независимых переменных, с точностью до 0,001 равен

Ответ 0,417

Задача 6.

С помощью иерархического кластерного анализа классифицируйте тридцать подшкал теста NEO PI-R на пять классов, используя данные только для множества респондентов с 51 до 350. В качестве метода кластеризации примените метод **Внутригрупповые связи**, а расстояние вычисляйте по формуле **Расстояние Евклида**.

По результатам классификации выясните, какие из приведенных ниже психологических показателей относятся к кластеру 2

- Ответ 1. N1\_Тревожность  
 Ответ 2. N2\_Враждебность  
 Ответ 3. +E1\_Доброжелательность  
 Ответ 4. E2\_Общительность  
 Ответ 5. +O1\_Фантазия  
 Ответ 6. +O2\_Эстетичность  
 Ответ 7. +A1\_Доверие  
 Ответ 8. +A2\_Прямота  
 Ответ 9. C1\_Компетентность  
 Ответ 10. C2\_Организованность

Задача 7.

Для множества респондентов с 1 до 300 постройте наилучшую теоретическую классификацию студентов на две группы - “мужчины” и “женщины”, используя метод пошагового дискриминантного анализа с включением и исключением независимых переменных, в качестве которых рассматривайте все тридцать подшкал теста NEO PI-R. При вычислении учитывайте относительные размеры групп.

Для построенной классификации процент правильно теоретически распознанных респондентов девушек с точностью до 0,1% равен

Ответ 97,2%

## Задача 8.

Выполните факторный анализ для множества респондентов с 51 до 350, используя данные по всем тридцати подшкалам теста NEO PI-R. Для выделения факторов примените метод Главных компонент с последующим Варимакс вращением.

Накопленный процент объясненной дисперсии данных для 4 извлеченных факторов с точностью до 0,001 равен

Ответ 54,378%

## Задача 9.

Психометрическая подшкала E1\_Доброжелательность теста NEO PI-R равна сумме восьми переменных (пунктов подшкалы) I2, inv\_I32, I62, inv\_I92, I122, I152, I182, I212. Выполните анализ пригодности этой подшкалы.

Показатель надёжности альфа Кронбаха для этой подшкалы с точностью до 0,001 равен

Ответ 0,715

## Задача 10.

С помощью многомерного шкалирования (процедура ALSCAL) постройте двумерную модель множества всех подшкал теста NEO PI-R, используя данные только для респондентов с 51 до 350. При этом учитывайте, что шкала измерения данных Интервальная, а расстояние вычисляйте по формуле Расстояние Евклида.

В построенной модели расстояние в двухмерном пространстве от подшкалы N2\_Враждебность до ближайшей к ней подшкалы с точностью до 0,001 равно

Ответ 0,093

### ***Примерные задания для тестирования № 3***

#### ***по теме «Вычислительная среда и язык программирования R»:***

Решите задачи, используя вычислительную среду R и файл с данными NEO, который содержит результаты социологического опроса и личностные психологические показатели студентов РГГУ.

## Задача 1.

Инициализировать датчик случайных чисел с номером 2013000 и сгенерировать выборку объёма  $n=230$  из генеральной совокупности, имеющей показательный закон распределения с параметром  $rate=0.4$ . Найти с точностью до 0.01 выборочную квантиль на уровне 0.95.

Ответ 8.06

## Задача 2.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, найти число юношей, для которых значение переменной N6\_Уязвимость больше 28.

Ответ 4

## Задача 3.

Инициализировать датчик случайных чисел с номером 2013000 и сгенерировать выборку объёма  $n=230$  из генеральной совокупности, имеющей закон распределения

Пуассона с параметром  $\lambda=8.1$ . По полученной выборке найти методом моментов с точностью до 0.01 точечную оценку параметра  $\lambda$ , используя центральный момент второго порядка.

Ответ 7.71

Задача 4.

Инициализировать датчик случайных чисел с номером 2013000 и сгенерировать выборку объёма  $n=300$  из генеральной совокупности, имеющей нормальный закон распределения с параметрами  $\text{mean}=172$ ,  $\text{sd}=6.4$ . По полученной выборке найти с надёжностью  $p=0.95$  бутстреп-оценку доверительного интервала для математического ожидания, используя в качестве точечной оценки среднее арифметическое. Вычисления выполнить на основе 10000 вторичных выборок с объёмом 300 элементов каждая. В ответе указать длину доверительного интервала с точностью до 0.01.

Ответ 1.50

Задача 5.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, с помощью критерия Колмогорова-Смирнова проверить статистическую гипотезу о том, что для респондентов юношей переменная O1\_Фантазия имеет закон распределения, который статистически значимо не отличается от нормального закона распределения. В ответе задачи указать значение  $p\text{-value}$  с точностью до 0.001.

Ответ 0.132

Задача 6.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, с помощью критерия Краскела-Уоллиса проверить статистическую гипотезу о том, что уровень переменной O1\_Фантазия не зависит от семейного дохода респондентов. В ответе задачи указать значение  $p\text{-value}$  с точностью до 0.001.

Ответ 0.286

Задача 7.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, с помощью Хи-квадрат критерия Пирсона проверить статистическую гипотезу о том, что ответы респондентов на пункт П\_41 опросника NEO PI-R не зависят от степени религиозности. В ответе задачи указать значение  $p\text{-value}$  с точностью до 0.001.

Ответ 0.346

Задача 8.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, выяснить на уровне значимости 0.05, какие из перечисленных ниже порядковых демографических переменных имеют статистически значимый коэффициент корреляции Кендалла с психологическим показателем E5\_Непоседливость.

Ответ 1. +возраст

Ответ 2. +обр\_род (образование родителей)

Ответ 3. степ\_рел (степень религиозности)

Ответ 4. сем\_дох (семейный доход)

Ответ 5. лич\_дох (личный доход)

#### Задача 9.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, построить оптимальную линейную регрессионную модель m26, содержащую 8 предикторов и переменную отклика N6\_Уязвимость, пошаговым методом добавления независимых переменных, в качестве которых рассматривать все подшкалы теста NEO PI-R кроме показателя N6\_Уязвимость. Найти с точностью до 0.001 коэффициент детерминации модели m26.

Ответ 0.690

#### Задача 10.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, построить оптимальную линейную регрессионную модель m25 для психологического показателя N6\_Уязвимость пошаговым методом добавления независимых переменных, в качестве которых рассматривать все остальные подшкалы теста NEO PI-R. Используя модель m25, найти с точностью до 0.01 прогноз значения зависимой переменной N6\_Уязвимость для респондента с номером 208.

Ответ 23.40

### **Примерные задания для контрольной работы по теме «Проверка статистических гипотез в пакете SPSS»:**

Решите задачи, используя систему SPSS и файл с данными NEO PI-R SPSS.sav, который содержит результаты социологического опроса и личностные психологические показатели студентов РГГУ.

#### Задача 1.

С помощью критерия Колмогорова-Смирнова выясните, на каком уровне значимости закон распределения показателя A6\_Отзывчивость отличается от нормального закона распределения.

Ответ 0,017

#### Задача 2.

С помощью критерия Стьюдента (Т-критерия) выясните, на каком уровне значимости различаются генеральные средние показателя N4\_Застенчивость для юношей и девушек.

Ответ 0,023

#### Задача 3.

С помощью критерия Стьюдента (Т-критерия) выясните, на каком уровне значимости различаются генеральные средние показателя O6\_Ценности для студентов с сильной и слабой степенью религиозности.

Ответ 0,002



## Задача 4.

С помощью однофакторного дисперсионного анализа выясните, какие из приведенных ниже психологических показателей статистически значимо зависят от степени религиозности.

- Ответ 1. N1\_Тревожность
- Ответ 2. N2\_Враждебность
- Ответ 3. E1\_Доброжелательность
- Ответ 4. E2\_Общительность
- Ответ 5. O1\_Фантазия
- Ответ 6. O2\_Эстетичность
- Ответ 7. A1\_Доверие
- Ответ 8. +A2\_Прямота
- Ответ 9. C1\_Компетентность
- Ответ 10. C2\_Организованность

## Задача 5.

С помощью однофакторного дисперсионного анализа выясните, от каких из приведенных ниже утверждений опросника NEO PI-R статистически значимо зависит показатель O1\_Фантазия.

- Ответ 1. I6
- Ответ 2. +I7
- Ответ 3. +I8
- Ответ 4. +I9
- Ответ 5. +I10

## Задача 6.

С помощью критерия Манна-Уитни выясните, на каком уровне значимости различаются генеральные средние показателя E3\_Настойчивость для юношей и девушек.

Ответ 0,003

## Задача 7.

С помощью критерия Колмогорова-Смирнова выясните, на каком уровне значимости различаются генеральные средние показателя E6\_Жизнерадостность для юношей и девушек.

Ответ 0,002

## Задача 8.

С помощью критерия Манна-Уитни выясните, какие из приведенных ниже психологических показателей статистически значимо различаются для студентов факультета информатики (ФИ) и историко-филологического факультета (ИФФ).

- Ответ 1. +N3\_Депрессивность
- Ответ 2. +N4\_Застенчивость
- Ответ 3. +E3\_Настойчивость
- Ответ 4. E4\_Активность
- Ответ 5. +O3\_Чувства
- Ответ 6. +O4\_Действия
- Ответ 7. A3\_Альтруизм
- Ответ 8. A4\_Уступчивость
- Ответ 9. C3\_Ответственность

Ответ 10. С4\_Целеустремленность

Задача 9.

С помощью критерия Крускала-Уоллеса выясните, какие из приведенных ниже психологических показателей статистически значимо зависят от степени религиозности.

- Ответ 1. N3\_Депрессивность
- Ответ 2. N4\_Застенчивость
- Ответ 3. E3\_Настойчивость
- Ответ 4. E4\_Активность
- Ответ 5. O3\_Чувства
- Ответ 6. O4\_Действия
- Ответ 7. A3\_Альтруизм
- Ответ 8. A4\_Уступчивость
- Ответ 9. C3\_Ответственность
- Ответ 10. С4\_Целеустремленность

Задача 10.

С помощью критерия Крускала-Уоллеса выясните, от каких из приведенных ниже утверждений опросника NEO PI-R статистически значимо зависит показатель

A3\_Альтруизм.

- Ответ 1. +I21
- Ответ 2. +I22
- Ответ 3. I23
- Ответ 4. I24
- Ответ 5. I25

### **Промежуточная аттестация (зачет с оценкой)**

#### ***Контрольные вопросы по дисциплине:***

1. Классификация задач прикладной статистики и методов их решения.
2. Виды статистических пакетов. Требования к статистическим пакетам.
3. Архитектура пакета Statistica. Интерфейс пользователя. Управление данными.
4. Встроенный язык программирования STATISTICA Visual Basic.
5. Многомерные статистические методы в пакете Statistica.
6. Моделирование структурными уравнениями (модуль SEPATH).
7. Анализ временных рядов и прогнозирование в системе Statistica.
8. Нейросетевой пакет STATISTICA Neural Networks и его применение.
9. Архитектура пакета SPSS. Специфика оконного интерфейса. Редактор данных и вывод результатов вычислений.
10. Собственные средства программирования системы SPSS. Интеграция SPSS с другими средствами статистических вычислений и языками программирования.
11. Многомерные статистические методы в пакете SPSS.
12. Анализ временных рядов и прогнозирование в системе SPSS.
13. Моделирование структурными уравнениями в SPSS.
14. Нейросетевой модуль Neural Networks и его применение
15. Общие сведения о среде статистических вычислений и языке программирования R.
16. Типы данных в R: векторы, факторы, пропущенные данные, матрицы, списки. Таблицы данных. Векторизованные вычисления.

17. Графические средства языка R. Два типа графических команд. Графические устройства и графические опции.
18. Статистическая обработка данных в системе R. Описательная статистика.
19. Проверка статистических гипотез в системе R.
20. Корреляционный анализ и анализ таблиц сопряженности. Регрессионный анализ в системе R.
21. Интеллектуальный анализ данных (Data Mining) с помощью R. Графический анализ многих переменных.
22. Статистические средства универсальных математических пакетов.
23. Microsoft Excel Analysis ToolPak. Другие статистические средства офисных пакетов.

***Примерные задания для итоговой контрольной работы:***

Для решения задач использовать фрейм данных NEO из файла с именем "NEO", который содержит результаты социологического опроса и личностные психологические показатели студентов.

***Вариант 1.***

**Задача 1.**

С помощью критерия Шапиро-Уилка проверить статистическую гипотезу о том, что для респондентов девушек переменная N1\_Тревожность имеет закон распределения, который статистически значимо не отличается от нормального закона распределения. В ответе задачи указать значение p-value с точностью до 0.001.

**Задача 2.**

С помощью критерия Колмогорова-Смирнова проверить статистическую гипотезу о том, что для респондентов юношей переменная N1\_Тревожность имеет закон распределения, который статистически значимо не отличается от нормального закона распределения. В ответе задачи указать значение p-value с точностью до 0.001.

**Задача 3.**

С помощью критерия Стьюдента проверить гипотезу о том, что математическое ожидание переменной N1\_Тревожность для студентов факультета «А» равно 27. В ответе задачи указать значение p-value с точностью до 0.001.

**Задача 4.**

С помощью критерия Стьюдента проверить гипотезу о том, что для юношей и девушек математические ожидания переменной N1\_Тревожность равны. В ответе задачи указать значение p-value с точностью до 0.001.

**Задача 5.**

С помощью критерия Уилкоксона проверить гипотезу о том, что положение переменной N1\_Тревожность для студентов факультета «А» равно 27. В ответе задачи указать значение p-value с точностью до 0.001.

**Задача 6.**

С помощью критерия Уилкоксона проверить гипотезу о том, что для респондентов с сильной и слабой степенью религиозности уровни переменной N1\_Тревожность равны. В ответе задачи указать значение p-value с точностью до 0.001.

**Задача 7.**

С помощью теста пропорций проверить статистическую гипотезу о том, что для студентов факультета «А» и студентов факультета «Б» доли респондентов, согласных с утверждением П\_31 опросника NEO PI-R, равны. В ответе задачи указать значение p-value с точностью до 0.001.

#### **Задача 8.**

С помощью Хи-квадрат критерия Пирсона проверить статистическую гипотезу о том, что ответы респондентов на пункт П\_31 опросника NEO PI-R не зависят от степени религиозности. В ответе задачи указать значение p-value с точностью до 0.001.

#### **Задача 9.**

С помощью однофакторного дисперсионного анализа проверить статистическую гипотезу о том, что математическое ожидание переменной N1\_Тревожность не зависит от семейного дохода респондентов. В ответе задачи указать значение p-value с точностью до 0.001.

#### **Задача 10.**

С помощью критерия Краскела-Уоллиса проверить статистическую гипотезу о том, что уровень переменной N1\_Тревожность не зависит от семейного дохода респондентов. В ответе задачи указать значение p-value с точностью до 0.001.

### **6. Учебно-методическое и информационное обеспечение дисциплины**

#### **6.1. Список источников и литературы**

##### **Литература**

##### *Основная*

1. Наследов А. SPSS 19: профессиональный статистический анализ данных. - СПб.: Питер, 2011. - 400 с.
2. Плехотников, К. Э. Основы эконометрики в пакете STATISTICA.: Учебное пособие / К.Э. Плехотников. - Москва : Вузовский учебник, 2010. - 298 с. + CDROM. ISBN 978-5-9558-0114-8. - Текст: электронный. - URL: <https://znanium.com/catalog/product/177719>

##### *Дополнительная*

1. Маркетинговые исследования с SPSS: Учебное пособие / Г. Моосмюллер, Н.Н. Ребик. - 2-е изд. - М.: ИНФРА-М, 2011. - 200 с.: 60х90 1/16. - (Высшее образование). (обложка) ISBN 978-5-16-004240-4 - Режим доступа: <http://znanium.com/catalog/product/257371>
2. Статистические методы обработки экспериментальных данных с использованием пакета MathCad: Учебное пособие/Ф.И.Карманов, В.А.Острейковский - М.: КУРС, НИЦ ИНФРА-М, 2015. - 208 с.: 60х90 1/16 (Обложка) ISBN 978-5-905554-96-4 - Режим доступа: <http://znanium.com/catalog/product/508241>
3. Тюрин Ю. Н. Анализ данных на компьютере: учеб. пособие по направлениям "Математика", "Математика. Прикладная математика" / Ю. Н. Тюрин, А. А. Макаров. - Изд. 4-е, перераб. - М.: Форум, 2013. - 366 с.- (Высшее образование)

#### **6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

1. Чубукова И.А. Data Mining. Учеб. курс НОУ ИНТУИТ [Электронный ресурс]. - Режим доступа: <http://www.intuit.ru/department/database/datamining/>
2. Учебно-образовательная физико-математическая библиотека на портале МИР МАТЕМАТИЧЕСКИХ УРАВНЕНИЙ [Электронный ресурс]. - Режим доступа: <http://eqworld.ipmnet.ru/ru/library.htm>
3. Официальный портал проекта R [Электронный ресурс]. - Режим доступа: <http://www.r-project.org/>
4. Сетевые архивы системы R (CRAN). [Электронный ресурс]. - Режим доступа: <http://cran.r-project.org/>
5. R — объектно-ориентированная статистическая среда [Электронный ресурс]. - Режим доступа: <http://herba.msu.ru/shipunov/software/r/r-ru.htm>
6. Язык и среда R [Электронный ресурс]. - Режим доступа: <http://r-statistics.livejournal.com/>

### Перечень современных профессиональных баз данных (БД) и информационно-справочных систем (ИСС)

№п/п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2017 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2017 г. Журналы Oxford University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Компьютерные справочные правовые системы Консультант Плюс, Гарант

### 7. Материально-техническое обеспечение дисциплины

Для материально-технического обеспечения дисциплины необходимы:

- лаборатория,
- компьютер или ноутбук для преподавателя,
- компьютеры для обучающихся,
- доска,
- проектор (стационарный или переносной),
- программное обеспечение (ПО).

### Перечень программного обеспечения (ПО)

№п/п	Наименование ПО	Производитель	Способ распространения
1	Microsoft Office 2010 Pro	Microsoft	лицензионное
2	Windows XP или Windows 7	Microsoft	лицензионное
3	Язык программирования R	сообщество	свободно распространяемое
4	Язык программирования	сообщество	свободно распространяемое

	Python		
5	Matlab	The MathWorks	лицензионное
6	Mathcad Education - University edition	PTC	лицензионное
7	Kaspersky Endpoint Security	Kaspersky	лицензионное

Для лабораторных занятий можно также использовать актуальные полнофункциональные демонстрационные версии профессиональных статистических пакетов SPSS и Statistica.

## **8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов**

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих:
  - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
  - письменные задания выполняются на компьютере со специализированным программным обеспечением, или могут быть заменены устным ответом;
  - обеспечивается индивидуальное равномерное освещение не менее 300 люкс;
  - для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств;
  - письменные задания оформляются увеличенным шрифтом;
  - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.
- для глухих и слабослышащих:
  - лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования;
  - письменные задания выполняются на компьютере в письменной форме;
  - экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.
- для лиц с нарушениями опорно-двигательного аппарата:
  - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
  - письменные задания выполняются на компьютере со специализированным программным обеспечением;
  - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих:
  - в печатной форме увеличенным шрифтом;
  - в форме электронного документа;
  - в форме аудиофайла.
- для глухих и слабослышащих:
  - в печатной форме;
  - в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата:
  - в печатной форме;
  - в форме электронного документа;
  - в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих:
  - устройством для сканирования и чтения с камерой SARA CE;
  - дисплеем Брайля PAC Mate 20;
  - принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих:
  - автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих;
  - акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата:
  - передвижными, регулируемые эргономическими партами СИ-1;
  - компьютерной техникой со специальным программным обеспечением.

## **9. Методические материалы**

### **9.1. Планы лабораторных занятий**

#### **Тема 1. Пакет Statistica.**

**Цель работы** - знакомство с технологией статистического анализа данных в пакете Statistica.

#### **Контрольные вопросы:**

1. Описательная статистика в пакете Statistica.
2. Проверка статистических гипотез в пакете Statistica.
3. Дисперсионный анализ в пакете Statistica.
4. Корреляционный анализ в пакете Statistica.
5. Множественная линейная регрессия в пакете Statistica.
6. Кластерный анализ в пакете Statistica.
7. Дискриминантный анализ в пакете Statistica.
8. Факторный анализ в пакете Statistica.
9. Анализ надежности в пакете Statistica.
10. Многомерное шкалирование в пакете Statistica.
11. Статистический анализ временных рядов в пакете Statistica.
12. Графические методы анализа данных в пакете Statistica.

#### **Примерные задачи для решения в аудитории:**

При решении задач рекомендуется использовать файл с данными NEO PI-R.sta, который содержит результаты социологического опроса и личностные психологические показатели студентов.

Задача 1.

Для девушек, степень религиозности которых слабая, среднее значение переменной E2\_Общительность (с точностью до 0,01) равно

Ответ 25,78

Задача 2.

С помощью критерия Стьюдента (Т-критерия) выясните, на каком уровне значимости (с точностью до 0,001) различаются генеральные средние показателя N2\_Враждебность для юношей и девушек.

Ответ 0,005

Задача 3.

С помощью критерия Манна-Уитни выясните, на каком уровне значимости различаются генеральные средние показателя E5\_Непоседливость для девушек с сильной и слабой степенью религиозности.

Ответ 0,024

Задача 4.

Коэффициент корреляции Спирмена пунктов I31 и I51 опросника NEO PI-R (с точностью до 0,001) равен

Ответ 0,184

Задача 5.

Для респондентов юношей постройте линейную регрессионную модель для психологического показателя N1\_Тревожность методом пошагового исключения независимых переменных, в качестве которых рассматривайте все остальные подшкалы теста NEO PI-R. Коэффициент детерминации для полученной оптимальной модели с точностью до 0,001 равен

Ответ 0,671

Задача 6.

С помощью кластерного анализа методом К средних классифицируйте юношей с низким личным доходом на четыре класса, используя утверждения теста NEO PI-R от I21 до I120. Для полученной классификации расстояние от респондента с номером 148 до центра кластера, в котором он находится, (с точностью до 0,001) равно

Ответ 0,705

Задача 7.

Для множества респондентов с 31 до 230 постройте наилучшую теоретическую классификацию студентов на две группы - "мужчины" и "женщины", используя метод пошагового дискриминантного анализа с включением независимых переменных, в качестве которых рассматривайте все тридцать подшкал теста NEO PI-R. Для построенной классификации процент правильно теоретически распознанных респондентов девушек с точностью до 0,1% равен

Ответ 94,0



Задача 8.

Выполните факторный анализ для множества респондентов с 31 до 230, используя данные по всем тридцати подшкалам теста NEO PI-R. Для выделения факторов примените метод Главных компонент с последующим Варимакс вращением. Накопленный процент объясненной дисперсии данных для 5 извлеченных факторов с точностью до 0,001 равен

Ответ 59,228%

Задача 9.

Психометрическая подшкала **N4\_Застенчивость** теста NEO PI-R равна сумме восьми переменных (пунктов подшкалы) 116, inv\_146, 176, inv\_1106, 1136, inv\_1166, 1196, inv\_1226. Выполните анализ пригодности этой подшкалы. Показатель надёжности альфа Кронбаха для этой подшкалы с точностью до 0,001 равен

Ответ 0,753

Задача 10.

С помощью многомерного шкалирования (процедура ALSCAL) постройте двумерную модель множества всех подшкал теста NEO PI-R, используя данные только для респондентов с 51 до 350. При этом учитывайте, что шкала измерения данных **Интервальная**, а расстояние вычисляйте по формуле **Расстояние Евклида**. В построенной модели расстояние в двумерном пространстве от подшкалы **O1\_Фантазия** до ближайшей к ней подшкалы с точностью до 0,001 равно

Ответ 0,257

### Справочные материалы

Шкалы теста NEO PI-R	Их подшкалы
N Нейротизм	N1 Тревожность N2 Враждебность N3 Депрессивность N4 Застенчивость N5 Импульсивность N6 Уязвимость
E Экстраверсия	E1 Доброжелательность E2 Общительность E3 Настойчивость E4 Активность E5 Непоседливость E6 Жизнерадостность
O Открытость опыту	O1 Фантазия O2 Эстетичность O3 Чувства O4 Действия O5 Идеи O6 Ценности
A Согласие	A1 Доверие A2 Прямота A3 Альтруизм A4 Уступчивость A5 Скромность A6 Отзывчивость
C Сознательность	C1 Компетентность

	C2 Организованность C3 Ответственность C4 Целеустремленность C5 Самодисциплина C6 Осмотрительность
--	--

## Тема 2. Пакет SPSS.

**Цель работы** - знакомство с технологией статистического анализа данных в пакете SPSS.

### Контрольные вопросы:

1. Описательная статистика в пакете SPSS.
2. Проверка статистических гипотез в пакете SPSS.
3. Дисперсионный анализ в пакете SPSS.
4. Корреляционный анализ в пакете SPSS.
5. Множественная линейная регрессия в пакете SPSS.
6. Кластерный анализ в пакете SPSS.
7. Дискриминантный анализ в пакете SPSS.
8. Факторный анализ в пакете SPSS.
9. Анализ надежности в пакете SPSS.
10. Многомерное шкалирование в пакете SPSS.
11. Статистический анализ временных рядов в пакете SPSS.
12. Графические методы анализа данных в пакете SPSS.

### Примерные задачи для решения в аудитории:

При решении задач рекомендуется использовать файл с данными NEO PI-R SPSS.sav, который содержит результаты социологического опроса и личностные психологические показатели студентов.

#### Задача 1.

Число респондентов, семейный доход которых низкий, равно

Ответ 22

#### Задача 2.

С помощью критерия Стьюдента (Т-критерия) выясните, какие из приведенных ниже психологических показателей статистически значимо различаются для юношей и девушек.

- Ответ 1. +N1\_Тревожность  
 Ответ 2. +N2\_Враждебность  
 Ответ 3. E1\_Доброжелательность  
 Ответ 4. E2\_Общительность  
 Ответ 5. +O1\_Фантазия  
 Ответ 6. +O2\_Эстетичность  
 Ответ 7. A1\_Доверие  
 Ответ 8. A2\_Прямота  
 Ответ 9. +C1\_Компетентность  
 Ответ 10. +C2\_Организованность

#### Задача 3.

С помощью критерия Колмогорова-Смирнова выясните, какие из приведенных ниже психологических показателей статистически значимо различаются для студентов факультета информатики (ФИ) и историко-филологического факультета (ИФФ).

- Ответ 1. +N1\_Тревожность
- Ответ 2. +N2\_Враждебность
- Ответ 3. E1\_Доброжелательность
- Ответ 4. E2\_Общительность
- Ответ 5. +O1\_Фантазия
- Ответ 6. +O2\_Эстетичность
- Ответ 7. A1\_Доверие
- Ответ 8. A2\_Прямота
- Ответ 9. C1\_Компетентность
- Ответ 10. +C2\_Организованность

Задача 4.

Выясните, какие из перечисленных ниже порядковых демографических переменных имеют статистически значимый коэффициент корреляции Спирмена с психологическим показателем E2\_Общительность.

- Ответ 1. +возраст
- Ответ 2. +обр\_род (образование родителей)
- Ответ 3. степ\_рел (степень религиозности)
- Ответ 4. сем\_дох (семейный доход)
- Ответ 5. +лич\_дох (личный доход)

Задача 5.

Постройте линейную регрессионную модель для психологического показателя C6\_Осмотрительность методом пошагового включения независимых переменных, в качестве которых рассматривайте все остальные подшкалы теста NEO PI-R.

Коэффициент детерминации для модели, содержащей 7 самых важных независимых переменных, с точностью до 0,001 равен

Ответ 0,448

Задача 6.

С помощью иерархического кластерного анализа классифицируйте тридцать подшкал теста NEO PI-R на пять классов, используя данные только для множества респондентов с 51 до 350. В качестве метода кластеризации примените метод Внутригрупповые связи, а расстояние вычисляйте по формуле Расстояние Евклида.

По результатам классификации выясните, какие из приведенных ниже психологических показателей относятся к кластеру 2

- Ответ 1. N3\_Депрессивность
- Ответ 2. N4\_Застенчивость
- Ответ 3. E3\_Настойчивость
- Ответ 4. E4\_Активность
- Ответ 5. +O3\_Чувства
- Ответ 6. O4\_Действия
- Ответ 7. +A3\_Альтруизм
- Ответ 8. A4\_Уступчивость
- Ответ 9. C3\_Ответственность
- Ответ 10. C4\_Целеустремленность

## Задача 7.

Для множества респондентов с 51 до 350 постройте наилучшую теоретическую классификацию студентов на две группы - “мужчины” и “женщины”, используя метод пошагового дискриминантного анализа с включением и исключением независимых переменных, в качестве которых рассматривайте все тридцать подшкал теста NEO PI-R. При вычислении учитывайте относительные размеры групп.

Используя построенную классификацию, укажите номера респондентов из приведенного ниже списка, для которых принадлежность к группе распознана неверно

- Ответ 1. +92
- Ответ 2. 93
- Ответ 3. 94
- Ответ 4. 95
- Ответ 5. 96
- Ответ 6. +97
- Ответ 7. 98
- Ответ 8. +99
- Ответ 9. 100
- Ответ 10. +101

## Задача 8.

Выполните факторный анализ для множества респондентов с 51 до 350, используя данные по всем тридцати подшкалам теста NEO PI-R. Для выделения факторов примените метод Главных компонент с последующим Варимакс вращением. Классифицируйте подшкалы теста NEO PI-R, включив каждую из них в свою группу, соответствующую фактору, с которым у этой подшкалы наибольший (по абсолютной величине) коэффициент корреляции.

Используя построенную классификацию, укажите подшкалы из приведенного ниже списка, которые включены в группу, соответствующую фактору 2

- Ответ 1. +A1\_Доверие
- Ответ 2. +A2\_Прямота
- Ответ 3. +A3\_Альтруизм
- Ответ 4. A4\_Уступчивость
- Ответ 5. A5\_Скромность
- Ответ 6. +A6\_Отзывчивость
- Ответ 7. C1\_Компетентность
- Ответ 8. C2\_Организованность
- Ответ 9. C3\_Ответственность
- Ответ 10. C4\_Целеустремленность

## Задача 9.

Психометрическая подшкала A3\_Альтруизм теста NEO PI-R равна сумме восьми переменных (пунктов подшкалы) inv\_114, 144, inv\_174, 1104, inv\_1134, 1164, 1194, 1224. Выполните анализ пригодности этой подшкалы.

Показатель надёжности альфа Кронбаха для этой подшкалы с точностью до 0,001 равен

- Ответ 0,658

## Задача 10.

С помощью многомерного шкалирования (процедура ALSCAL) постройте двумерную модель множества всех подшкал теста NEO PI-R, используя данные только

для респондентов с 51 до 350. При этом учитывайте, что шкала измерения данных Интервальная, а расстояние вычисляйте по формуле Расстояние Евклида.

Из приведенных ниже психологических показателей укажите три подшкалы, которые в построенной модели находятся дальше остальных (из этого списка) от подшкалы O1\_Фантазия

- Ответ 1. N1\_Тревожность
- Ответ 2. +N2\_Враждебность
- Ответ 3. N3\_Депрессивность
- Ответ 4. N4\_Застенчивость
- Ответ 5. N5\_Импульсивность
- Ответ 6. +N6\_Уязвимость
- Ответ 7. E1\_Доброжелательность
- Ответ 8. E2\_Общительность
- Ответ 9. +E3\_Настойчивость
- Ответ 10. E4\_Активность

### Тема 3. Вычислительная среда и язык программирования R.

**Цель работы** - знакомство с технологией статистического анализа данных в среде R.

#### Контрольные вопросы:

1. Описательная статистика в среде R.
2. Графические методы анализа данных в среде R.
3. Проверка статистических гипотез в среде R.
4. Дисперсионный анализ в среде R.
5. Корреляционный анализ в среде R.
6. Регрессионный анализ в среде R.

#### Примерные задачи для решения в аудитории:

При решении задач рекомендуется использовать файл с данными NEO, который содержит результаты социологического опроса и личностные психологические показатели студентов.

#### Задача 1.

Загрузить в рабочее пространство системы R данные из файла “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов. Используя фрейм данных с именем “NEO”, выполнить следующие задания: 1) вывести на экран имена всех переменных и найти количество строк и столбцов таблицы данных “NEO”; 2) выяснить структуру данных части таблицы “NEO”, содержащей первые пять столбцов; 3) получить сводную информацию о переменных с номерами 3, 12 и 17; 4) записать сводную информацию обо всех переменных таблицы “NEO” в текстовый файл “NEO.summary.txt”; 5) создать подмножество фрейма данных “NEO”, которое содержит информацию о студентках факультета “Б” с сильной степенью религиозности, и найти количество строк полученного фрейма данных; 6) графически исследовать степень религиозности респондентов (переменная “СТЕП\_РЕЛ”); 7) построить график переменной N6\_Уязвимость; 8) построить график зависимости личного дохода от пола респондента; 9) построить график зависимости показателя N6\_Уязвимость от пола респондента; 10) построить график зависимости показателя N2\_Враждебность от показателя N6\_Уязвимость.

## Задача 2.

Используя фрейм данных с именем “NEO”, выполнить следующие задания: 1) построить таблицу частот и таблицу относительных частот для переменной СТЕП\_РЕЛ; 2) составить список, компонентами которого являются таблицы частот для переменных с номерами от 2 до 7 (включительно); 3) построить таблицу частот для переменной N6\_Уязвимость; 4) для группированного вариационного ряда переменной N6\_Уязвимость построить таблицы частот, относительных частот, накопленных и накопленных относительных частот.

## Задача 3.

Используя фрейм данных с именем “NEO”, выполнить следующие задания: 1) построить полигон относительных частот для переменной N6\_Уязвимость и добавить на график кривую плотности распределения вероятностей нормального закона, параметрами которого считать выборочное среднее и выборочное стандартное отклонение переменной N6\_Уязвимость; 2) построить полигон накопленных относительных частот для переменной N6\_Уязвимость и добавить к нему график функции распределения нормального закона с такими же параметрами, как в задании 1; 3) выполнить задание 1 для группированного вариационного ряда переменной N6\_Уязвимость с 10 интервалами группировки; 4) выполнить задание 2 для группированного вариационного ряда переменной N6\_Уязвимость с 10 интервалами группировки.

## Задача 4.

Используя фрейм данных с именем “NEO”, найти эмпирическую функцию распределения для переменной N6\_Уязвимость. Построить её график. С помощью эмпирической функции распределения вычислить статистические вероятности следующих событий: А — переменная N6\_Уязвимость принимает значение не больше 20; В — переменная N6\_Уязвимость принимает значение на отрезке [25, 30]; С — переменная N6\_Уязвимость принимает значение больше 35.

## Задача 5.

Используя фрейм данных с именем “NEO”, найти для переменной N6\_Уязвимость следующие числовые характеристики выборки: объём  $n$ ; наименьшее и наибольшее значения  $\min$  и  $\max$ ; выборочное среднее (арифметическое)  $m$ ; медиану  $me$ ; нижнюю и верхнюю квартили  $q1$  и  $q3$ ; (исправленную) выборочную дисперсию  $var$ ; среднее квадратическое отклонение  $sd$ ; размах выборки  $r$ ; межквартильный размах  $iqr$ ; моду  $mo$ , начальные  $\alpha[i]$  и центральные  $\mu[i]$  моменты до 4 порядка включительно ( $i=1,2,3,4$ ), асимметрию  $as$ , эксцесс  $ex$ .

## Задача 6.

Используя фрейм данных с именем “NEO”, выполнить следующие задания: 1) для девушек факультета «Б», которые младше 19 лет, найти наибольшее значение переменной N6\_Уязвимость; 2) для юношей со слабой степенью религиозности найти с точностью до 0.001 среднее арифметическое переменной N6\_Уязвимость; 3) указать знак зодиака, для которого выборочная дисперсия показателя N6\_Уязвимость является наибольшей; 4) выбрать знаки зодиака, для которых выборочное среднее показателя N6\_Уязвимость меньше 22.

## Задача 7.

Используя фрейм данных с именем “NEO”, построить с помощью функции `boxplot()` следующие графики: 1) диаграмму распределения переменных `N6_Уязвимость`, `E6_Жизнерадостность`, `A6_Отзывчивость`, `C6_Осмотрительность`; 2) диаграмму зависимости распределения переменной `N6_Уязвимость` от уровня фактора `ФАКУЛЬТ`.

## Задача 8.

Используя фрейм данных с именем “NEO”, найти с надёжностью  $p=0.95$  интервальную бутстреп-оценку `bcir1` для коэффициента корреляции Пирсона случайных величин `N6_Уязвимость` и `C6_Осмотрительность`. Вычисление интервальной бутстреп-оценки выполнить на основе 10000 вторичных выборок с объёмом, равным объёму исходной выборки.

## Задача 9.

Инициализировать датчик случайных чисел с номером 2013000 и сгенерировать выборку объёма  $n=300$  из генеральной совокупности, имеющей нормальный закон распределения с параметрами `mean=172`, `sd=6.4`. По полученной выборке найти с надёжностью  $p=0.95$  интервальные бутстреп-оценки квантилей на уровнях 0.05, 0.25, 0.5, 0.75, 0.95. Вычисление интервальных бутстреп-оценок выполнить на основе 10000 вторичных выборок с объёмом 300 элементов каждая. Построить график зависимости интервальной бутстреп-оценки верхней квантили от доверительной вероятности.

## Задача 10.

Инициализировать датчик случайных чисел с номером 2013000 и сгенерировать выборку объёма  $n=300$  из генеральной совокупности, имеющей нормальный закон распределения с параметрами `mean=172`, `sd=6.4`. По полученной выборке найти с надёжностью  $p=0.95$  интервальную бутстреп-оценку асимметрии генеральной совокупности. Вычисление интервальной бутстреп-оценки выполнить на основе 10000 вторичных выборок с объёмом 300 элементов каждая. Построить график зависимости этой интервальной бутстреп-оценки от доверительной вероятности.

## Задача 11.

Используя фрейм данных с именем “NEO”, который содержит результаты социологического опроса и личностные психологические показатели студентов, с помощью критерия Шапиро-Уилка проверить следующие статистические гипотезы: 1) переменная `N6_Уязвимость` имеет закон распределения, который статистически значимо не отличается от нормального закона распределения; 2) для респондентов юношей переменная `N6_Уязвимость` имеет закон распределения, который не отличается от нормального закона распределения.

## Задача 12.

Используя фрейм данных с именем “NEO”, с помощью критерия Колмогорова-Смирнова проверить следующие статистические гипотезы: 1) переменная `N6_Уязвимость` имеет нормальный закон распределения с параметрами, которые равны выборочному среднему и выборочному стандартному отклонению; 2) для респондентов юношей переменная `N6_Уязвимость` имеет нормальный закон распределения с параметрами, которые равны выборочному среднему и выборочному стандартному отклонению.

## Задача 13.

Используя фрейм данных с именем “NEO”, с помощью критерия Стьюдента проверить гипотезу о том, что для юношей и девушек математические ожидания переменной N6\_Уязвимость равны.

## Задача 14.

Используя фрейм данных с именем “NEO”, с помощью критерия Уилкоксона проверить следующие статистические гипотезы: 1) уровень переменной N6\_Уязвимость равен 22; 2) уровень переменной N6\_Уязвимость для студентов факультета «А» равен 22.

## Задача 15.

Используя фрейм данных с именем “NEO”, с помощью критерия Уилкоксона проверить следующие статистические гипотезы: 1) для юношей и девушек уровень различий переменной N6\_Уязвимость равен нулю; 2) для респондентов с сильной и слабой степенью религиозности уровень различий переменной N6\_Уязвимость равен нулю.

## Задача 16.

Используя фрейм данных с именем “NEO”, проверить следующие статистические гипотезы: 1) коэффициент корреляции Пирсона переменных С6\_Осмотрительность и N6\_Уязвимость равен нулю; 2) коэффициент корреляции Кендалла переменных С6\_Осмотрительность и N6\_Уязвимость равен нулю; 3) коэффициент корреляции Спирмена переменных С6\_Осмотрительность и N6\_Уязвимость равен нулю.

## Задача 17.

Используя фрейм данных с именем “NEO”, с помощью хи-квадрат критерия Пирсона проверить следующие статистические гипотезы: 1) степень религиозности не зависит от пола респондентов; 2) степень религиозности не зависит от семейного дохода; 3) переменные П\_3 и П\_12 независимы.

## Задача 18.

Используя фрейм данных с именем “NEO”, с помощью однофакторного дисперсионного анализа проверить статистическую гипотезу о том, что математическое ожидание переменной N6\_Уязвимость не зависит от семейного дохода респондентов.

## Задача 19.

Используя фрейм данных с именем “NEO”, с помощью критерия Краскела-Уоллиса проверить статистическую гипотезу о том, что уровень переменной N6\_Уязвимость не зависит от семейного дохода респондентов.

## Задача 20.

Среди случайно взятых 10000 новорождённых оказалось 5143 мальчика. С помощью теста пропорций проверить статистическую гипотезу о том, что вероятность рождения мальчика равна 0.5.



## Задача 21.

Используя фрейм данных с именем “NEO”, с помощью теста пропорций проверить следующие статистические гипотезы: 1) для юношей и девушек вероятности сильной степени религиозности равны; 2) для студентов разных факультетов вероятности сильной степени религиозности равны.

## Задача 22.

Используя фрейм данных с именем “NEO”, с помощью критерия Бартлетта проверить следующие статистические гипотезы: 1) дисперсия переменной N6\_Уязвимость одинаковая для разных факультетов; 2) дисперсия переменной N6\_Уязвимость одинаковая для разных уровней семейного дохода респондентов; 3) дисперсия переменной N6\_Уязвимость одинаковая для разных знаков зодиака.

## Задача 23.

Используя фрейм данных с именем “NEO”, найти выборочные коэффициенты корреляции Пирсона переменных C6\_Осмотрительность и N6\_Уязвимость для девушек и юношей отдельно. На уровне значимости 0.05 проверить являются ли эти коэффициенты корреляции статистически значимыми.

## Задача 24.

Загрузить фрейм данных NEO. На уровне значимости 0.1 найти статистически значимые коэффициенты корреляции Спирмена показателя A6\_Отзывчивость с порядковыми демографическими переменными ВОЗРАСТ, СТЕП\_РЕЛ, СЕМ\_ДОХ, ЛИЧ\_ДОХ.

## Задача 25.

Загрузить фрейм данных NEO. На уровне значимости 0.1 найти статистически значимые коэффициенты корреляции Кендалла переменной СТЕП\_РЕЛ со следующими показателями: N1\_Тревожность, N2\_Враждебность, E1\_Доброжелательность, E2\_Общительность, O1\_Фантазия, O2\_Эстетичность, A1\_Доверие, A2\_Прямота, C1\_Компетентность, C2\_Организованность.

## Задача 26.

Загрузить фрейм данных NEO. Для студентов факультета “Б” с помощью критерия Фишера на уровне значимости 0.05 найти количество статистически значимых зависимостей среди первых тридцати пунктов опросника NEO PI-R. Решить ту же задачу с помощью хи-квадрат критерия Пирсона. Сравнить результаты вычислений.

## Задача 27.

Загрузить фрейм данных NEO, который содержит результаты социологического опроса и личностные психологические показатели студентов. Построить линейную регрессионную модель m1 зависимости психологического показателя N6\_Уязвимость от показателей N1\_Тревожность, A1\_Доверие, O2\_Эстетичность. Найти основные характеристики модели m1 и построить модель m2, удалив из модели m1 независимую переменную A1\_Доверие. Оценить характеристики модели m2 и построить модель m3, удалив из модели m2 независимую переменную O2\_Эстетичность. Сравнить качество моделей m2 и m3. Построить линейную регрессионную модель m4 зависимости

психологического показателя N6\_Уязвимость от всех остальных 29 подшкал теста NEO PI-R. Оценить характеристики модели m4.

Задача 28.

Загрузить фрейм данных NEO. Построить линейную регрессионную модель m21 зависимости показателя N6\_Уязвимость от переменных N2\_Враждебность, O2\_Эстетичность, C2\_Организованность, A2\_Прямота, C5\_Самодисциплина. Оптимизировать модель m21 и построить модель m22, используя пошаговый метод автоматического исключения переменных на основе информационного критерия AIC. Сравнить качество моделей m21 и m22 по скорректированному коэффициенту детерминации.

Задача 29.

Загрузить фрейм данных NEO. Построить оптимальную линейную регрессионную модель m25 для психологического показателя N6\_Уязвимость пошаговым методом добавления независимых переменных, в качестве которых рассматривать все остальные подшкалы теста NEO PI-R. Найти число независимых переменных, включенных автоматически в модель m25.

Задача 30.

Загрузить фрейм данных NEO. Выполнить многофакторный дисперсионный анализ, построив следующие линейные модели: 1) двухфакторную модель m33 зависимости показателя цинизм от степени религиозности и пола респондентов; 2) модель m34 зависимости показателя цинизм от степени религиозности и семейного дохода респондентов с учётом взаимодействия этих факторов; 3) модель m35 зависимости показателя цинизм от степени религиозности, семейного дохода и факультета обучения с учётом всех взаимодействий между этими факторами; 4) модель m36, полученную из модели m35 пошаговым методом автоматического исключения независимых переменных.

## Приложения

### Приложение 1

#### АННОТАЦИЯ ДИСЦИПЛИНЫ

Дисциплина «Статистические пакеты прикладных программ» реализуется на факультете информационных систем и безопасности кафедрой фундаментальной и прикладной математики.

Цель дисциплины: формирование у студентов адекватных представлений о современных программных средствах прикладной статистики для их эффективного практического применения в различных областях научных исследований и инженерной практики.

Задачи: познакомить студентов с профессиональными статистическими пакетами (R, SPSS, Statistica и другими) для разработки алгоритмов и инструментальных средств обработки данных, а также для создания пользовательских интерфейсов к вычислительным процедурам; обсудить примеры применения этих пакетов для решения прикладных задач.

Дисциплина направлена на формирование следующих компетенций:

- ПК-1 - способность использовать стандартные пакеты прикладных программ для решения практических задач на электронных вычислительных машинах, отлаживать, тестировать прикладное программное обеспечение.

В результате освоения дисциплины обучающийся должен:

*Знать:* основные современные статистические пакеты прикладных программ и технологии их использования

*Уметь:* применять статистические пакеты для решения прикладных задач, разрабатывать алгоритмы и статистические инструментальные средства для анализа и обработки данных, а также создавать интерфейсы к вычислительным процедурам

*Владеть:* навыками работы с программными системами для статистического анализа данных и разработки новых инструментальных средств при решении прикладных задач.

По дисциплине предусмотрена промежуточная аттестация в форме зачета с оценкой.

Общая трудоемкость освоения дисциплины составляет 3 зачетные единицы.

**ЛИСТ ИЗМЕНЕНИЙ**

№	Текст актуализации или прилагаемый к РПД документ, содержащий изменения	Дата	№ протокола
1	Приложение к листу изменений №1	20.06.18	13
2	Приложение к листу изменений №2	28.06.19	13
3	Приложение к листу изменений №3	22.06.20	13

**1. Перечень программного обеспечения (ПО) (к п.7 на 2018г.)***Таблица 1*

№п/п	Наименование ПО	Производитель	Способ распространения
1	Microsoft Office 2010 Pro	Microsoft	лицензионное
2	Windows XP или Windows 7	Microsoft	лицензионное
3	Язык программирования R	сообщество	свободно распространяемое
4	Язык программирования Python	сообщество	свободно распространяемое
5	Matlab	The MathWorks	лицензионное
6	Mathcad Education - University edition	PTC	лицензионное
7	Kaspersky Endpoint Security	Kaspersky	лицензионное

Для лабораторных занятий можно также использовать актуальные полнофункциональные демонстрационные версии профессиональных статистических пакетов SPSS и Statistica.

**2. Перечень современных профессиональных баз данных (БД) и информационно-справочных систем (ИСС) (к п.6.2 на 2018г.)***Таблица 2*

№ п/п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2018 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2018 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis Электронные издания издательства Springer
3	Компьютерные справочные правовые системы Консультант Плюс, Гарант

**3. Структура дисциплины (п.2 для студентов приема 2018г.)**

Общая трудоёмкость дисциплины составляет 4 з.е., 144 ч., в том числе контактная работа обучающихся с преподавателем 56 ч., самостоятельная работа обучающихся 88 ч.

*Таблица 3*

№ п/п	Раздел дисциплины/темы	Семестр	Виды учебной работы (в часах)				Формы текущего контроля успеваемости, форма промежуточной аттестации (по семестрам)
			контактная		Промежуточная аттестация	Самостоятельная работа	
			Лекции	Практические занятия			
1	Обзор современных программных средств статистического анализа данных	7	4	4		12	Опрос
2	Пакет Statistica	7	4	4		12	Опрос
3	Пакет SPSS	7	4	6		14	Опрос Тестирование № 1
4	Вычислительная среда и язык программирования R	7	4	6		14	Опрос Тестирование № 2
5	Статистические средства универсальных математических пакетов	7	4	6		14	Опрос Тестирование № 3
6	Статистические средства офисных пакетов	7	4	4		14	Опрос Контрольная работа
	Зачёт с оценкой	7		2		8	Ответы на теоретические вопросы, итоговая контрольная работа
	<b>Итого:</b>		<b>24</b>	<b>32</b>		<b>88</b>	

4. Для студентов приема 2018г. в разделах рабочей программы дисциплины:

#### ОГЛАВЛЕНИЕ

4. Образовательные технологии

5. Оценка планируемых результатов обучения

7. Материально-техническое обеспечение дисциплины

9.1. Планы лабораторных занятий

«Лабораторные занятия» заменить на «Практические занятия», «Лабораторная работа» заменить на «Практическая работа», «лаборатория» на «лаборатория или компьютерный класс».

**1. Перечень программного обеспечения (ПО) (к п.7 на 2019г.)***Таблица 1*

№п/п	Наименование ПО	Производитель	Способ распространения
1	Microsoft Office 2010 Pro	Microsoft	лицензионное
2	Windows XP или Windows 7	Microsoft	лицензионное
3	Пакет Statistica	проприетарное	лицензионное
4	Язык программирования R	сообщество	свободно распространяемое
5	Язык программирования Python	сообщество	свободно распространяемое
6	Matlab	The MathWorks	лицензионное
7	Mathcad Education - University edition	PTC	лицензионное
8	Kaspersky Endpoint Security	Kaspersky	лицензионное
9	Zoom	проприетарное	лицензионное

Для лабораторных занятий можно также использовать актуальную полнофункциональную демонстрационную версию профессионального статистического пакета SPSS.

**2. Перечень современных профессиональных баз данных (БД) и информационно-справочных систем (ИСС) (к п.6.2 на 2019г.)***Таблица 2*

№ п/п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2019 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2019 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Компьютерные справочные правовые системы Консультант Плюс, Гарант

**3. Список источников и литературы (п.6.1)****Литература***Основная*

1. Наследов А. SPSS 19: профессиональный статистический анализ данных. - СПб.: Питер, 2011. - 400 с.
2. Основы эконометрики в пакете STATISTICA: Учебное пособие / Плохотников К.Э. - М.:Вузовский учебник, 2018. - 298 с. (Переплёт) ISBN 978-5-9558-0114-8 - Режим доступа: <http://znanium.com/catalog/product/914118>

*Дополнительная*

1. Маркетинговые исследования с SPSS: Учебное пособие / Г. Моосмюллер, Н.Н. Ребик. - 2-е изд. - М.: ИНФРА-М, 2011. - 200 с.: 60х90 1/16. - (Высшее образование). (обложка) ISBN 978-5-16-004240-4 - Режим доступа: <http://znanium.com/catalog/product/257371>
2. Многомерные статистические методы в экономике : учебник / Л.И. Ниворожкина, С.В. Арженовский. — М. : РИОР : ИНФРА-М, 2018. — 203 с. — (Высшее образование). — [www.dx.doi.org/10.12737/21773](http://www.dx.doi.org/10.12737/21773). - Режим доступа: <http://znanium.com/catalog/product/975772>
3. Статистические методы обработки экспериментальных данных с использованием пакета MathCad: Учебное пособие/Ф.И.Карманов, В.А.Острейковский - М.: КУРС, НИЦ ИНФРА-М, 2015. - 208 с.: 60х90 1/16 (Обложка) ISBN 978-5-905554-96-4 - Режим доступа: <http://znanium.com/catalog/product/508241>
4. Статистическая обработка данных в учебно-исследовательских работах: Учебное пособие / Волкова П.А., Шипунов А.Б. - М.: Форум, 2016. - 96 с.: 60х90 1/16 (Обложка. КБС) ISBN 978-5-91134-576-1 - Режим доступа: <http://znanium.com/catalog/product/556479>
5. Тюрин Ю. Н. Анализ данных на компьютере: учеб. пособие по направлениям "Математика", "Математика. Прикладная математика" / Ю. Н. Тюрин, А. А. Макаров. - Изд. 4-е, перераб. - М.: Форум, 2013. - 366 с.- (Высшее образование)



## 1. Образовательные технологии (к п.4 на 2020г.)

В период временного приостановления посещения обучающимися помещений и территории РГГУ для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:

- видео-лекции;
- онлайн-лекции в режиме реального времени;
- электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
- системы для электронного тестирования;
- консультации с использованием телекоммуникационных средств.

## 2. Список источников и литературы (п.6.1)

### Литература

#### Основная

1. Наследов А. SPSS 19: профессиональный статистический анализ данных. - СПб.: Питер, 2011. - 400 с.
2. Основы эконометрики в пакете STATISTICA: Учебное пособие / Плохотников К.Э. - М.: Вузовский учебник, 2018. - 298 с. (Переплёт) ISBN 978-5-9558-0114-8 - Режим доступа: <http://znanium.com/catalog/product/914118>

#### Дополнительная

1. Маркетинговые исследования с SPSS: Учебное пособие / Г. Моосмюллер, Н.Н. Ребик. - 2-е изд. - М.: ИНФРА-М, 2011. - 200 с.: 60x90 1/16. - (Высшее образование). (обложка) ISBN 978-5-16-004240-4 - Режим доступа: <http://znanium.com/catalog/product/257371>
2. Многомерные статистические методы в экономике : учебник / Л.И. Ниворожкина, С.В. Арженовский. — М. : РИОР : ИНФРА-М, 2018. — 203 с. — (Высшее образование). — [www.dx.doi.org/10.12737/21773](http://www.dx.doi.org/10.12737/21773). - Режим доступа: <http://znanium.com/catalog/product/975772>
3. Статистические методы обработки экспериментальных данных с использованием пакета MathCad: Учебное пособие/Ф.И.Карманов, В.А.Острейковский - М.: КУРС, НИЦ ИНФРА-М, 2015. - 208 с.: 60x90 1/16 (Обложка) ISBN 978-5-905554-96-4 - Режим доступа: <http://znanium.com/catalog/product/508241>
4. Статистическая обработка данных в учебно-исследовательских работах: Учебное пособие / Волкова П.А., Шипунов А.Б. - М.: Форум, 2016. - 96 с.: 60x90 1/16 (Обложка. КБС) ISBN 978-5-91134-576-1 - Режим доступа: <http://znanium.com/catalog/product/556479>
5. Тюрин Ю. Н. Анализ данных на компьютере: учеб. пособие по направлениям "Математика", "Математика. Прикладная математика" / Ю. Н. Тюрин, А. А. Макаров. - Изд. 4-е, перераб. - М.: Форум, 2013. - 366 с.- (Высшее образование)
6. Груздев, А. В. Прогнозное моделирование в IBM SPSS Statistics, R и Python: метод деревьев решений и случайный лес : руководство / А. В. Груздев. — Москва : ДМК Пресс, 2018. — 642 с. — ISBN 978-5-97060-539-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/123700>

### 3. Перечень современных профессиональных баз данных (БД) и информационно-справочных систем (ИСС) (к п. 6.2 на 2020г.)

Таблица 1

№ п/п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2020 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2020 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Компьютерные справочные правовые системы Консультант Плюс, Гарант

### 4. Перечень программного обеспечения (ПО) (к п.7 на 2020г.)

Таблица 2

№п/ п	Наименование ПО	Производитель	Способ распространения
1	Microsoft Office 2010 Pro	Microsoft	лицензионное
2	Windows XP или Windows 7	Microsoft	лицензионное
3	Пакет Statistica	проприетарное	лицензионное
4	Язык программирования R	сообщество	свободно распространяемое
5	Язык программирования Python	сообщество	свободно распространяемое
6	Matlab	The MathWorks	лицензионное
7	Mathcad Education - University edition	PTC	лицензионное
8	Kaspersky Endpoint Security	Kaspersky	лицензионное
9	Zoom	Zoom	лицензионное

Для лабораторных занятий можно также использовать актуальную полнофункциональную демонстрационную версию профессионального статистического пакета SPSS.