

**«Финансовый университет при Правительстве Российской Федерации»  
(Финуниверситет)**

**Калужский филиал Финуниверситета**

**Кафедра «Бизнес – информатика и высшая математика»**

**«УТВЕРЖДАЮ»**

Директор Калужского филиала  
Финуниверситета



**В.А. Матчинов**

**30» июня 2022 г.**

**Никаноркина Н.В.**

## **АНАЛИЗ ДАННЫХ**

### **Рабочая программа дисциплины**

для студентов, обучающихся по направлению подготовки  
38.03.01 «Экономика»

Образовательная программа «Экономика и финансы»  
Очная и очно-заочная формы обучения

*Рекомендовано Ученым советом Калужского филиала Финуниверситета  
(протокол №56 от 30.06. 2022 г.)*

Одобрено кафедрой «Бизнес – информатика и высшая математика»  
Калужского филиала Финуниверситета  
(протокол № 12 от 28 июня 2022 г.)


**КАЛУГА 2022**


Рабочая программа предназначена для преподавания дисциплины «Анализ данных» студентам, обучающимся по направлению подготовки 38.03.01 «Экономика», образовательная программа «Экономика и финансы» по очной и очно-заочной формам обучения.

В рабочей программе излагаются планируемые результаты освоения дисциплины, содержание дисциплины, тематика и содержание семинаров и практических занятий, технологии их проведения. В рабочей программе дисциплины приводится перечень учебно-методического обеспечения для самостоятельной работы обучающихся, фонд оценочных средств для проведения промежуточной аттестации обучающихся, перечень основной и дополнительной литературы, а также ресурсов информационно-телекоммуникационной сети Интернет.

СОГЛАСОВАНО:

Заместитель директора  
по учебно-методической работе  /Орловцева О.М./  
«28» июня 2022 г.

Начальник учебно-методического отдела  /Толстикова В.С./  
«28» июня 2022 г.

Заведующий кафедрой  
«Бизнес-информатика и высшая математика»  /Дробышева И.В./  
«28» июня 2022 г.

## СОДЕРЖАНИЕ

1. Наименование дисциплины.....	3
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине .....	3
3. Место дисциплины в структуре образовательных программ .....	5
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся .....	5
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий .....	6
5.1. Содержание дисциплины.....	6
5.2. Учебно-тематический план .....	11
5.3. Содержание семинаров, практических занятий .....	12
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине .....	14
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы .....	14
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю .....	15
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по данной дисциплине.....	19
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины .....	33
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины .....	34
10. Методические указания для обучающихся по освоению дисциплины .....	36
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем .....	37
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	38

## 1. Наименование дисциплины

Б.1.1.2.2. «Анализ данных».

## 2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Таблица 1

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
ПКН-1	Владение основными научными понятиями и категориальным аппаратом современной экономики и их применение при решении прикладных задач	1. Демонстрирует знание современных экономических концепций, моделей, ведущих школ и направлений развития экономической науки, использует категориальный и научный аппарат при анализе экономических явлений и процессов.	<b>Знает:</b> концептуальные постановки прикладных экономических задач, решаемых с применением методов математической статистики; <b>Умеет:</b> выбирать адекватные методы анализа экономических явлений и процессов
		2. Выявляет сущность и особенности современных экономических процессов, их связь с другими процессами, происходящими в обществе, критически переосмысливает текущие социально-экономические проблемы.	<b>Знает:</b> основные понятия и методы теории вероятностей и прикладной статистики, необходимые для анализа социально-экономических задач и процессов; <b>Умеет:</b> решать социально-экономические задачи вероятностными и статистическими методами
		3. Грамотно и результативно пользуется российскими и зарубежными источниками научных знаний и экономической информации, знает основные направления экономической политики государства.	<b>Знает:</b> источники актуальных данных о состоянии экономических субъектов; <b>Умеет:</b> осуществлять поиск и импорт данных, необходимых для решения конкретных задач

ПКН-3	Способность осуществлять сбор, обработку и статистический анализ данных, применять математические методы для решения стандартных профессиональных финансово-экономических задач, интерпретировать полученные результаты	1.Проводит сбор, обработку и статистический анализ данных для решения финансово-экономических задач.	<b>Знает:</b> тенденции развития новых программных продуктов и IT-сервисов, ориентированных на повышение эффективности сбора и обработки данных; <b>Умеет:</b> осуществлять поиск, сбор и первичную подготовку данных для экономических исследований
		2.Формулирует математические постановки финансово-экономических задач, переходит от экономических постановок задач к математическим моделям.	<b>Знает:</b> основные принципы спецификации экономико-математических моделей <b>Умеет</b> формализовать постановки прикладных задач из сфер экономики и финансов
		3.Системно подходит к выбору математических методов и информационных технологий для решения конкретных финансово-экономических задач в профессиональной области.	<b>Знает:</b> основные понятия теории вероятностей и математической статистики; методики расчетов, используемые при анализе данных; вероятностные и статистические методы. <b>Умеет:</b> использовать инструменты описательной статистики и визуализации данных, вероятностные и статистические методы для решения профессиональных задач.
		4.Анализирует результаты исследования математических моделей финансово-экономических задач и делает на их основании количественные и качественные выводы и рекомендации по принятию финансово-экономических решений.	<b>Знает:</b> основные принципы программирования алгоритмов статистической обработки данных в табличных процессорах и R(RStudio) для анализа и прогноза внутри- и внешнеэкономических процессов <b>Умеет:</b> применять технические и инструментальные средства для решения прикладных экономических задач и подготовки обоснований управленческих решений

ПКН-4	Способность оценивать показатели деятельности экономических субъектов	1.Проводит анализ внешней и внутренней среды ведения бизнеса, выявляет основные факторы экономического роста, оценивает эффективность формирования и использования производственного потенциала экономических субъектов.	<b>Знает:</b> методы анализа тенденций развития экономических процессов; <b>Умеет:</b> использовать табличные процессоры и R(RStudio) в статистическом анализе данных, характеризующих внутри- и внешнеэкономические процессы
		2. Рассчитывает и интерпретирует показатели деятельности экономических субъектов.	<b>Знает:</b> условия и ограничения основных моделей в статистическом анализе <b>Умеет:</b> сопоставлять реальные данные с математическими моделями и уметь интерпретировать результаты статистических исследований

### 3. Место дисциплины в структуре образовательной программы

Дисциплина «Анализ данных» является обязательной дисциплиной Цикла математики и информатики по направлению подготовки 38.03.01 «Экономика» ОП "Экономика и финансы".

### 4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Таблица 2

Вид учебной работы по дисциплине	Всего (в з./е. и часах)	Семестр 3 (в часах)	Семестр 4 (в часах)
Общая трудоемкость дисциплины	7 з./е., 252 ч.	134/110	118/142
<b>Контактная работа - Аудиторные занятия</b>	116/84	66/34	50/50
Лекции	32/32	16/16	16/16
Семинары, практические занятия	84/52	50/18	34/34
<b>Самостоятельная работа</b>	136/168	68/76	68/92
Вид текущего контроля	Домашнее творческое задание, расчетно-аналитическая работа	Домашнее творческое задание	Расчетно-аналитическая работа
Вид промежуточной аттестации	Зачет, экзамен	Зачет	Экзамен

## 5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

### 5.1. Содержание дисциплины

#### *Тема 1. Данные в экономике, их визуализация и предварительная обработка*

**1.1. Данные в экономике.** Объекты, признаки и таблицы. Типы признаков в экономике: интервальные, порядковые, ранговые, дихотомические. Форматирование наборов данных как таблиц в табличных процессорах. Гистограммы в табличных процессорах. Условное форматирование в табличных процессорах. Графики и диаграммы рассеяния в табличных процессорах.

**1.2. Инструменты описательной статистики в табличных процессорах.** Измерение центра распределения. Измерение разброса данных. Описательная статистика и диаграммы размаха в табличных процессорах.

**1.3. Предварительная обработка данных.** Выбросы и их обработка в табличных процессорах. Пропущенные значения и их обработка в табличных процессорах. Повторяющиеся строки и их обработка в табличных процессорах. Синтетические признаки.

#### *Тема 2. Случайные события*

**2.1. Определение вероятности. Случайные события, их виды. Операции**

над событиями как операции над множествами. Комбинаторика. Классическая вероятностная схема. Схема геометрических вероятностей. Статистическая вероятность. Аксиомы теории вероятностей. Метод Монте-Карло. Моделирование случайных событий.

**2.2. Основные теоремы теории вероятностей.** Теорема сложения вероятностей. Условная вероятность. Независимость событий. Формула полной вероятности. Формула Байеса.

### ***Тема 3. Случайные величины***

**3.1. Определение случайной величины.** Понятие случайной величины. Функция распределения случайной величины. Свойства функции распределения. Индикатор события как простейшая случайная величина. Функция распределения индикатора события. Смеси распределений.

**3.2. Дискретные случайные величины и их важнейшие числовые характеристики.** Дискретная случайная величина. Ряд распределения и функция распределения дискретной случайной величины. Математическое ожидание, дисперсия и среднее квадратическое отклонение дискретной случайной величины.

**3.3. Дискретные случайные величины, часто встречающиеся в практике.** Последовательности испытаний. Биномиальная схема. Отрицательное биномиальное распределение (биномиальный закон распределения, геометрические законы распределения). Закон распределения Пуассона. Простейший поток событий. Гипергеометрический закон распределения.

**3.4. Абсолютно непрерывные случайные величины и их важнейшие числовые характеристики.** Абсолютно непрерывная случайная величина. Функция распределения и функция плотности распределения абсолютно непрерывной случайной величины. Свойства функции плотности распределения. Математическое ожидание, дисперсия и среднее квадратическое отклонение абсолютно непрерывной случайной величины.

**3.5. Абсолютно непрерывные случайные величины, часто встречающиеся в практике.** Равномерный закон распределения. Экспоненциальный закон



распределения. Нормальный закон распределения. Законы распределения, важные в математической статистике (законы распределения Стюдента,  $\chi^2$ , Фишера — Снедекора). Реализация моделей абсолютно непрерывных случайных величин в табличных процессорах.

**3.6. Моменты и критические границы случайной величины.** Начальные и центральные моменты случайной величины. Асимметрия и эксцесс случайной величины. Квантили и процентные точки случайной величины. Вычисление квантилей и процентных точек в табличных процессорах. Медиана и мода случайной величины.

**3.7. Системы случайных величин.** Случайные векторы и условные законы распределения. Условный ряд распределения (для дискретных случайных величин), условная плотность распределения (для непрерывных случайных величин). Условное математическое ожидание. Формула полного математического ожидания. Формула полной дисперсии. Ковариация и коэффициент корреляции.

#### ***Тема 4. Предельные теоремы теории вероятностей***

**4.1. Закон больших чисел.** Массовые случайные явления в экономике. Теорема Чебышёва и оценка математического ожидания. Теорема Бернулли и оценка вероятности. Обсуждение условий статистической устойчивости.

**4.2. Центральная предельная теорема.** Теорема Леви. Интегральная теорема Муавра — Лапласа. Место центральной предельной теоремы в изучении статистических закономерностей.

#### ***Тема 5. Выборочный метод математической статистики***

**5.1. Основы выборочного метода.** Предмет и задачи математической статистики. Генеральная и выборочная совокупности. Случайная и конкретная выборки. Случайная повторная и случайная бесповторная выборки. Соотношение между предельной ошибкой среднего, уровнем значимости (риском) и объемом выборки. Использование этого соотношения в организации выборочных обследований.

**5.2. Оценка плотности распределения и функции распределения.**

Вариационный ряд. Числовые характеристики выборки. Интервальный вариационный ряд. Полигон частот, кумулята. Оценка числовых характеристик генеральной случайной величины (выборочное среднее как оценка математического ожидания, относительная частота как оценка вероятности, выборочная дисперсия как оценка дисперсии).

**5.3. Точечные оценки параметров.** Понятие точечной оценки параметра генеральной совокупности. Свойства точечных оценок: состоятельность, несмещенность, эффективность. Выборочное среднее как состоятельная, несмещенная и эффективная оценка математического ожидания генеральной случайной величины. Смещенность выборочной дисперсии как оценки дисперсии генеральной случайной величины. Исправленная выборочная дисперсия как несмещенная и состоятельная оценка дисперсии генеральной случайной величины (для повторной и бесповторной выборок). Методы построения точечных оценок: метод моментов, метод максимального правдоподобия. Построение оценок параметров распределений случайных величин.

**5.4. Интервальные оценки параметров.** Понятие интервальной оценки параметра генеральной совокупности. Точные интервальные оценки вероятности, математического ожидания, дисперсии и коэффициента корреляции. Поправка на конечный объем генеральной совокупности (для повторной и бесповторной выборок). Асимптотический подход к интервальному оцениванию.

## ***Тема 6. Проверка статистических гипотез***

**6.1. Статистические гипотезы.** Понятие статистической гипотезы. Виды статистических гипотез: параметрические и непараметрические, простые и сложные. Критерий проверки гипотезы, критическое множество. Проверка гипотез с помощью интервальных оценок. Ошибки первого и второго рода. Мощность критерия. Наиболее мощный критерий. Достигаемый уровень значимости (p-value).

**6.2. Критерии случайности, независимости, однородности.** Критерий Граббса (отсутствия аномальных значений). Таблицы сопряженности. Критерии однородности  $\chi^2$  и Смирнова. Критерий независимости  $\chi^2$ , критерий значимости

коэффициента корреляции. Критерий инверсий.

**6.3. Критерии согласия.** Критерий согласия  $\chi^2$  Пирсона. Критерий  $\chi^2$  Пирсона для дискретных распределений. Критерии согласия Колмогорова и Смирнова.

**6.4. Критерии согласия для сложных гипотез.** Критерии для проверки нормальности (Лиллиефорса, Шапиро-Уилка и другие)

**6.5. Параметрические критерии.** Проверка гипотезы о равенстве математического ожидания теоретическому значению. Проверка гипотезы о равенстве двух математических ожиданий. Проверка гипотезы о равенстве дисперсии теоретическому значению. Проверка гипотезы о равенстве двух дисперсий. Проверка гипотезы о равенстве вероятности события теоретическому значению. Проверка гипотезы о равенстве двух вероятностей. Проверка гипотез для связанных выборок.

## ***Тема 7. Основы дисперсионного анализа***

**7.1. Задачи дисперсионного анализа.** Результативный признак и факторы. Существенность факторов. Предпосылки и границы применения дисперсионного анализа.

**7.2. Однофакторный дисперсионный анализ.** Задача однофакторного дисперсионного анализа. Разложение общей вариации результативного признака на межгрупповую и внутригрупповую. Реализация процедуры однофакторного дисперсионного анализа в табличных процессорах. Интерпретация результатов.

**7.3. Двухфакторный дисперсионный анализ.** Задача двухфакторного дисперсионного анализа. Двухфакторный анализ с повторениями и без. Межфакторное взаимодействие и отражение его в модели. Реализация процедуры двухфакторного дисперсионного анализа в табличных процессорах.

## ***Тема 8. Основы непараметрической статистики***

**8.1. Таблицы сопряженности.** Критерий  $\chi^2$  для проверки независимости компонент случайной величины.

**8.2. Непараметрические критерии.** Критерий знаков. Распределение Вилкоксона и его критические границы. Критерий Вилкоксона — Манна — Уитни

(непараметрический критерий сравнения математических ожиданий для независимых выборок).

**8.3. Ранговая корреляция.** Коэффициент ранговой корреляции Спирмена. Коэффициент ранговой корреляции Кендалла. Коэффициент конкордации. Проверка гипотез о значимости ранговых коэффициентов корреляции.

## 5.2. Учебно-тематический план

Таблица 3

п/ п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успева- емости
		Всего	Контактная работа – Аудиторная работа			Самостоя- тельная работа	
			Общая, в т.ч.:	Лекции	Семина- ры, прак- тические занятия		
1.	Данные в эконо- мике, их ви- зуализация и предваритель- ная обработка	32/32	15/11	4/4	11/7	17/21	Решение задач на практических заня- тиях. Домашние задания. Собеседования по материалу.
2.	Случайные со- бытия	32/32	15/11	4/4	11/7	17/21	Решение задач на практических заня- тиях. Домашние задания. Собеседования по материалу.
3.	Случайные ве- личины	32/32	15/11	4/4	11/7	17/21	Решение задач на практических заня- тиях. Домашние задания. Собеседования по материалу.
4.	Предельные теоремы теории вероятностей	32/32	15/11	4/4	11/7	17/21	Решение задач на практических заня- тиях. Домашние задания. Собеседования по материалу.
5.	Выборочный метод матема- тической стати- стики	31/31	14/10	4/4	10/6	17/21	Решение задач на практических заня- тиях. Домашние задания. Собеседования по материалу.
6.	Проверка ста- тистических гипотез	31/31	14/10	4/4	10/6	17/21	Решение задач на практических заня- тиях. Домашние задания. Собеседования по материалу.
7.	Основы дис- персионного	31/31	14/10	4/4	10/6	17/21	Решение задач на практических заня-

	анализа						тиях. Домашние задания. Собеседования по материалу.
8.	Основы непараметрической статистики	31/31	14/10	4/4	10/6	17/21	Решение задач на практических занятиях. Домашние задания. Собеседования по материалу
В целом по дисциплине							Согласно учебному плану: домашнее творческое задание, расчетно-аналитическая работа
Итого		252/252	116/84	32/32	84/52	136/168	

### 5.3. Содержание семинаров, практических занятий

Таблица 4

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Данные в экономике, их визуализация и предварительная обработка	Представление данных в виде форматированных таблиц. Гистограммы, диаграммы рассеяния, диаграммы размаха. Сводные таблицы. Визуализация качественных признаков. Импорт, предварительная обработка и визуализация данных в R.[8.1], [9.2]	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы
Случайные события	Основы комбинаторики. Функции подсчета количества комбинаций в табличных процессорах и R. Классическая вероятностная схема. Схема геометрических вероятностей. Моделирование статистической вероятности. Вероятность суммы. Условные вероятности. Вероятность произведения. Полная вероятность. Формула Байеса. Последовательности испытаний. Схема Бернулли. [8.1,4],[9.2]	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы
Случайные величины	Дискретные случайные величины и их важнейшие числовые характеристики. Дискретные случайные величины, встречающиеся в экономической практике. Абсолютно непрерывные случайные величины и их важнейшие числовые характеристики. Абсолютно непрерывные величины, встречающиеся в экономической практике. Равномерный, показательный, нормальный законы распределения. Моменты случайных величин. Квантили и 100 $\alpha$ -процентные точки. Формулы полного математического ожидания и полной дисперсии. Формулы корреляции и	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы

	ковариации. [8.1,2],[9.2]	
Предельные теоремы теории вероятностей	Применение неравенств Маркова и Чебышёва для решения практических задач. Центральная предельная теорема. [8.6], [9.2]	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы
Выборочный метод математической статистики	Выборочные оценки функций плотности и распределения. Точечные оценки параметров. Описательная статистика средствами табличных процессоров и R. Интервальные оценки параметров. [8.2-5],[9.2]	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы
Проверка статистических гипотез	Формулировка статистически гипотез. Критерии проверки статистических гипотез. Параметрические гипотезы. Одно-и двухвыборочные тесты. Функции проверки статистических гипотез в табличных процессорах и R. Критерии согласия. Проверка значимости коэффициента корреляции Пирсона. Непараметрическая точечная оценка математического ожидания. Непараметрическая интервальная оценка математического ожидания. Критерий Вилкоксона (парный критерий знаковых рангов). Критерий Вилкоксона —Манна —Уитни (непараметрический критерий сравнения математических ожиданий для независимых выборок). Коэффициент ранговой корреляции Спирмена.[8.2-5], [9.2]	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы
Основы дисперсионного анализа	Реализация процедуры однофакторного дисперсионного анализа в табличных процессорах. Интерпретация результатов. Двухфакторный анализ с повторениями и без. Межфакторное взаимодействие и отражение его в модели. Реализация процедуры двухфакторного дисперсионного анализа в табличных процессорах. [8.2-5], [9.2]	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы
Основы непараметрической статистики	Непараметрическая точечная оценка математического ожидания. Непараметрическая интервальная оценка математического ожидания. Критерий Вилкоксона (парный критерий знаковых рангов). Критерий Вилкоксона —Манна —Уитни (непараметрический критерий сравнения математических ожиданий для независимых выборок). Коэффициент ранговой корреляции Спирмена. Коэффициент ранговой корреляции Кендалла. Проверка гипотез о значимости ранговых коэффициентов корреляции. Примеры использования ранговой корреляции в экономических исследования [8.2-5], [9.2]	Решение задач в аудитории, ответы «у доски», проверка самостоятельной работы

## 6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

### 6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Таблица 5

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Данные в экономике, их визуализация и предварительная обработка	Выбросы и их обработка в табличных процессорах. Пропущенные значения и их обработка в табличных процессорах. Повторяющиеся строки и их обработка в табличных процессорах. Синтетические признаки	Выполнение домашнего задания по сбору и предварительной обработке данных с применением сервисов создания опросов, табличных процессоров, R
Случайные события	Реализация сочетаний и размещений. Алгебра событий. Диаграммы Эйлера-Венна	Выполнение домашних работ по реализации комбинаций в R, имитации случайных событий в R. Диаграммы Эйлера-Венна в R
Случайные величины	Условные распределения. Полное математическое ожидание и полная дисперсия. Многомерные случайные величины и условные распределения в экономике. Смеси распределений.	Решение задач на смеси распределений.
Предельные теоремы теории вероятностей	Метод Монте-Карло. Моделирование случайных величин в табличных процессорах и R.	Выполнение домашних работ по моделированию случайных величин, распределенных по основным законам в R. Вычисление характеристик случайных величин в процессе моделирования.
Выборочный метод математической статистики	Практические задачи оценки параметров генеральной совокупности	Выполнение самостоятельной работы по теме на портале campus.fa.ru
Проверка статистических гипотез	Критерии согласия	Выполнение самостоятельной работы по теме на портале campus.fa.ru
Основы дисперсионного анализа	Функции языка R для проведения дисперсионного анализа	Решение задач дисперсионного анализа в среде R
Основы непараметрической статистики	Задачи непараметрической статистики в экономической практике	Решение задач на применение критериев: Уилкоксона, знаков, Манна-Уитни в R.



## **6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю**

### **Примерная тематика домашнего творческого задания**

Проведите свое небольшое социологическое исследование по одной из следующих тем:

1. Здоровье
2. Досуг
3. Планы дальнейшего обучения или трудоустройства
4. Природа и экология
5. Транспорт
6. Иностранные языки
7. Благосостояние населения
8. Качество и доступность услуг
9. Технологии и интернет
10. Психология и общение

1. Для этого составьте небольшую анкету (3-4 вопроса), ответы на которую помогут Вам ответить на вопросы исследования. Вопросы должны содержать как категориальные ответы, так и числовые.

2. Проверьте однозначность вопросов и корректность предлагаемых ответов.

3. Запрограммируйте свою анкету в каком-либо сервисе создания опросов.

4. Запустите интернет-анкетирование среди респондентов (это могут быть друзья, студенты, случайные интернет-пользователи, люди с форумов или иных групп). Соберите не менее 50, а лучше 100 ответов респондентов.

5. Скачайте файл ответов и визуализируйте результаты опроса с помощью табличного процессора или R(RStudio) на Ваше усмотрение. При этом для категориальных переменных должны быть построены гистограммы, а для числовых переменных – гистограммы и ящики с усами, проинтерпретируйте результаты.

Оформите и распечатайте результаты в каком-либо текстовом редакторе с краткими выводами Вашего исследования.

## **Примерное задание расчётно-аналитической работы**

### **(звездочкой обозначены задания на дополнительную оценку)**

1. В соответствии с вариантом, выданным преподавателем, соберите дневные данные о ценах закрытия и объемах торгов по акциям трех компаний, обращающимся на Московской бирже, с начальной даты, указанной в задании, по конечную дату, указанную в задании.

Дальнейшая обработка разбивается на две части:

- Обработка совместных данных для всех трёх компаний (диаграммы рассеяния в п.3.5\*, корреляционный анализ, двух выборочные тесты). Для этой части создается отдельный лист, на котором совместно анализируются данные всех компаний

- Обработка данных отдельно для каждой компании. Для этой части создаются три отдельных листа, каждый из которых соответствует одной компании.

2. Удалите строки, соответствующие датам, в которые не было торгов хотя бы у одного из трех эмитентов (пропущенные, или нулевые значения объемов торгов)

Тренировочное задание:

3. Рассчитайте для каждой компании дневные логдоходности.

3.1 Для каждой компании посчитайте количество строк, в которых значение логдоходности оказалось строго равно 0.

3.2. Скопируйте данные по каждой компании на три отдельных листа (один лист – одна компания) и удалите строки с нулями логдоходности.

3.3 \*. Постройте для каждой компании графики цен от времени. Прослеживаются ли какие-нибудь зависимости цен от времени?

3.4 \*. Для одной из компаний постройте по три диаграммы рассеяния для каждой пары признаков. Между какими признаками можно предположить наличие зависимости?

3.5 \*. Постройте три диаграммы рассеяния логдоходности для каждой пары компаний.

3.6 \*. Рассчитайте корреляционную матрицу для тройки компаний.

3.7\*. Проверьте гипотезы о незначимости коэффициентов корреляции для тройки компаний

3.8\*. Постройте диаграммы размаха логдоходности для каждой компании.

3.9\*. Для тройки компаний постройте диаграммы оптимальности по Парето и определите акции, оптимальные по Парето

4. Постройте гистограммы интервальных частот цен и логдоходностей и опишите их: являются ли они симметричными или скошенными, похожи ли они на «колокол» нормального распределения.

4.1\*. Укажите на графиках асимметрию и эксцесс, полученные в п. 5, а также результаты проверки гипотез о нормальности распределений, полученные в п. 10.

5. С помощью программы «Описательная статистика» определите основные статистические показатели для логдоходностей.

6. Скопируйте данные для каждой отдельной компании на отдельный лист (всего три дополнительных листа) и удалите строки с выбросами логдоходности.

7. Постройте гистограммы интервальных частот логдоходности без выбросов.

7.1\*. Постройте графики эмпирической плотности для логдоходности в R с выбросами и без выбросов

7.2\*. Постройте графики выборочных функций распределения для логдоходности с выбросами и без выбросов.

7.3\*\*. В предположении нормального закона логдоходности для данных с не удалёнными выбросами замените параметры нормального закона  $\mu$  и  $\sigma$  их несмещёнными состоятельными оценками и рассчитайте значения функции плотности нормального закона и функции распределения в правых концах интервалов. Для каждого из анализируемых признаков постройте на одном рисунке гистограмму частот и график теоретической функции плотности распределения, на другом эмпирическую и теоретическую функцию распределения.

7.4\*\*. Выполните то же, что и в п. 7.3\*\* для данных с удалёнными выбросами.

8. Постройте 95%-ные доверительные интервалы для математических ожиданий и средних квадратических отклонений логдоходностей с выбросами.

8.1\*. То же для данных с удалёнными выбросами.

9. Для каждой компании проверьте на 5%-ном и 1%-ном уровне значимости гипотезу о том, что математическое ожидание логдоходности равно нулю при альтернативной гипотезе о том, что оно больше нуля. В каждом случае вычислите также наблюдаемый уровень значимости.

9.1\* Для каждой пары компаний проверьте на 5%-ном и 1%-ном уровне значимости гипотезу о том, что математические ожидания логдоходностей этих эмитентов равны между собой (без какого-либо предположения о равенстве дисперсий) при альтернативной гипотезе о том, что они не равны.

9.2\* Для каждой пары компаний проверьте на 5%-ном и 1%-ном уровне значимости гипотезу о том, что дисперсии логдоходностей этих эмитентов равны между собой при альтернативной гипотезе о том, что они не равны.

10. Для каждой компании проверьте на 5%-ном уровне значимости гипотезы о нормальном законе распределения логдоходности для данных без удаления выбросов и нулей (критерий Хи-квадрат в табличном процессоре или критерий Лиллие-форса в R или и то и другое).

10.1\*. Проведите эту проверку для данных с удалёнными выбросами

10.2\*. Проверьте гипотезу о нормальности в табличном процессоре с помощью критерия Лиллиефорса.

Критерии бальной оценки различных форм текущего контроля успеваемости содержится в соответствующих методических рекомендациях кафедры.

## 7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по данной дисциплине

Перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы содержится в разделе 2. *«Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине».*

### Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, знаний и умений

Таблица 6

Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
Владение основными научными понятиями и категориальным аппаратом современной экономики и их применение при решении прикладных задач (ПКН-1)	1. Демонстрирует знание современных экономических концепций, моделей, ведущих школ и направлений развития экономической науки, использует категориальный и научный аппарат при анализе экономических явлений и процессов.	<b>Знает</b> концептуальные постановки прикладных экономических задач, решаемых с применением методов математической статистики; <b>Умеет</b> выбирать адекватные методы анализа экономических явлений и процессов	По результатам опроса в рамках исследования состояния малого предпринимательства ответы респондентов на определенный вопрос анкеты представлены в виде выборки (задается выборка) Используя табличный процессор или R (RStudio), очистите выборку от пропусков, обозначенных как "NA", и ответьте на следующие ниже вопросы. 1. Определите объем исходной выборки 2. Определите количество пропущенных данных "NA" в исходной выборке 3. Определите объем очищенной от "NA" выборки 4. Определите количество различных вариантов ответов респондентов, встречающиеся в очищенной выборке 5. Определите количество респондентов, которые дали ответ (приводится конкретная варианта)

			<p>6. Определите долю респондентов, которые дали ответ (приводится конкретная варианта)</p> <p>7. Определите левую границу 0,95-доверительного интервала для истинной доли ответов (приводится конкретная варианта)</p> <p>8. Введите правую границу 0,95-доверительного интервала для истинной доли ответов ответ (приводится конкретная варианта)</p>
	<p>2. Выявляет сущность и особенности современных экономических процессов, их связь с другими процессами, происходящими в обществе, критически переосмысливает текущие социально-экономические проблемы.</p>	<p><b>Знает</b> основные понятия и методы теории вероятностей и прикладной статистики, необходимые для анализа социально-экономических задач и процессов;</p> <p><b>Умеет</b> решать социально-экономические задачи вероятностными и статистическими методами</p>	<p>Дана трехмерная выборка показателей X,Y,Z предприятия ABC. По данным выборки вычислите коэффициенты корреляции Пирсона для каждой пары признаков, проверьте гипотезы о незначимости каждого из коэффициентов</p>
	<p>3. Грамотно и результативно пользуется российскими и зарубежными источниками научных знаний и экономической информации, знает основные направления экономической политики государства.</p>	<p><b>Знает</b> источники актуальных данных о состоянии экономических субъектов;</p> <p><b>Умеет</b> осуществлять поиск и импорт данных, необходимых для решения конкретных задач</p>	<p>Представьте данные, полученные из Глобальной сети Интернет в формате csv, в виде таблицы. Сколько пропущенных значений (N/A) содержит признак, размещенный в первом столбце? Каков объем многомерной выборки после удаления пропусков?</p>
<p>Способность осуществлять сбор, обработку и статистический анализ данных, применять математические методы для решения стандартных профес-</p>	<p>1.Проводит сбор, обработку и статистический анализ данных для решения финансово-экономических задач.</p>	<p><b>Знает</b> тенденции развития новых программных продуктов и ИТ-сервисов, ориентированных на повышение эффективности сбора и обработки данных;</p>	<p>По данным об изменении социально – экономического показателя за последние 24 месяца (поиск и импорт данных осуществляется самостоятельно) постройте доверительный интервал прогноза этого показателя на следующий месяц</p>

сиональных финансово-экономических задач, интерпретировать полученные результаты (ПКН-3)		<b>Умеет</b> осуществлять поиск, сбор и первичную подготовку данных для экономических исследований	
	2.Формулирует математические постановки финансово-экономических задач, переходит от экономических постановок задач к математическим моделям.	<b>Знает</b> основные принципы спецификации экономико-математических моделей <b>Умеет</b> формализовать постановки прикладных задач из сфер экономики и финансов	Клиент может обратиться за получением кредита в один из трёх банков (1, 2, 3). Вероятности обращения в каждый банк зависят от их удалённости и равны соответственно <b>0.3, 0.3 и 0.4</b> . Вероятности того, что данному клиенту будет отказано в выдаче кредита равны 0.3 – в банке "1", 0.3 – в банке "2" и 0.05 – в банке "3". Найдите вероятность $P$ того, что при первом же обращении кредит клиентом получен. Указать, в каком из банков (1,2,3) с наибольшей вероятностью он мог бы быть получен.
	3.Системно подходит к выбору математических методов и информационных технологий для решения конкретных финансово-экономических задач в профессиональной области.	<b>Знает</b> основные понятия теории вероятностей и математической статистики; методики расчетов, используемые при анализе данных; вероятностные и статистические методы. <b>Умеет</b> использовать инструменты описательной статистики и визуализации данных, вероятностные и статистические методы для решения профессиональных задач.	Даны три вектора с характеристиками торгов финансового актива. В первом векторе - номера дат, во втором – соответствующие этим датам цены и в третьем -соответствующие этим датам объёмы. 1. Расположите эти векторы в трёх столбцах на листе табличного процессора, начиная с ячеек B2, C2, D2 и озаглавьте столбцы словами "Дата", "Цена", "Объём". 2. Удалите строки, соответствующие датам, когда не было торгов. Признаком отсутствия торгов являются нулевые значения объёма. 3. Для оставшейся части временного ряда удалите строки, в которых присутствуют значения NA и/или нулевые значения цен. 4. Отсортируйте полученный временной ряд в порядке воз-

			<p>растания номеров дат.</p> <p>5. Введите в столбец А номера по порядку (в ячейке А2 - 1, А3 - 2 и т.д.)</p> <p>6. Введите в форму длину временного ряда N</p> <p>7. Введите в форму номер даты, цену и объём, соответствующие строке нового ряда.</p> <p>8. Вычислите среднее значение цены</p> <p>9. Вычислите среднее значение объёма</p> <p>10. Вычислите стандартное отклонение цены</p> <p>11. Вычислите стандартное отклонение объёма</p> <p>12. Вычислите стандартную ошибку выборки для цены</p> <p>13. Вычислите минимальное значение цены</p> <p>14. Вычислите максимальное значение объёма</p> <p>15. Вычислите значение коэффициента эксцесса цены</p>
	<p>4. Анализирует результаты исследования математических моделей финансово-экономических задач и делает на их основании количественные и качественные выводы и рекомендации по принятию финансово-экономических решений.</p>	<p><b>Знает</b></p> <p>основные принципы программирования алгоритмов статистической обработки данных в табличных процессорах и R(RStudio) для анализа и прогноза внутри- и внешне-экономических процессов</p> <p><b>Умеет</b></p> <p>применять технические и инструментальные средства для решения прикладных экономических задач и подготовки обоснований управленческих решений</p>	<p>На основании ряда совместных наблюдений цен акций трех компаний определите, какие из инвестиционных операций по покупке акций указанных компаний будут оптимальными по Парето. Указание: меры прибыльности и риска оценить по значениям логдоходностей акций соответствующих компаний. Для каждой позиции в таблице ниже введите 1, если соответствующая операция оптимальна по Парето, введите 0 – в противном случае</p>



Способность оценивать показатели деятельности экономических субъектов (ПКН-4)	1. Проводит анализ внешней и внутренней среды ведения бизнеса, выявляет основные факторы экономического роста, оценивает эффективность формирования и использования производственного потенциала экономических субъектов.	<b>Знает</b> методы анализа тенденций развития экономических процессов; <b>Умеет</b> использовать табличные процессоры и R(RStudio) в статистическом анализе данных, характеризующих внутри- и внешнеэкономические процессы	По заданному временному ряду (приводится ряд для цены акции компании X) постройте ряд логдоходности (логдоходность равна натуральному логарифму отношения цены акции на текущую дату к цене акции на предыдущую дату). Вычислите статистические характеристики (математическое ожидание и среднее квадратическое отклонение) построенной логдоходности
	2. Рассчитывает и интерпретирует показатели деятельности экономических субъектов.	<b>Знает</b> условия и ограничения основных моделей в статистическом анализе <b>Умеет</b> сопоставлять реальные данные с математическими моделями и интерпретировать результаты статистических исследований	По данным многомерной выборки, содержащей сведения о поле, возрасте, месте проживания и уровне доходов жителей региона N (приводятся данные) 1) проверьте гипотезу о равенстве распределения мужчин и женщин по возрасту, используя критерий однородности Колмогорова-Смирнова 2) проверьте гипотезу о равенстве математических ожиданий доходов респондентов, проживающих в городской и в сельской местностях

### Примерные вопросы для подготовки к зачету:

1. Данные в экономике. Типы признаков в экономике и управлении: интервальные, порядковые, ранговые, дихотомические.
2. Измерение центра распределения. Измерение разброса данных.
3. Выбросы и их обработка. Пропущенные значения и их обработка. Повторяющиеся строки и их обработка. Синтетические признаки.

4. Основы комбинаторики. Правила суммы и произведения. Перестановки, размещения и сочетания без повторений. Перестановки, размещения и сочетания с повторениями.

5. Определение вероятности. Случайные события, их виды. Операции над событиями как операции над множествами. Классическая вероятностная схема.

6. Схема геометрических вероятностей.

7. Статистическая вероятность.

8. Аксиоматическое построение теории вероятностей.

9. Теорема сложения вероятностей.

10. Условные вероятности. Независимость событий. Формула полной вероятности. Формула Байеса.

11. Последовательности испытаний. Биномиальная схема. Формула Бернулли. Последовательности испытаний в экономике и управлении.

12. Определение случайной величины. Понятие случайной величины.

13. Функция распределения случайной величины. Свойства функции распределения.

14. Индикатор события как простейшая случайная величина. Функция распределения индикатора события.

15. Дискретные случайные величины и их важнейшие числовые характеристики. Дискретная случайная величина. Ряд распределения и функция распределения дискретной случайной величины.

16. Математическое ожидание, дисперсия и среднее квадратическое отклонение дискретной случайной величины.

17. Биномиальный закон распределения.

18. Геометрический закон распределения.

19. Закон распределения Пуассона.

20. Простейший поток событий.

21. Гипергеометрический закон распределения.

22. Сравнение случайных величин: отношение предпочтения, ожидаемая полезность, оптимальность по Парето.

23. Абсолютно непрерывные случайные величины и их важнейшие числовые характеристики. Абсолютно непрерывная случайная величина. Функция распределения и функция плотности распределения абсолютно непрерывной случайной величины. Свойства функции плотности распределения.

24. Математическое ожидание, дисперсия и среднее квадратическое отклонение абсолютно непрерывной случайной величины.

25. Равномерный закон распределения.

26. Показательный закон распределения.

27. Нормальный закон распределения.

28. Логарифмически нормальный закон распределения и ценообразование финансовых инструментов.

29. Закон распределения Парето и задачи налогообложения.

30. Законы распределения, важные в математической статистике (законы распределения Стьюдента,  $\chi^2$ , Фишера — Снедекора).

31. Смеси распределений.

32. Начальные и центральные моменты случайной величины. Асимметрия и эксцесс случайной величины.

33. Квантили и процентные точки случайной величины.

34. Ценность под риском.

35. Медиана и мода случайной величины.

36. Случайные векторы и условные законы распределения. Условный ряд распределения (для дискретных случайных величин), условная плотность распределения (для непрерывных случайных величин).

37. Условное математическое ожидание. Формула полного математического ожидания. Формула полной дисперсии.

38. Ковариация и коэффициент корреляции.

39. Портфель финансовых инструментов

40. Функции случайных величин. Функции одной случайной величины. Функции нескольких случайных величин. Формула композиции. Композиция равномерных случайных величин.

41. Закон больших чисел. Массовые случайные явления в экономике. Теорема Чебышёва и оценка математического ожидания. Теорема Бернулли и оценка вероятности. Обсуждение условий статистической устойчивости.

42. Центральная предельная теорема. Теорема Леви. Интегральная теорема Муавра — Лапласа. Место центральной предельной теоремы в изучении статистических закономерностей в экономике, финансах и управлении.

43. Математические основы теории страхования.

44. Метод Монте-Карло. Моделирование случайных величин.

### **Примерные вопросы для подготовки к экзамену:**

1. Основы выборочного метода. Генеральная и выборочная совокупности. Случайная и конкретная выборки. Случайная повторная и случайная бесповторная выборка.

2. Предмет и задачи математической статистики.

3. Соотношение между предельной ошибкой выборки, уровнем значимости (риском) и объемом выборки. Использование этого соотношения в организации выборочных обследований.

4. Оценка плотности распределения и функции распределения. Вариационный ряд. Выборочная случайная величина (статистический ряд распределения). Интервальный вариационный ряд. Полигон частот, кумулята.

5. Оценка числовых характеристик генеральной случайной величины с помощью выборочной случайной величины. Выборочное среднее как оценка математического ожидания. Относительная частота как оценка вероятности. Выборочная дисперсия как оценка дисперсии.

6. Точечные оценки параметров. Понятие точечной оценки параметра генеральной совокупности. Свойства точечных оценок: состоятельность, несмещенность, эффективность.

7. Выборочное среднее как состоятельная, несмещенная и эффективная оценка математического ожидания генеральной случайной величины.

8. Смещенность выборочной дисперсии как оценки дисперсии генеральной случайной величины. Исправленная выборочная дисперсия как несмещенная и состоятельная оценка дисперсии генеральной случайной величины.

9. Методы построения точечных оценок: метод моментов, метод максимального правдоподобия. Примеры построения оценок параметров распределений случайных величин, применяемых в экономике и управлении.

10. Интервальные оценки параметров. Понятие интервальной оценки параметра генеральной совокупности. Точные интервальные оценки вероятности, математического ожидания, дисперсии и коэффициента корреляции. Поправка на конечный объем генеральной совокупности. Асимптотический подход к интервальному оцениванию.

11. Статистические гипотезы. Понятие статистической гипотезы. Виды статистических гипотез: параметрические и непараметрические, простые и сложные.

12. Критерий проверки гипотезы, критическое множество. Проверка гипотез с помощью интервальных оценок. Ошибки первого и второго родов. Мощность критерия. Наиболее мощный критерий.

13. Проверка гипотезы о равенстве математического ожидания теоретическому значению. Проверка гипотезы о равенстве двух математических ожиданий.

14. Проверка гипотезы о равенстве дисперсии теоретическому значению. Проверка гипотезы о равенстве двух дисперсий.

15. Проверка гипотезы о равенстве вероятности события теоретическому значению. Проверка гипотезы о равенстве двух вероятностей.

16. Проверка гипотез о значимости коэффициента корреляции.

17. Критерии согласия. Критерий согласия  $\chi^2$  Пирсона. Критерий  $\chi^2$  Пирсона при неизвестных параметрах распределения.

18. Однофакторный дисперсионный анализ.

19. Двухфакторный дисперсионный анализ.

20. Критерий  $\chi^2$  для проверки однородности данных.

21. Непараметрические критерии. Проверка гипотез на малых выборках.

Критерий знаков.

22. Распределение Вилкоксона и его критические границы.

Непараметрическая точечная оценка математического ожидания.

23. Непараметрическая интервальная оценка математического ожидания.

24. Ранговая корреляция. Коэффициент ранговой корреляции Спирмена.

Коэффициент ранговой корреляции Кендалла. Коэффициент конкордации.

25. Проверка гипотез о значимости ранговых коэффициентов корреляции.

Примеры использования ранговой корреляции в экономике.

## Пример экзаменационного билета

Билеты генерируются в системе Moodle.

Пример билета приводится в аутентичной транскрипции данной системы постранично.

**Вопрос 1**  
Пока нет ответа  
Балл: 30,00  
Отметить вопрос  
Редактировать вопрос

Независимые наблюдения нормально распределенной случайной величины  $X$ , описывающей приращение стоимости акций некоторой компании, представлены в виде выборки: {NA; NA; NA; -147.14; -158.18; NA; -257.29; -168.67; -123.1; -176.29; -160.78; -158.28; -173.04; -190.61; -164.68; -161.87; -175.23; -163.61; NA; -130.52; -138; -184.18; -137.21; -160.62; -108.5; -186.51; NA; -192.64; -200.66; -135.2; -129.28; -149.95; -173.23; -135.99; -173.13; -161.34; -170.02; -174.53; NA; -174.86; NA; -189.47; -170.93; -146.73; NA; -99.43; NA; -130.19; -163.24; -174.33; -178.57; -202.39; -124.74; -163.31; -114.8; -177.37; NA; -122.79; -155.05; -162.74; -174.64; NA; -149.37; -188.37; NA; -144.6; -118.11; -167.92; -148.17; -161.38; -133.12; -191.19; NA; -208.31; -179.24; -127.52; -149.36; -151.94; -124.52; NA; -139.49; -204.18; -157.63; -187.98; -158.74; -154.04; NA; -125.65; -172.15; -145.13; -138.48; -155.31; -163.84; -154.95; -134.23; -169.13; -156.23; NA; -109.99; -129.92; -188.98; -161.42; -158.07; -131.82; NA; -177.58; NA; NA; -184.32; -159.49; -172.18; -131.57; -139.35; -135.43; -139.13; -166.47; -176.03; -185.28; -168.78; -155.34; -180.87; NA; -160.34; -167.18; NA; -116.73; -152.29; NA; -145.97; -150.13; -167.71; -254.77; -185.26; -150.3; -187.65; -204.59; -129.49; -162.78; -162.62; -111.04; NA; -172.9; -141.73; -172.69; -171.27; -174.14; -154.32; -170.7; -115.48; -107.68; NA; -145.76; -166.04; -156.34; -151.22; -142.44; -130.36; -159.56; -154.11; -156.68; -198.53; -185.93; -176.52; -136.8; -139.02; -152.56; -124.09; -124.92; -166.08; -153.1; -156.07; -128.13; -138.13; NA; -166.58; -117.82; -167.81; -163.07; -176.73; -263.72; -161.46; -144.77; -156.56; -156.34; -156.89; -109.34; NA; -197.82; -156.81; NA; -178.56; -161.02; -169.33; -140.55; -117.48; -154.58; -170.76; -195.78; -171.87; NA; -150.94; -145.75; NA; NA; -137.43; -128.99; -136.38; -178.31; -158.46; -152.95; -160.1; -151.91; -180; -132.63; NA; -145.85; -211.5; -170.88; -171.91; -194.47; -182.26; -179.42; -166.17; NA; -193.83; -171.98; NA; -158.81; NA; -144.42; -144.91; -189; -151.31; -158.13; -163.54; -123.12; -109.8; -182.86; -115.39; -157.91; -145.43; -171.55; -176.57; -138.31; -153.23; -153.25; -177.86; -124.74; NA; -163.32; -162.97; -173.05; NA; -147.49; -127.81; -162.61; -138.07; -154.85; -187; -141.97; -151.6; NA; NA; -137.06; -118.72; -169.03; -117.61; NA; -121.47; -158.77; -194.19; -186.28; -139.04; -157.03; -169.59; -157.45; -154.29; NA; -172.49; -141.43; -205.98; NA; -180.68; -180.08; -153.77; -158.14; -171.02; -160.84; -115.28; -182.75; -179.52; -121.12; -193.54; -160.76; -187.04; NA; -185.71; -166.86; -189.84; -155.72; -189.29; -181.34; -185.98; -171.21; -138.2; NA; -192.74; -144.44; -137.23; -164.1}.

Скопируйте и преобразуйте в столбец "A" данные выборки на лист "Лист1" Excel-файла и, используя Excel или R(RStudio), вычислите требуемые ниже величины.

1. Введите количество пропущенных значений в исходной выборке, обозначенные как "NA"

Очистите исходную выборку от пропусков "NA", преобразуйте её в вариационный ряд и работайте далее с полученным рядом.

2. Введите объем очищенной от пропусков выборки

3. Введите минимальное значение в вариационном ряду

4. Введите максимальное значение в вариационном ряду

5. Введите первую квартиль

6. Введите медиану

7. Введите третью квартиль

8. Введите квартильный размах

9. Введите среднее значение

10. Введите стандартное отклонение (несмещенное)

11. Введите исправленную дисперсию

12. Введите ошибку выборки

13. Введите эксцесс (формула по умолчанию в Excel)

14. Введите коэффициент асимметрии (формула по умолчанию в Excel)

15. Введите левую границу 0.99-доверительного интервала для  $E(X)$

16. Введите правую границу 0.99-доверительного интервала для  $E(X)$

17. Введите левую границу 0.99-доверительного интервала для  $Var(X)$

18. Введите правую границу 0.99-доверительного интервала для  $Var(X)$

19. Введите квантиль уровня 0.6

20. Введите общее количество выбросов

21. Постройте на листе "Лист1" гистограмму и диаграмму "ящик с усами" для исходной выборки, очищенной от "NA" и выбросов. Если построения произведены в R(RStudio), то скопируйте полученные диаграммы из RStudio на "Лист1".



**Вопрос 2**

Пока нет ответа

Балл: 20,00

Отметить  
вопросРедактировать  
вопрос

По результатам социологического исследования ответы респондентов на определенный вопрос анкеты представлены в виде выборки: {D; D; C; D; NA; D; C; B; C; D; NA; D; B; B; C; D; D; D; B; B; D; A; B; D; NA; B; D; B; A; B; D; C; D; D; C; B; D; D; NA; A; D; D; B; A; D; D; NA; NA; NA; A; A; D; C; A; NA; NA; D; D; D; B; B; C; B; NA; D; C; NA; B; B; D; NA; B; D; D; D; D; D; B; D; D; B; D; A; D; D; D; D; B; NA; D; A; NA; D; B; D; NA; D; NA; NA; A; D; B; D; D; D; D; D; A; A; D; A; B; B; D; D; NA; NA; D; D; NA; D; NA; B; B; B; D; D; D; NA; D; D; D; D; D; C; D; D; C; D; A; D; D; NA; D; D; A; D; B; D; D; C; D; D; D; D; D; NA; C; D; D; D; A; D; D; NA; D; B; NA; D; D; NA; C; A; D; D; C; B; D; NA; A; C; NA; B; B; C; D; D; D; NA; D; D; C; NA; D; D; D; A; D; D; D; D; B; D; B; A; D; D; NA; B; B; D; A; B; B; D; B; D; D; D; D; D; D; D; NA; B; D; C; A; B; D; B; A; B; A; D; D; B; NA; B; B; A; D; D; D; D; B; D; A; B; C; D; A; B; D; B; D; NA; C; D; B; A; D; D; D; A; D; D; D; NA; D; D; C; NA; D; NA; D; D; B; NA; D; D; D; B}.

Скопируйте и преобразуйте в столбец "A" данные выборки на лист "Лист2" Excel-файла.

Используя Excel или R(RStudio), очистите выборку от пропусков, обозначенных как "NA", и ответьте на следующие ниже вопросы.

1. Введите количество различных вариантов ответов респондентов, встречающиеся в очищенной выборке

2. Введите объем очищенной от "NA" выборки

3. Введите количество пропущенных данных "NA" в исходной выборке

4. Введите долю респондентов, которые дали ответ "B"

5. Введите правую границу 0.99-доверительного интервала для истинной доли ответов "B"

6. Введите левую границу 0.99-доверительного интервала для истинной доли ответов "B"

На уровне значимости 0.01 проверьте критерием согласия (Хи-квадрат критерием Пирсона) гипотезу о равномерном распределении ответов респондентов.

7. Введите критическое значение статистики хи-квадрат

8. Введите количество степеней свободы

9. Введите наблюдаемое значение хи-квадрат

10. Введите 1, если есть основания отвергнуть гипотезу о равномерном распределении ответов, или введите 0, если таких оснований нет.

11. Постройте на листе "Лист2" гистограмму для исходной выборки, очищенной от "NA". Если построения произведены в R(RStudio), то скопируйте полученные диаграммы из RStudio на "Лист2".


**Вопрос 3**

Пока нет ответа

Балл: 10,00

Отметить

вопрос

 Редактировать  
вопрос

Ряд совместных наблюдений независимых нормально распределенных случайных величин  $X$  и  $Y$ , описывающих некоторый финансовый показатель двух фирм, задан двумерной выборкой:

{(NA, -200.166); (NA, -147.812); (-208.4021, -150.193); (-247.5678, -182.045); (-165.426, -148.965); (-211.5128, -129.144); (-220.7977, -177.937); (-217.9882, -167.629); (-224.1744, -137.087); (-122.5397, -147.198); (-192.2301, -167.659); (-221.5835, NA); (-220.378, -182.103); (-235.512, NA); (-240.679, -155.163); (-172.3632, -163.97); (-140.9888, -129.314); (NA, -161.071); (-179.7534, -163.171); (-181.0125, -141.497); (-193.4819, -97.889); (-206.7622, -157.369); (-160.6644, -135.882); (-128.6632, -166.681); (-201.6559, NA); (-222.3484, -157.992); (-201.5761, -140.886); (-184.4567, NA); (-199.4676, -115.91); (-224.093, -184.092); (NA, -131.163); (NA, -144.17); (-217.6648, -161.813); (-212.1995, -146.888); (-213.5616, NA); (-195.9949, -165.285); (-209.9585, -147.199); (-212.0849, -154.144); (-175.7262, -149.01); (-208.0548, -155.185); (-223.6774, -131.436); (-192.2471, -175.21); (NA, -109.6); (-187.2649, -146.37); (-234.4358, -146.447); (-207.0624, -178.951); (-187.1979, -129.548); (-175.0766, -123.593); (-209.2384, -162.745); (-183.9404, -153.508); (-211.9154, -92.389); (NA, -164.114); (-233.6673, -145.726); (-198.64, -162.603); (-200.003, -148.838); (-190.9473, -142.562); (-219.6943, -204.87); (-221.2966, -155.57); (-170.0186, -195.506); (NA, -176.575); (-201.7801, -156.711); (-195.7176, -140.589); (NA, -163.016); (-231.7002, -126.977); (-224.5886, NA); (-192.4474, -147.955); (-223.6718, -154.911); (-209.795, -180.448); (-188.5438, -137.721); (-191.1335, -175.409); (NA, -142.667); (-156.1438, -131.749); (-227.4467, -147.128); (NA, -188.359); (-210.5232, -128.753); (NA, -157.582); (-197.5018, -134.994); (-236.1859, -162.574); (NA, NA); (-203.1479, -117.507); (-193.546, -171.226); (NA, -162.768); (-223.9803, -153.473); (-207.4509, -143.656); (-212.3123, -170.853); (-199.1342, -110.963); (-263.3709, -170.754); (-200.1349, NA); (-203.6951, -157.855); (-227.7833, -143.801); (-189.8768, -116.997); (-201.7683, -152.184); (-200.9583, -144.026); (-237.1278, -156.199); (-212.2596, -150.202); (-216.6751, -130.428); (-189.8154, -155.996); (-194.5055, -144.686); (-201.7344, -172.07); (-215.7982, -148.016); (-171.0562, -146.041); (-183.9788, -119.36); (-170.5053, -185.588); (-196.9646, -135.598); (-209.2727, -134.338); (-217.5302, -169.638); (-169.0613, -149.659); (-176.0433, -141.122); (NA, -170.243); (-237.6997, -142.279); (-195.8273, -139.075); (-183.4436, -160.424); (NA, -160.026); (-220.5686, -104.783); (NA, -160.683); (-186.5226, -151.203); (-218.6891, -190.973); (-209.1348, -173.84); (-227.5963, -148.452); (-219.1326, -111.848); (-186.9232, -119.916); (-235.9756, -162.446); (-212.9752, -146.357); (-141.4253, -185.335); (-202.2192, -117.827); (-237.1769, -109.862); (-206.4806, -186.751); (-206.4279, NA); (-218.2599, -135.402); (-213.9562, -144.986); (-223.4841, -144.963); (-217.6743, -203.344); (-181.5788, -129.673); (-184.8199, -128.654); (NA, -150.983); (-204.1213, -177.369); (-205.3603, -124.538); (-182.9048, NA); (-227.9288, -166.395); (-192.0839, -124.399); (-221.5064, -164.629); (-189.8635, -133.957); (-250.9996, -216.395); (-215.2317, -173.513); (-179.4822, -162.965); (-190.6926, -144.343); (-246.6572, -129.085); (-240.9944, -132.645); (NA, -140.787); (-194.2856, -141.749)}.

Скопируйте данную выборку на лист "Лист3" и преобразуйте ее в столбцы "A" и "B" соответственно для первой и второй фирмы. При этом связанные значения показателей должны располагаться в одной строке.

Используя Excel или R(RStudio), очистите исходную выборку от пропущенных данных, обозначенных как "NA", и вычислите требуемые ниже величины.

1. Введите выборочный коэффициент корреляции Пирсона между  $X$  и  $Y$

2.1 Введите значение P-value в проверке гипотезы о равенстве средних значений показателей фирм при альтернативной гипотезе об их неравенстве (без каких-либо предположений о равенстве дисперсий)

2.2 На уровне значимости 0.1 можно ли утверждать, что средние значения показателей у фирм различны?

Введите 1 - если да, и 0 - если нет

3.1 Введите значение P-value в проверке гипотезы о равенстве дисперсий показателей двух фирм при альтернативной гипотезе об их неравенстве

3.2 На уровне значимости 0.05 можно ли утверждать, что дисперсии показателей фирм различны? Введите 1 - если да, и 0 - если нет

**Вопрос 4**

Пока нет ответа

Не оценен

Отметить

Приложите файл(ы) Excel и/или R с вычислениями итоговой контрольной работы.

**Внимание! Без прикрепленных файлов с вычислениями работа считается невыполненной и не подлежит проверке!**

Максимальный размер новых файлов: 20Мбайт, максимальное количество прикрепленных файлов: 3

## 8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

### Рекомендуемая литература

#### а) основная:

1. Соловьев, В. И. Анализ данных в экономике: теория вероятностей, прикладная статистика, обработка и визуализация данных в Microsoft Excel: учебник для направления бакалавриата "Экономика и управление" / В. И. Соловьев; Финуниверситет. - Москва: Кнорус, 2019. - 498 с. - Текст : непосредственный. - То же. — 2021. - ЭБС BOOK.ru. - URL: <https://book.ru/book/938856> (дата обращения: 02.11.2022). — Текст: электронный.

2. Теория вероятностей и математическая статистика: учебник-практикум / А. В. Браилов, В. И. Глебов, С. Я. Криволапов, П. Е. Рябов. — Москва. — Ижевск : НИЦ Регулярная и хаотическая динамика ; Институт компьютерных исследований, 2016. — 414 с. — [org.fa.ru](http://org.fa.ru). — URL: [https://org.fa.ru/app/umm/tree?login=yes&\\_url=%2Fumm%2Ftree&documentId=%7BAEF22B18-ACDC-4743-859D-C7FB61A343BF%7D](https://org.fa.ru/app/umm/tree?login=yes&_url=%2Fumm%2Ftree&documentId=%7BAEF22B18-ACDC-4743-859D-C7FB61A343BF%7D) (дата обращения: 02.11.2022). - Текст: электронный.

#### б) дополнительная:

3. Браилов А. В. Практикум для самостоятельной работы студентов по дисциплине "Теория вероятностей и математическая статистика" / А. В. Браилов, С. А. Задаев, П. Е. Рябов. — Москва: Финуниверситет, 2014. — ЭБ Финуниверситета. — URL: <http://elib.fa.ru/rbook/praktikum.pdf/view> (дата обращения: 02.11.2022). — Текст: электронный.

4. Зададаев, С. А. Применение R (RStudio) в математической статистике: учебное пособие для проведения семинаров по дисциплине «Анализ данных»/ С. А. Зададаев, М. В. Солдаткина. Применение R-studio в математической статистике. – Москва: Финансовый университет, Департамент анализа данных, принятия решений и финансовых технологий, 2019. – 29 с. — org.fa.ru. - URL:[https://org.fa.ru/app/umm/tree?login=yes&\\_url=%2Fumm%2Ftree&documentId=%7B1008D5E1-A4D5-49DA-9548-D148C78A187E%7D](https://org.fa.ru/app/umm/tree?login=yes&_url=%2Fumm%2Ftree&documentId=%7B1008D5E1-A4D5-49DA-9548-D148C78A187E%7D) (дата обращения: 02.11.2022).— Текст: электронный.

5. Расчетно-аналитическая работа по математической статистике "Практикум для студентов, обучающихся по направлениям подготовки 38.03.01 «Экономика», 38.03.02 «Менеджмент», 39.03.01 «Социология» (все профили)" / Т. Г. Апалькова, В. И. Глебов, С. Я. Криволапов, К. Г. Левченко. - Москва: Финуниверситет, Департамент математики, 2021. - 126с. - org.fa.ru. - URL: [https://org.fa.ru/app/umm/tree?login=yes&\\_url=%2Fumm%2Ftree&documentId=acb641b2-6161-44a8-a063-7ee66edddd6e8](https://org.fa.ru/app/umm/tree?login=yes&_url=%2Fumm%2Ftree&documentId=acb641b2-6161-44a8-a063-7ee66edddd6e8) (дата обращения: 02.11.2022). - Текст: электронный.

6. Анализ данных. Часть 2. учебное пособие для студентов, обучающихся по направлениям подготовки 38.03.01 «Экономика», заочная и заочная ускоренная формы обучения / А. В. Потемкин, М. Н. Фридман, И. И. Цыганок, И. М. Эйсымонт. — Москва: Финансовый университет; Департамент анализа данных, принятия решений и финансовых технологий, 2018.- 59 с.- org.fa.ru.- URL:[https://org.fa.ru/app/umm/tree?login=yes&\\_url=%2Fumm%2Ftree&documentId=%7BD0A63C98-BBFA-4D1A-A5F3-584A72263258%7D](https://org.fa.ru/app/umm/tree?login=yes&_url=%2Fumm%2Ftree&documentId=%7BD0A63C98-BBFA-4D1A-A5F3-584A72263258%7D) (дата обращения: 02.11.2022). — Текст: электронный.

## **9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины**

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <https://org.fa.ru/>

2. Система дистанционного обучения Финансового Университета

URL: <https://campus.fa.ru>

3. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>

4. Электронно-библиотечная система BOOK.RU <http://www.book.ru>

5. Электронно-библиотечная система «Университетская библиотека ОН-ЛАЙН» <http://biblioclub.ru/>

6. Электронно-библиотечная система Znanium <http://www.znanium.com>

7. Электронно-библиотечная система издательства «ЮРАЙТ» <https://urait.ru/>

8. Электронно-библиотечная система издательства Проспект <http://ebs.prospekt.org/books>

9. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>

10. Электронная библиотека Издательского дома «Гребенников» <https://grebennikon.ru/>

## **10. Методические указания для обучающихся по освоению дисциплины**

Обучающимся в рамках самостоятельной работы следует использовать Методические рекомендации по планированию и организации внеаудиторной самостоятельной работы студентов по образовательным программам бакалавриата и магистратуры в Финансовом университете, утвержденные Приказом ректора №1040/о от 11.05.2021 г.

Самостоятельная работа студентов проходит внеаудиторно. Организации самостоятельной работы служит учебно-тематический план изучения дисциплины. В данном плане указана тематика лекций, семинаров, вопросы и задания для самостоятельного изучения. Во время лекций необходимо конспектировать содержание лекции. После лекции необходимо отредактировать записи, оформить конспект, дополняя его содержание дополнительной информацией. При оформлении конспекта целесообразно выделять названия тем и формулировки вопросов, основные определения, примеры.

При подготовке к семинару необходимо изучить вопросы семинара, соответствующий теоретический материал, делая для себя необходимые записи в рабочей тетради. После занятий необходимо просмотреть записанные решения и восстановить в решениях имеющиеся пробелы.

При затруднении в решении практических вопросов (задач), можно обратиться за консультацией (помощью) к преподавателю. Семинары проходят, как правило, в интерактивной форме и преподаватель учитывает активность обучающихся,

направленную на решение предложенных вопросов (вариантов задач), а также вариантов ответов на решаемые вопросы (проблемы).

Не следует бояться дать неверный ответ или допустить иную ошибку: исправление и анализ ошибок в режиме общения с преподавателем и сокурсниками в ходе семинара способствует более глубокому освоению учебного материала и предупреждает возникновение ошибок в дальнейшем. Домашние задания (подготовку к занятиям) следует осуществлять регулярно. Если то или иное задание, при подготовке к семинару вызвало затруднение, необходимо обратиться к преподавателю за консультацией. Регулярность в выполнении домашних заданий (подготовке к занятиям) - важный фактор качественного освоения дисциплины.

### **Методические рекомендации по обучению лиц с ограниченными возможностями здоровья**

Профессорско-преподавательский состав знакомится с психолого-физиологическими особенностями обучающихся инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ). При необходимости осуществляется дополнительная поддержка преподавания тьюторами, психологами, социальными работниками, прошедшими подготовку ассистентами.

В соответствии с методическими рекомендациями Минобрнауки РФ (утв. 8 апреля 2014 г. № АК-44/05вн) в курсе предполагается использовать социально-активные и рефлексивные методы обучения, технологии социокультурной реабилитации с целью оказания помощи в установлении полноценных межличностных отношений с другими студентами, создании комфортного психологического климата в студенческой группе. Подбор и разработка учебных материалов производятся с учётом предоставления материала в различных формах: аудиальной, визуальной, с использованием специальных технических средств и информационных систем. Освоение дисциплины лицами с ОВЗ осуществляется с использованием средств обучения общего и специального назначения (персонального и коллективного использования). Материально-техническое обеспечение предусматривает приспособление аудиторий к нуждам лиц с ОВЗ. Форма проведения аттестации для студентов-инвалидов устанавливается с учётом индивидуальных психофизических особенностей.

Для студентов с ОВЗ предусматривается доступная форма предоставления заданий оценочных средств, а именно:

- в печатной или электронной форме (для лиц с нарушениями опорно-двигательного аппарата);
  - в печатной форме или электронной форме с увеличенным шрифтом и контрастностью (для лиц с нарушениями слуха, речи, зрения);
  - методом чтения ассистентом задания вслух (для лиц с нарушениями зрения).
- Студентам с инвалидностью увеличивается время на подготовку ответов на контрольные вопросы.

Для таких студентов предусматривается доступная форма предоставления ответов на задания, а именно:



- письменно на бумаге или набором ответов на компьютере (для лиц с нарушениями слуха, речи);
- выбором ответа из возможных вариантов с использованием услуг ассистента (для лиц с нарушениями опорно-двигательного аппарата);
- устно (для лиц с нарушениями зрения, опорно-двигательного аппарата). При необходимости для обучающихся с инвалидностью процедура оценивания результатов обучения может проводиться в несколько этапов.

### **Методические рекомендации по выполнению контрольной работы, РАР**

Методические рекомендации по выполнению **контрольной работы, РАР** предусмотрены в «Методических рекомендациях по подготовке написанию и оформлению контрольной работы, РАР», разрабатываемой преподавателем кафедры на учебный год, в котором реализуется учебная дисциплина

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

#### **11.1 Комплект лицензионного программного обеспечения**

1. Антивирусная защита ESET NOD32
2. Windows, Microsoft Office
3. Astra Linux

11.2 Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Консультант Плюс»;
2. Информационно-правовая система «Гарант»;

11.3 Сертифицированные программные и аппаратные средства защиты информации

Сертифицированные программные и аппаратные средства защиты информации не предусмотрены

### **12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине**

- аудитории для проведения лекционных и семинарских занятий, оборудованные видеопроекционным оборудованием для презентаций, средствами звуковоспроизведения, экраном, доской меловой/интерактивной;
- библиотеку, имеющую рабочие места для студентов, оснащенные компьютерами с доступом к базам данных и сети Интернет
- компьютерные классы с набором лицензионного базового программного обеспечения для проведения практических занятий и выходом в глобальную сеть Internet;

### **Образовательные технологии**

В процессе изучения дисциплины «Анализ данных» предполагается:

- сопровождение курса лекций наглядной презентацией, включающей практические примеры, схемы, графики, табличный материал;
- рассмотрение на семинарских занятиях интерактивных ситуационных задач по проблематике дисциплины;
- деловые игры;
- разбор конкретных ситуаций, коллективное обсуждение проблем российской и зарубежной практики по изучаемым темам;
- виртуальное общение в течение срока изучения курса в целях обеспечения лекций и практических занятий необходимым материалом и также контроля самостоятельной работы студентов.