

Федеральное государственное образовательное бюджетное учреждение
высшего образования
**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финуниверситет)**

Тульский филиал Финуниверситета

Кафедра «Математика и информатика»

УТВЕРЖДАЮ

Директор филиала

Г.В. Кузнецов

«30» апреля 2024 г.



Б.П. Ваньков

АНАЛИЗ ДАННЫХ

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

38.03.02 «Менеджмент»

Образовательная программа «Финансовый менеджмент»

Профиль «Финансовый менеджмент»

Рекомендовано Ученым советом Тульского филиала Финуниверситета
(протокол от 23 апреля 2024 г. № 14)

Одобрено кафедрой «Математика и информатика»
(протокол от 5 марта 2024 г. № 8)

Тула 2024

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 образовательной программы	5
4. Объем дисциплины 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	5
5.1. Содержание дисциплины.....	5
5.2 Учебно-тематический план	8
5.3 Содержание семинаров, практических занятий	10
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	13
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	13
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю (согласно таблице 2)	15
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	19
7.1. Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний.....	19
7.2. Иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний	32
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.....	32
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	33
10. Методические указания для обучающихся по освоению дисциплины.....	33
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем.....	35
11.1 Комплект лицензионного программного обеспечения.....	35
11.2 Современные профессиональные базы данных и информационные справочные системы.....	35
12 Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	35

1. Наименование дисциплины

Анализ данных.

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

В совокупности с дисциплинами базовой части математического и естественнонаучного цикла дисциплина «Анализ данных» обеспечивает формирование следующих компетенций:

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
УК-4	Способность использовать прикладное программное обеспечение при решении профессиональных задач	1.Использует основные методы и средства получения, представления, хранения и обработки данных.	Знать: основные способы получения, отбора, хранения и обработки данных, способы доступа к различным баз с данными. Уметь: использовать техническое программное обеспечение для получения, отбора, хранения и обработки данных; уметь обращаться к внешним ресурсам для получения данных.
		2. Демонстрирует владение профессиональными пакетами прикладных программ.	Знать: основные принципы работы персонального компьютера; функциональные настройки инструментальных программных средств. Уметь: использовать пакеты прикладных программ общего назначения.
		3.Выбирает необходимое прикладное программное обеспечение в зависимости от решаемой задачи.	Знать: основные современные средства и информационные технологии при решении управленческих задач. Уметь: применять соответствующее программное обеспечение для моделирования ситуаций и их анализа.
		4. Использует прикладное программное обеспечение для решения конкретных прикладных задач.	Знать: современный инструментарий обработки реальных данных. Уметь: применять информационные инструменты к анализу результатов исследования реальных данных.

УК-15	Способность релевантно решаемым задачам использовать информационные ресурсы и информационно-коммуникационные технологии для достижения целей, связанных с профессиональной деятельностью, обучением, участием в жизни общества и других сферах жизни	1. Самостоятельно выбирает и использует цифровые средства общения, осуществляет поиск и/или создание контента в соответствии с целью взаимодействия, в том числе для организации совместной деятельности.	Знать: возможности информационных ресурсов для решения задач анализа данных. Уметь: использовать информационные ресурсы для решения задач анализа данных.
		2. Владеет навыками организации взаимодействия и коммуникации с помощью информационных систем и/или цифровых сервисов и технологий.	Знать: возможности цифровых сервисов и технологий для взаимодействия при решении задач анализа данных. Уметь: использовать цифровые сервисы и технологии для взаимодействия при решении задач анализа данных.
		3. Осуществляет подбор и применение различных информационно-коммуникационных средств для решения образовательных и профессиональных задач.	Знать: различные информационно-коммуникационные средства для решения задач анализа данных. Уметь: использовать различные информационно-коммуникационные средства для решения задач анализа данных.
ПКН-10	Владение методами количественного и качественного анализа информации, а также навыками построения моделей, применяя для анализа, моделирования и поддержки принятия решений современные информационные технологии и программные средства, включая инструменты бизнес-аналитики, обработки и анализа данных	1. Использует методы получения информации, ее анализа для построения моделей и интерпретации результатов моделирования	Знать: методы поиска информации, ее системного и критического анализа. Уметь: применять современные информационные технологии для сбора информации, использовать для построения адекватных математических моделей.
		2. Применяет приемы квалификации и выбора подходящих измерительных шкал при описании организационных систем, происходящих в них процессов и явлений	Знать: существующие измерительные шкалы, виды данных. Уметь: выбрать и применить подходящие измерительные шкалы, исследовать организационные системы.
		3. Использует навыки организации и проведения качественных и количественных исследований анализа информации, подготовки аналитических отчетов о состоянии и динамике развития рынков товаров и услуг.	Знать: методы и программы исследования, информационный инструментарий для анализа данных. Уметь: обосновать сущность происходящего, выявить закономерности, понять природу вариабельности, использовать инстру-

			менты анализа данных, исследовать динамику развития рынка, уметь составлять аналитический отчет.
--	--	--	--

3. Место дисциплины в структуре образовательной программы

Дисциплина «Анализ данных» входит в состав цикла математики и информатики обязательной части дисциплин учебного плана направления подготовки 38.03.02 «Менеджмент».

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

Таблица 1

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр (модуль) 3 (в часах)	Семестр (модуль) 4 (в часах)
Общая трудоёмкость дисциплины	7/252	126	126
<i>Контактная работа</i>	34	16	18
<i>Аудиторные занятия</i>			
Лекции (Л)	16	8	8
Семинары (ПЗ)	18	8	10
Самостоятельная работа (СР)	218	110	108
Вид текущего контроля		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

Данные в экономике. Объекты, признаки и таблицы. Типы признаков в экономике: интервальные, порядковые, ранговые, дихотомические. Форматирование наборов данных как электронных таблиц. Гистограммы. Условное форматирование. Графики и диаграммы рассеяния. Инструменты описательной статистики. Измерение центра распределения. Измерение разброса данных. Диаграммы размаха. Предварительная обработка данных. Выбросы и их обработка. Пропущенные значения и их обработка. Повторяющиеся строки и их обработка. Синтетические признаки.

Тема 2. Случайные события

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. Основные теоремы теории вероятностей. Теорема сложения вероятностей. Условная вероятность. Независимость событий. Формула полной вероятности. Формула Байеса. Простейшие примеры применения теории вероятностей.

Тема 3. Случайные величины

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. 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. 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. 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. 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 Стьюдента, χ^2 , Фишера — Снедекора). Реализация моделей абсолютно непрерывных случайных величин в электронных таблицах. 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. Центральные и начальные моменты, асимметрия, эксцесс

Моменты случайной величины. Асимметрия и эксцесс случайной величины. Квантили и процентные точки случайной величины. Сумма под риском (VaR). Медиана и мода случайной величины.

Тема 5. Центральная предельная теорема

Неравенство Чебышева. Закон больших чисел. Теорема Чебышева. Теорема Бернулли. Центральная предельная теорема Ляпунова. Центральная предельная теорема для одинаково распределенных слагаемых. Роль нормального закона в приложениях теории вероятностей. Теорема Леви.

Тема 6. Случайные процессы

Определение и способы задания цепей Маркова. Вероятности и матрица переходов. Многошаговые вероятности переходов и теорема о матрице многошаговых переходов. Предельные вероятности. Теорема Маркова о предельных вероятностях. Моделирование случайных величин методом Монте-Карло.

Тема 7. Выборочный метод

Статистические методы обработки экспериментальных данных. Генеральная совокупность и выборка. Допустимый объем выборки для обеспечения ее репрезентативности. Вариационный ряд. Оценка функции распределения (эмпирическая функция распределения) и плотности распределения. Гистограмма.

Тема 8. Точечные оценки параметров распределения

Понятие статистики, оценки числовой характеристики. Свойства точечных оценок: состоятельность, несмещенность и эффективность. Методы построения точечных оценок. Метод моментов. Метод максимального правдоподобия.

Тема 9. Интервальные оценки параметров.

Доверительный интервал для математического ожидания при известном значении дисперсии. Доверительный интервал для математического ожидания при неизвестном значении дисперсии. Доверительный интервал для вероятности. Доверительный интервал для дисперсии. Доверительный интервал для прогноза.

Тема 10. Статистические гипотезы.

Понятие статистической гипотезы. Виды гипотез: параметрические и непараметрические. Нулевая гипотеза. Альтернативная гипотеза. Статистический критерий. Критическая область. Общая схема проверки гипотезы. Риски при проверке гипотезы. Уровень значимости. Мощность. Надежность. Доверительные интервалы.

Тема 11. Проверка параметрических гипотез

Проверка гипотезы о числовом значении математического ожидания нормального закона распределения (с известным и неизвестным значением σ). Проверка гипотезы о числовом значении вероятности. Проверка гипотезы о числовом значении дисперсии нормального закона распределения. Проверка гипотезы о равенстве дисперсий двух генеральных совокупностей. Проверка гипотезы о равенстве математических ожиданий двух генеральных совокупностей. Проверка гипотезы о равенстве двух генеральных долей. Проверка гипотезы о равенстве нескольких дисперсий генеральных совокупностей по выборкам различного объема. Критерий Бартлетта. Проверка гипотезы о равенстве нескольких дисперсий генеральных совокупностей по выбор-

кам одинакового объема. Критерий Кочрена. Сравнение двух вероятностей биномиального распределения. Проверка гипотезы о значимости выборочного коэффициента корреляции Пирсона.

Тема 12. Проверка непараметрических гипотез

Критерии согласия. Критерии независимости качественных признаков. Таблицы сопряженности. Критерии сдвига. Проверка гипотезы о значимости выборочного ранговой корреляции Спирмена. Проверка гипотезы о значимости выборочного ранговой корреляции Кендалла. Проверка гипотезы об однородности двух выборок по критерию Вилкоксона. Проверка гипотезы о нормальном распределении генеральной совокупности по критерию Пирсона. Проверка гипотезы о показательном распределении генеральной совокупности. Проверка гипотезы о распределении генеральной совокупности по биномиальному распределению. Проверка гипотезы о равномерном распределении генеральной совокупности. Проверка гипотезы о распределении генеральной совокупности по закону Пуассона.

Тема 13. Дисперсионный анализ

Модель однофакторного дисперсионного анализа. Факторы. Результативный признак. Общая вариация. Межгрупповая вариация. Детерминированная модель двухфакторного дисперсионного анализа с повторениями.

Тема 14. Корреляционно-регрессионный анализ

Парная линейная регрессия. Взаимосвязи экономических переменных. Метод наименьших квадратов. Случайное отклонение в регрессионной модели. Спецификация модели. Выборочный коэффициент корреляции между переменными уравнения регрессии. Проверка качества уравнения регрессии. Коэффициент детерминации R^2 . Условия оптимальности метода наименьших квадратов (теорема Гаусса-Маркова). Интервальные оценки коэффициентов линейного уравнения регрессии. Доверительные интервалы для зависимой переменной. Нелинейные модели: гиперболическая зависимость на примере зависимости курса доллара от цены нефти; логарифмическая зависимость; зависимость степенного типа на примере производственной функции; логистическая регрессия. Гетероскедастичность. Автокорреляция. Мультиколлинеарность. Множественная линейная регрессия. Стандартизированная форма множественной регрессии. Проверка предпосылок метода наименьших квадратов. Статистика Дарбина-Уотсона.

Тема 15. Элементы теории временных рядов

Временной ряд, его характеристики. Корреляция временных рядов. Определение тренда временного ряда. Оценка периодических колебаний временного ряда.

5.2 Учебно-тематический план

Таблица 2

№ п/	Наименование темы (раздела)	Трудоемкость в часах			Формы текущего
		Все-	Контактная работа - Аудитор-	Само-	

п	дисциплины	го	ная работа			стоя- тельная работа	контроля успеваемо- сти
			Об- щая	Лек- ции	Семинары, практиче- ские заня- тия		
1	Данные в эконо- мике, их визуали- зация и предва- рительная обра- ботка	18	2	1	1	16	Выполнение практических заданий
2	Случайные собы- тия	24	4	2	2	20	Выполнение практических заданий Срезовая ра- бота
3	Случайные вели- чины	26	4	2	2	22	Выполнение практических заданий. Срезовая ра- бота
4	Центральные и начальные мо- менты, асиммет- рия, эксцесс	17	2	1	1	15	Выполнение практических заданий
5	Центральная пре- дельная теорема	17	2	1	1	15	Выполнение практических заданий
6	Случайные про- цессы	24	2	1	1	22	Выполнение практических заданий
	Итого 3 семестр	126	16	8	8	110	Согласно учебному плану: до- машнее твор- ческое зада- ние
7	Выборочный ме- тод	12	1	1	-	11	Выполнение практических заданий
8	Точечные оценки параметров рас- пределения	14	2	1	1	12	Выполнение практических заданий. Срезовая ра- бота
9	Интервальные оценки парамет- ров распределе- ния	14	2	1	1	12	Выполнение практических заданий
10	Статистические гипотезы	14	1	1	-	13	Выполнение практических заданий

11	Проверка параметрических гипотез	14	3	1	2	11	Выполнение практических заданий. Срезовая работа
12	Проверка непараметрических гипотез	14	2	1	1	12	Выполнение практических заданий
13	Дисперсионный анализ	14	2	1	1	12	Выполнение практических заданий
14	Корреляционно-регрессионный анализ	16	3	1	2	13	Выполнение практических заданий
15	Элементы теории временных рядов	14	2	-	2	12	Выполнение практических заданий
	Итого 4 семестр	126	18	8	10	108	Согласно учебному плану: расчетно-аналитическая работа
	В целом по дисциплине	252	34	16	18	218	Согласно учебному плану: расчетно-аналитическая работа
	Итого в %	13	6	7	87	13	-

5.3 Содержание семинаров, практических занятий

Таблица 3

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Данные в экономике, их визуализация и предварительная обработка	Данные в экономике. Форматирование наборов данных как таблиц. Гистограммы. Условное форматирование. Графики и диаграммы рассеяния. Инструменты описательной статистики. Визуализация качественных признаков. Предварительная обработка данных. Рекомендуемые источники: 8, 1-2	Выполнение заданий на компьютере
Случайные события	Основы комбинаторики. Определение вероятности. Классическая вероятностная схема. Схема геометрических вероятностей. Статистическая вероятность. Аксиоматическое построение теории вероятностей.	Решение задач

	Условные вероятности. Последовательности испытаний. Рекомендуемые источники: 8, 1-2	
Случайные величины	Определение случайной величины. Дискретные случайные величины и их важнейшие числовые характеристики. Дискретные случайные величины, часто встречающиеся в экономической практике. Абсолютно непрерывные случайные величины и их важнейшие числовые характеристики. Абсолютно непрерывные случайные величины, часто встречающиеся в экономической практике. Равномерный закон распределения. Показательный закон распределения. Нормальный закон распределения. Логарифмически нормальный закон распределения. Моменты и критические границы случайной величины. Меры связи случайных величин. Функции случайных величин. Рекомендуемые источники: 8, 1-2	Решение задач
Центральные и начальные моменты, асимметрия, эксцесс	Моменты случайной величины. Асимметрия и эксцесс случайной величины. Квантили и процентные точки случайной величины. Сумма под риском (VaR). Медиана и мода случайной величины. Рекомендуемые источники: 8, 1-2	Решение задач
Центральная предельная теорема	Закон больших чисел и центральная предельная теорема. Рекомендуемые источники: 8, 1-2	Решение задач
Случайные процессы	Определение и способы задания цепей Маркова. Вероятности и матрица переходов. Многошаговые вероятности переходов и теорема о матрице многошаговых переходов. Предельные вероятности. Теорема Маркова о предельных вероятностях. Моделирование случайных величин методом Монте-Карло. Рекомендуемые источники: 8, 1-2	Решение задач
Выборочный метод	Статистические методы обработки экспериментальных данных. Генеральная совокупность и выборка. Допустимый объем выборки для обеспечения ее репрезентативности. Вариационный ряд. Оценка функции распределения (эмпирическая функция распределения) и плотности распределения. Гистограмма. Рекомендуемые источники: 8, 1-2	Решение задач Выполнение заданий на компьютере
Точечные оценки параметров распределения	Понятие статистики, оценки числовой характеристики. Свойства точечных оценок: состоятельность, несмещенность и эффективность. Методы построения точечных оценок. Метод моментов. Метод максимального правдоподобия. Рекомендуемые источники: 8, 1-2	Решение задач Выполнение заданий на компьютере
Интервальные оценки параметров распределения	Однофакторный дисперсионный анализ. Двухфакторный дисперсионный анализ. Рекомендуемые источники: 8, 1-2	Решение задач Выполнение заданий на компьютере

Статистические гипотезы	Доверительный интервал для математического ожидания при известном значении дисперсии. Доверительный интервал для математического ожидания при неизвестном значении дисперсии. Доверительный интервал для вероятности. Доверительный интервал для дисперсии. Доверительный интервал для прогноза. Рекомендуемые источники: 8, 1-2	Решение задач Выполнение заданий на компьютере
Проверка параметрических гипотез	Проверка гипотезы о числовом значении математического ожидания нормального закона распределения (с известным и неизвестным значением σ). Проверка гипотезы о числовом значении вероятности. Проверка гипотезы о числовом значении дисперсии нормального закона распределения. Проверка гипотезы о равенстве дисперсий двух генеральных совокупностей. Проверка гипотезы о равенстве математических ожиданий двух генеральных совокупностей. Проверка гипотезы о равенстве двух генеральных долей. Проверка гипотезы о значимости выборочного коэффициента корреляции Пирсона. Рекомендуемые источники: 8, 1-2	Решение задач Выполнение заданий на компьютере
Проверка непараметрических гипотез	Критерии согласия. Критерии независимости качественных признаков. Таблицы сопряженности. Критерии сдвига. Проверка гипотезы о нормальном распределении генеральной совокупности по критерию Пирсона. Проверка гипотезы о показательном распределении генеральной совокупности. Проверка гипотезы о распределении генеральной совокупности по биномиальному распределению. Проверка гипотезы о равномерном распределении генеральной совокупности. Проверка гипотезы о распределении генеральной совокупности по закону Пуассона. Рекомендуемые источники: 8, 1-2	Решение задач
Дисперсионный анализ	Модель однофакторного дисперсионного анализа. Факторы. Результативный признак. Общая вариация. Межгрупповая вариация. Многофакторный дисперсионный анализ. Рекомендуемые источники: 8, 1-2	Выполнение заданий на компьютере
Корреляционно-регрессионный анализ	Парная линейная регрессия. Взаимосвязи экономических переменных. Метод наименьших квадратов. Случайное отклонение в регрессионной модели. Спецификация модели. Выборочный коэффициент корреляции между переменными уравнения регрессии. Проверка качества уравнения регрессии. Коэффициент детерминации R^2 . Условия оптимальности метода наименьших квадратов (теорема Гаусса-Маркова). Интервальные оценки коэффициентов линейного уравнения регрессии. Доверительные интервалы для зависимой переменной. Нелинейные модели: гиперболическая зависимость на примере зависимости курса доллара от цены нефти; логарифмическая зависимость; зависимость степенного типа на	Решение задач

	примере производственной функции; логистическая регрессия. Гетероскедастичность. Автокорреляция. Мультиколлинеарность. Множественная линейная регрессия. Стандартизированная форма множественной регрессии. Рекомендуемые источники: 8, 1-2	
Элементы теории временных рядов	Временной ряд, его характеристики. Корреляция временных рядов. Определение тренда временного ряда. Рекомендуемые источники: 8, 1-2	Выполнение заданий на компьютере

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 4

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Данные в экономике, их визуализация и предварительная обработка	Разработка макросов для первичной обработки данных	Работа с учебной литературой, решение задач. Работа за компьютером.
Случайные события	Свойства операций над событиями. Законы де Моргана. Парадокс Монти Холла. Бином Ньютона и свойства биномиальных коэффициентов.	Работа с учебной литературой, решение задач. Работа за компьютером.
Случайные величины	Выбор случайных величин, оптимальных по Парето. Биномиальная модель ценообразования финансовых инструментов	Работа с учебной литературой, решение задач. Работа за компьютером.
Центральные и начальные моменты, асимметрия, эксцесс	Квантили и процентные точки случайной величины	Работа с учебной литературой, решение задач. Работа за компьютером.
Центральная предельная теорема	Роль нормального закона в приложениях теории вероятностей. Теорема Леви.	Работа с учебной литературой, решение задач. Работа за компьютером.
Случайные процессы	Моделирование случайных величин.	Работа с учебной литературой, решение задач. Работа за компьютером.

Выборочный метод	Способы отбора элементов из генеральной совокупности в выборочную. Принципы выборочного наблюдения. Контроль статистической информации.	Работа с учебной литературой, решение задач. Работа за компьютером. Выполнение самостоятельной расчетно-аналитической работы.
Точечные оценки параметров распределения	Свойства выборочных моментов.	Работа с учебной литературой, решение задач. Работа за компьютером. Выполнение самостоятельной расчетно-аналитической работы.
Интервальные оценки параметров распределения	Односторонние доверительные интервалы.	Работа с учебной литературой, решение задач. Работа за компьютером.
Статистические гипотезы	Непараметрическая интервальная оценка теоретической величины сдвига. Критерий Вилкоксона — Манна — Уитни (непараметрический критерий сравнения математических ожиданий для независимых выборок)	Работа с учебной литературой, решение задач. Работа за компьютером.
Проверка параметрических гипотез	Проверка гипотезы о равенстве нескольких дисперсий генеральных совокупностей по выборкам различного объема. Критерий Бартлетта. Проверка гипотезы о равенстве нескольких дисперсий генеральных совокупностей по выборкам одинакового объема. Критерий Кочрена. Сравнение двух вероятностей биномиального распределения.	Работа с учебной литературой, решение задач. Работа за компьютером.
Проверка непараметрических гипотез	Проверка гипотезы о значимости выборочного ранговой корреляции Спирмена. Проверка гипотезы о значимости выборочного ранговой корреляции Кендалла. Проверка гипотезы об однородности двух выборок по критерию Вилкоксона.	Работа с учебной литературой, решение задач. Работа за компьютером.
Дисперсионный анализ	Детерминированная модель двухфакторного дисперсионного анализа с повторениями.	Работа с учебной литературой, решение задач. Работа за компьютером.
Корреляционно-регрессионный анализ	Проверка предпосылок метода наименьших квадратов. Статистика Дарбина-Уотсона.	Работа с учебной литературой, решение задач. Работа за компьютером.
Элементы теории временных рядов	Оценка периодических колебаний временного ряда.	Работа с учебной литературой, решение задач. Работа за компьютером.

6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю (согласно таблице 2)

Вопросы

для подготовки к текущему контролю по дисциплине

1. Виды данных – количественные, порядковые и номинальные.
2. Задачи анализа данных: классификация, группировка, прогнозирование, нахождение ассоциаций и зависимостей, визуализация.
3. Основные разделы, на которых базируется анализ данных: статистика, базы данных и знаний, распознавание образов, искусственный интеллект.
4. Классификация методов анализа данных. Этапы анализа данных: выявление закономерностей, прогнозирование, анализ исключений.
5. Сферы применения анализа данных: финансы и банковское дело, маркетинг, медицина, генетика, биоинформатика, интернет.
6. Наиболее важные законы распределения, их свойства.
7. Законы распределения: равномерное, нормальное (гауссово), Стьюдента, «хи-квадрат», экспоненциальное, Фишера.
8. Числовые характеристики случайных величин, характеристики центра группирования и вариации.
9. Теоретические моменты.
10. Основные определения и понятия выборочного метода. Повторные и бесповторные выборки. Первичный анализ данных, группировка.
11. Дискретные и интервальные вариационные ряды. Формула Стерджеса.
12. Генеральные характеристики: среднее, дисперсия, моменты высших порядков (асимметрия, эксцесс). Мода и Медиана.
13. Эмпирическая функция распределения, полигон и гистограмма.
14. Точечные оценки. Несмещенность, состоятельность и эффективность точечных оценок.
15. Оценивание параметров функции распределения.
16. Метод моментов. Метод максимального правдоподобия.
17. Интервальные оценки. Точность и надежность выборочных оценок.
18. Доверительная вероятность. Доверительный интервал.
19. Определение объема репрезентативной выборки для однородной и стратифицированной генеральной совокупности.
20. Ошибки первого и второго рода. Уровень значимости и мощность критерия.
21. Описание гипотез и критерии их проверки. Простые и сложные гипотезы.

22. Проверка гипотез о равенстве средних и дисперсий двух нормально распределенных генеральных совокупностей.
23. Хи-квадрат критерий Пирсона: проверка гипотезы о соответствии наблюдаемых значений предполагаемому распределению вероятностей (дискретному или непрерывному).
24. Проверка гипотез о вероятностной природе данных (стационарности, нормальности, независимости, однородности).
25. Функциональная и статистическая зависимости. Корреляционная таблица. Групповые средние. Понятие корреляционной зависимости.
26. Эмпирическая ковариация. Выборочный коэффициент корреляции, его свойства. Основные задачи теории корреляции: определение формы и оценка тесноты связи.
27. Виды корреляционной связи (парная и множественная, линейная и нелинейная).
28. Линейная корреляция. Уравнения прямых регрессии для парной корреляции.
29. Определение параметров прямых регрессии методом наименьших квадратов.
30. Значимость коэффициентов по критерию Стьюдента.
31. Однофакторный дисперсионный анализ с одинаковым числом испытаний на различных уровнях.
32. Однофакторный дисперсионный анализ с различным числом испытаний на различных уровнях.
33. Двух- и многофакторный дисперсионный анализ. Критерий адекватности Фишера.
34. Формирование частных и интегральных показателей результативности функционирования сложных систем

Примерные задания домашнего творческого задания

Соберите недельные данные о ценах закрытия и объемах торгов по пяти акциям, обращающимся на Московской бирже, с 01.01.2017 по сегодняшний день. Для каждой акции:

- рассчитайте недельные доходности и натуральные логарифмы объемов торгов;
- постройте и опишите гистограммы всех показателей;
- проверьте гипотезы о нормальном и логарифмически нормальном законе распределения всех показателей;
- постройте 95%-ный доверительный интервал для доходности
- проверьте гипотезу о том, что доходность по нему равна среднерыночной при альтернативной гипотезе о том, что доходность превышает среднерыночную;
- постройте модель прогнозирования цен закрытия, дав точечные и интервальные прогнозы.

Примерное задание расчетно-аналитической работы

Каждый студент выполняет индивидуальный вариант расчетно-аналитического задания.

1. В соответствии с вариантом, выданным преподавателем, соберите не-дельные данные о ценах закрытия и объемах торгов по трем акциям, обращающимся на Московской бирже, с 01.01.2015 г. по сегодняшний день. Названия акций в соответствии с вариантами приведены в файле «Варианты задания по анализу котировок акций.xlsx».

2. Удалите строки с пропущенными и нулевыми данными хотя бы в одном из столбцов.

3. Рассчитайте для каждой акции недельные доходности и логдоходности, а также натуральные логарифмы цен и объемов торгов.

4. Удалите строки с выбросами хотя бы в одном из столбцов.

5. Постройте гистограммы всех показателей и опишите их: являются ли они симметричными или скошенными, похожи ли они на «колокол» нормального распределения.

Укажите на графиках асимметрию и эксцесс, полученные в п. 7, а также результаты проверки гипотез о нормальности распределений, полученные в п. 18.

6. Проведите условное форматирование данных об акциях.

7. Постройте графики всех показателей от времени. Прослеживаются ли какие-нибудь зависимости данных признаков от времени?

С помощью методов условного форматирования определите даты, соответствующие аномальным значениям, обнаруженным на графиках, и выбросам. Попытайтесь найти объяснения этим аномалиям.

8. Постройте диаграммы рассеяния для каждой пары признаков. Между какими признаками можно предположить наличие зависимости?

9. С помощью программы «Описательная статистика» определите основные статистические показатели для данных об акциях. Постройте диаграммы Парето и определите акции, оптимальные по Парето.

10. Постройте диаграммы размаха по данным о ценах, объемах торгов и доходностях акций. Есть ли в данных выбросы? Чем можно их объяснить? Проанализируйте выбросы отдельно, исключив их из общего набора данных.

Дальнейший анализ необходимо проводить для данных с исключёнными выбросами.

11. Постройте интервальные вариационные ряды, гистограммы, полигоны частот и графики выборочных функций распределения для каждого из признаков.

12. На основе первых 7 лет найдите портфели акций с минимальным риском и с максимальной доходностью из пятёрки, 5 четвёрок и 10 троек акций. Выберите из них один с наименьшим риском и один с наименьшей доходностью.

13. Проследите траектории этих портфелей на основе дневных доходностей за последний год.

14. Сопоставьте результаты исследования портфелей с результатами анализа оптимальности по Парето.

15. В предположении нормального закона всех показателей замените для каждого из этих признаков параметры нормального закона μ и σ их несмещенными состоятельными оценками рассчитайте значения функции плотности нормального закона в серединах интервалов и функции распределения в правых концах интервалов. Для каждого из анализируемых признаков постройте на одном рисунке полигон частот и график теоретической функции плотности распределения, на другом – кумуляту и график теоретической функции распределения, а на третьем – эмпирическую и теоретическую функцию распределения.

16. Постройте 95%-ные доверительные интервалы для математических ожиданий и средних квадратических отклонений всех показателей.

17. Постройте 95%-ные доверительные интервалы для прогнозов на следующую неделю значений всех показателей.

18. Для каждой акции рассчитайте оценку 95% ной суммы под риском на ближайшую неделю для пакета акций, приобретаемого сегодня за 100 000 руб.

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

20. Для каждой акции проверьте на 5%-ном уровне значимости гипотезы о нормальном законе распределения каждого из признаков.

Примеры тестовых заданий для текущего контроля

1. Математическое ожидание случайной величины – это ...
2. Какие случайные величины являются дискретными:
 - а) рост студентов в группе
 - б) значение числа очков при бросании игрального кубика
 - в) выручка торговой организации за месяц
 - г) число работников организации

Примеры практических заданий для текущего контроля

1. Учебник издан тиражом 100 000 экземпляров. Вероятность того, что учебник сброшюрован неправильно, равна 0,0001. Найти вероятность того, что тираж содержит ровно пять бракованных книг.

2. Произведено 200 испытаний, в каждом из которых неизвестная вероятность p появления события A постоянна. Событие A появилось в 150 испытаниях. Найти доверительный интервал, покрывающий неизвестную вероятность p с надежностью 0,95.

Критерии балльной оценки различных форм текущего контроля успеваемости

Организация текущего контроля успеваемости обучающихся по итогам семестра (модуля), по программам бакалавриата и магистратуры в соответствии с учебными планами по направлениям подготовки высшего образования утверждена приказом ректора Финуниверситета по основной деятельности от 23.03.2017 №0557/о «Об утверждении Положения о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся по программам бакалавриата и магистратуры в Финансовом университете».

Таблица 5 - Виды работ обучающегося, формирующие текущий контроль по дисциплине

Вид учебной деятельности	Баллы	Максимум за семестр (модуль)
Выполнение и защита ДТЗ/РАР	10	10
Аудиторная контрольная/срезовая работа	5	10
Тестирование/блиц опросы	0,2	10
Посещение занятий, ведение конспекта лекции/семинара и работа с ним	0,1	4
Активное участие в интерактивном процессе выполнения практических заданий на семинарском занятии	0,1	6
Всего за семестр (модуль)		40

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

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

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

Примерные задания, необходимые для оценки достижения компетенций.

Компетенция УК-4 Способность использовать прикладное программное обеспечение при решении профессиональных задач.

Индикатор 1. Использует основные методы и средства получения, представления, хранения и обработки данных.

Знать: основные способы получения, отбора, хранения и обработки данных, способы доступа к различным баз с данными.

Уметь: использовать техническое программное обеспечение для получения, отбора, хранения и обработки данных; уметь обращаться к внешним ресурсам для получения данных.

Задание 1-3. Менеджер компании, занимающейся прокатом автомобилей, хочет оценить среднюю величину пробега одного автомобиля в течение месяца. Из 280 автомобилей, принадлежащих компании, по схеме собственно-случайной бесповторной выборки отобрано 60. Результаты представлены в таблице:

Пробег (км)	Менее 1000	1000 - 2000	2000 - 3000	3000 - 4000	4000 - 5000	5000 - 6000	Более 6000	Итого
Число автомобилей	3	5	9	16	13	8	6	60

Определить:

1) вероятность того, что средний пробег автомобиля в месяц отличается от среднего их пробега в выборке не более чем на 400 км (по абсолютной величине);

2) границы, в которых с вероятностью 0,95 заключена доля автомобилей, пробег которых составляет менее 3000 км;

3) объем бесповторной выборки, при котором те же границы для указанной доли можно гарантировать с вероятностью 0,9876; дать ответ на тот же вопрос, если никаких предварительных сведений о рассматриваемой доли нет

Индикатор 2. Демонстрирует владение профессиональными пакетами прикладных программ.

Знать: основные принципы работы персонального компьютера; функциональные настройки инструментальных программных средств.

Уметь: использовать пакеты прикладных программ общего назначения.

Задание 1-2. Вероятность наличия нужного покупателю товара в первом магазине равна 0,6, во втором – 0,7, в третьем – 0,8, в четвертом – 0,85. Покупатель в указанной последовательности посещает эти магазины до тех пор, пока не найдет нужный ему товар.

Составить закон распределения случайной величины X – числа магазинов, которые придется посетить покупателю.

Найти:

1) функцию распределения случайной величины X и построить ее график;

2) ее математическое ожидание и дисперсию

Задание 3. Вероятность поражения мишени стрелком равна 0,8.

Что вероятнее: поразить мишень семь раз при десяти выстрелах или 140 раз при двухстах выстрелах?

Задание 4-6. По схеме собственно-случайной бесповторной выборки проведено 10%-ное обследование строительных организаций региона по объему выполненных работ. Результаты представлены в таблице.

Объем работ (млн. руб.)	Менее 56	56–60	60–64	64–68	68–72	Более 72	Итого
Число организаций	9	11	19	30	18	13	100

Найти:

а) границы, в которых с вероятностью 0,9973 заключен средний объем выполненных работ всех строительных организаций региона;

б) вероятность того, что доля всех строительных организаций, объем работ которых составляет не менее 60 млн. руб., отличается от доли таких организаций в выборке не более чем на 0,05 (по абсолютной величине);

в) объем бесповторной выборки, при котором те же границы для среднего объема выполненных работ (см. п. а) можно гарантировать с вероятностью 0,9876.

Задание 7. Построить эмпирическое распределение рейтинга студентов по результатам экзаменов, оцененных в баллах для следующей произвольной выборки: 48, 51, 64, 62, 55, 71, 74, 79, 80, 86, 91, 99, 83, 50.

Индикатор 3. Выбирает необходимое прикладное программное обеспечение в зависимости от решаемой задачи.

Знать: основные современные средства и информационные технологии при решении управленческих задач.

Уметь: применять соответствующее программное обеспечение для моделирования ситуаций и их анализа.

Задание 1 Банк выдает 5 кредитов. Вероятность невозврата кредита равна 0,2 для каждого из заемщиков. Требуется составить таблицу закона распределения количества заемщиков, не вернувших кредит по окончании срока кредитования.

Задание 2. Магазин продает мужские костюмы. Распределение спроса по размерам является нормальным с математическим ожиданием $M=48$ и сигма=2. Вычислите при наличии данных процент спроса на 50 размер при условии разброса значений этой величины в интервале (49,51).

Задание 3. Одна из случайных величин (X) задана законом распределения:

$X:$	x_i	0	1	3
	p_i	0,2	0,3	0,5

а другая (Y) имеет биномиальное распределение с параметрами $n = 2$, $p = 0,4$.

Составить закон распределения их разности. Найти математическое ожидание и дисперсию этой случайной величины.

Индикатор 4. Использует прикладное программное обеспечение для решения конкретных прикладных задач.

Знать: современный инструментарий обработки реальных данных.

Уметь: применять информационные инструменты к анализу результатов исследования реальных данных.

Задание 1-9. Предложенные экономические задачи решите с использованием табличного процессора MS Excel.

По предприятиям легкой промышленности региона получена информация, характеризующая зависимость объема выпуска продукции (Y , млн руб.) от объема капиталовложений (X , млн руб.).

Требуется:

1. Найти параметры уравнения линейной регрессии, дать экономическую интерпретацию коэффициента регрессии.
 2. Вычислить остатки; найти остаточную сумму квадратов; оценить дисперсию остатков; построить график остатков.
 3. Проверить выполнение предпосылок МНК.
 4. Осуществить проверку значимости параметров уравнения регрессии с помощью t -критерия Стьюдента ($\alpha = 0,05$).
 5. Вычислить коэффициент детерминации, проверить значимость уравнения регрессии с помощью F -критерия Фишера ($\alpha = 0,05$), найти среднюю относительную ошибку аппроксимации. Сделать вывод о качестве модели.
 6. Осуществить прогнозирование среднего значения показателя Y при уровне значимости $\alpha = 0,1$, если прогнозное значение фактора X составит 80% от его максимального значения.
 7. Представить графически фактические и модельные значения Y точки прогноза.
 8. Составить уравнения нелинейной регрессии:
 - гиперболической;
 - степенной;
 - показательной.
- Привести графики построенных уравнений регрессии.
9. Для указанных моделей найти коэффициенты детерминации, коэффициенты эластичности и средние относительные ошибки аппроксимации. Сравнить модели по этим характеристикам и сделать вывод.

X	66	58	73	82	81	84	55	67	81	59
Y	133	107	145	162	163	170	104	132	159	116

Компетенция УК-15 Способность релевантно решаемым задачам использовать информационные ресурсы и информационно-коммуникационные технологии для достижения целей, связанных с профессиональной деятельностью, обучением, участием в жизни общества и других сферах жизни.

Индикатор 1. Самостоятельно выбирает и использует цифровые средства общения, осуществляет поиск и/или создание контента в соответствии с целью взаимодействия, в том числе для организации совместной деятельности.

Знать: возможности информационных ресурсов для решения задач анализа данных.

Уметь: использовать информационные ресурсы для решения задач анализа данных.

Задание 1.

Аналитическая группа проводит анализ предлагаемых бизнес-планов различными инвесторами с целью принятия решений по инвестиционным проектам. Практические исследования показали, что в прошлые периоды 3% всех предлагаемых проектов являются неподходящими для инвестирования. Предложенная схема анализа определяет 95% «неподходящих» проектов как «действительно неподходящие», но при этом 20% проектов, пригодных для инвестиций, также определяет как «действительно неподходящие».

Найдите: 1) вероятность того, что проект не подходит для инвестирования, при условии, что после аналитического исследования всех бизнес-планов он был определен как «действительно подходящий»; 2) вероятность того, что проект подходит для инвестирования, при условии, что аналитикой он был определен как «действительно неподходящий». На основе полученных результатов прокомментируйте пригодность предлагаемой схемы анализа для принятия инвестиционных решений. Задание выполнить с помощью электронной таблицы Excel.

Индикатор 2. Владеет навыками организации взаимодействия и коммуникации с помощью информационных систем и/или цифровых сервисов и технологий.

Знать: возможности цифровых сервисов и технологий для взаимодействия при решении задач анализа данных.

Уметь: использовать цифровые сервисы и технологии для взаимодействия при решении задач анализа данных.

Задание 1.

Из 3000 выданных некоторым банком кредитов 570 не были возвращены в срок. Постройте 95%-ный доверительный интервал для вероятности невозврата кредита. С помощью 95%-ного доверительного интервала для вероятности невозврата кредита, проверьте на 5%-ном уровне значимости гипотезу о том, что вероятность невозврата кредита равна 0,2 (при альтернативе,

что она не равна 0,2). Задание выполнить с помощью электронной таблицы Excel.

Индикатор 3. Осуществляет подбор и применение различных информационно-коммуникационных средств для решения образовательных и профессиональных задач.

Знать: различные информационно-коммуникационные средства для решения задач анализа данных.

Уметь: использовать различные информационно-коммуникационные средства для решения задач анализа данных.

Задание 1.

В консалтинговом агентстве работает 8 сотрудников высшей категории, 5 сотрудников - первой категории, 9 - второй категории. Вероятности ошибочных заключений, выданных сотрудниками агентства, зависят от их квалификации и равны: 0,01 - для высшей категории, 0,05 - для первой категории и 0,07 – для второй категории. Найти вероятность того, что при обращении в данное агентство к случайно выбранному сотруднику клиенту будет выдано верное заключение. Найти вероятность того, что клиента обслуживал сотрудник первой категории. Задание выполнить с помощью электронной таблицы Excel.

Компетенция ПКН-10 Владение методами количественного и качественного анализа информации, а также навыками построения моделей, применяя для анализа, моделирования и поддержки принятия решений современные информационные технологии и программные средства, включая инструменты бизнес-аналитики, обработки и анализа данных.

Индикатор 1. Использует методы получения информации, ее анализа для построения моделей и интерпретации результатов моделирования.

Знать: методы поиска информации, ее системного и критического анализа.

Уметь: применять современные информационные технологии для сбора информации, использовать для построения адекватных математических моделей.

Задание 1. Имеются следующие данные о цене на нефть x (ден. ед.) и индексе акций нефтяных компаний y (усл. ед.).

x	17,28	17,05	18,30	18,80	19,20	18,50
y	537	534	550	555	560	552

Предполагая, что между переменными x и y существует линейная зависимость, найти эмпирическую формулу вида $y = ax + b$, используя метод наименьших квадратов.

Задание 2-4. Распределение 100 средних фермерских хозяйств по числу наемных рабочих X (чел.) и их среднемесячной заработной плате на одного человека Y (тыс. руб.) представлено в таблице.

$\begin{matrix} y \\ x \end{matrix}$	10–20	20–30	30–40	40–50	50–60	Свыше 60	Итого
102						10	10
103					6	15	21
104			10	11	8		29
105			8	3			11
106		5	6				11
107	5	9	4				13
Итого:	5	14	28	14	14	25	100

Необходимо:

1. Вычислить групповые средние, построить эмпирические линии регрессии.

2. Предполагая, что между переменными X и Y существует линейная корреляционная зависимость:

а) найти уравнения прямых регрессии, построить их графики на одном чертеже с эмпирическими линиями регрессии и дать экономическую интерпретацию полученных уравнений;

б) вычислить коэффициент корреляции; на уровне значимости $\alpha = 0,05$ оценить его значимость и сделать вывод о тесноте и направлении связи между переменными X и Y ;

в) используя соответствующее уравнение регрессии, оценить среднемесячную заработную плату одного рабочего фермерского хозяйства, в котором работает 10 наемных рабочих.

Задание 5. Дана выборка стоимости валюты: 27,70; 27,85; 28,12; 28,20; 28,10; 27,75; 28,25 (рублей). Необходимо определить границы 95% доверительного интервала для среднего.

Индикатор 2. Применяет приемы квалификации и выбора подходящих измерительных шкал при описании организационных систем, происходящих в них процессов и явлений.

Знать: существующие измерительные шкалы, виды данных.

Уметь: выбрать и применить подходящие измерительные шкалы, исследовать организационные системы.

Задание 1. Найти границы 90% интервала для среднего значения, если по результатам 24 торгов среднее значение стоимости доллара составило 28 рублей, а стандартное отклонение – 35 копеек.

Задание 2. Проверить соответствие выборочных данных результатов сдачи экзаменов, оцененных в баллах: 48, 51, 67, 70, 64, 71, 85, 79, 80, 83, 86, 91, 99, 56, 66, 65, 84, 84, 84, 75, 76, 77, 78, 80, 86, 88, 58, 69, 65, 81, 75, 78, 85,

80, 80, 83, 86, 80, 89, 60, 68, 55, 82, 64, 71, 72, 72, 73, 74, 74, 79 нормальному закону распределения.

Индикатор 3. Использует навыки организации и проведения качественных и количественных исследований анализа информации, подготовки аналитических отчетов о состоянии и динамики развития рынков товаров и услуг.

Знать: методы и программы исследования, информационный инструментарий для анализа данных.

Уметь: обосновать сущность происходящего, выявить закономерности, понять природу вариабельности, использовать инструменты анализа данных, исследовать динамику развития рынка, уметь составлять аналитический отчет.

Задание 1-3. Для планирования бюджета предприятия на следующий год было проведено выборочное обследование использования амортизационного фонда. По схеме собственно-случайной бесповторной выборки из 500 выплат были отобраны 100 и получены следующие данные.

Величина выплаты (руб.)	Менее 1000	1000–2000	2000–3000	3000–4000	4000–5000	5000–6000	Итого
Число выплат	3	13	33	26	17	8	100

Найти:

- 1) вероятность того, что средняя выплата отличается от средней выплаты в выборке не более чем на 100 руб.;
- 2) границы, в которых с вероятностью 0,9281 заключена доля всех выплат, величина которых не превышает 4000 руб.;
- 3) объем бесповторной выборки, при котором те же границы для доли можно гарантировать с вероятностью 0,9545.

Примерный перечень вопросов для подготовки к зачету

1. Данные в экономике. Объекты, признаки и таблицы. Типы признаков в экономике и управлении: интервальные, порядковые, ранговые, дихотомические.
2. Инструменты описательной статистики.
3. Визуализация качественных признаков.
4. Предварительная обработка данных.
5. Основы комбинаторики. Правила суммы и произведения.
6. Случайные события, их виды. Операции над событиями как операции над множествами.

7. Условные вероятности. Условная вероятность. Независимость событий. Формула полной вероятности. Формула Байеса. Простейшие примеры применения теории вероятностей в экономике, управлении и финансах.
8. Последовательности испытаний. Биномиальная схема. Формула Бернулли.
9. Понятие случайной величины. Функция распределения случайной величины. Свойства функции распределения.
10. Дискретные случайные величины и их важнейшие числовые характеристики.
11. Биномиальный закон распределения.
12. Биномиальная модель ценообразования финансовых инструментов.
13. Геометрический закон распределения.
14. Закон распределения Пуассона.
15. Простейший поток событий.
16. Гипергеометрический закон распределения.
17. Сравнение случайных величин: отношение предпочтения, ожидаемая полезность, оптимальность по Парето.
18. Абсолютно непрерывные случайные величины и их важнейшие числовые характеристики.
19. Равномерный закон распределения.
20. Показательный закон распределения.
21. Нормальный закон распределения.
22. Логарифмически нормальный закон распределения и ценообразование финансовых инструментов.
23. Закон распределения Парето и задачи налогообложения.
24. Законы распределения, важные в математической статистике (законы распределения Стьюдента, χ^2 , Фишера — Снедекора).
25. Смеси распределений.
26. Начальные и центральные моменты случайной величины. Асимметрия и эксцесс случайной величины.
27. Квантили и процентные точки случайной величины.
28. Ценность под риском.
29. Медиана и мода случайной величины.
30. Случайные векторы и условные законы распределения.
31. Условное математическое ожидание. Формула полного математического ожидания. Формула полной дисперсии.
32. Ковариация и коэффициент корреляции.
33. Портфель финансовых инструментов
34. Функции случайных величин.
35. Закон больших чисел. Массовые случайные явления в экономике.
36. Центральная предельная теорема. Место центральной предельной теоремы в изучении статистических закономерностей в экономике, финансах и управлении.
37. Математические основы теории страхования.

38. Метод Монте-Карло. Моделирование случайных величин.

Примерный перечень вопросов для подготовки к экзамену

1. Предмет и задачи математической статистики. Генеральная и выборочная совокупности.
2. Соотношение между предельной ошибкой выборки, уровнем значимости (риском) и объемом выборки. Использование этого соотношения в организации выборочных обследований.
3. Оценка плотности распределения и функции распределения.
4. Понятие точечной оценки параметра генеральной совокупности. Свойства точечных оценок: состоятельность, несмещенность, эффективность.
5. Выборочное среднее как состоятельная, несмещенная и эффективная оценка математического ожидания генеральной случайной величины.
6. Смещенность выборочной дисперсии как оценки дисперсии генеральной случайной величины. Исправленная выборочная дисперсия как несмещенная и состоятельная оценка дисперсии генеральной случайной величины.
7. Методы построения точечных оценок: метод моментов, метод максимального правдоподобия.
8. Понятие интервальной оценки параметра генеральной совокупности.
9. Понятие статистической гипотезы. Виды статистических гипотез: параметрические и непараметрические, простые и сложные.
10. Проверка гипотезы о равенстве математического ожидания теоретическому значению. Проверка гипотезы о равенстве двух математических ожиданий.
11. Проверка гипотезы о равенстве дисперсии теоретическому значению. Проверка гипотезы о равенстве двух дисперсий.
12. Проверка гипотезы о равенстве вероятности события теоретическому значению. Проверка гипотезы о равенстве двух вероятностей.
13. Проверка гипотез о значимости коэффициента корреляции.
14. Критерии согласия. Критерий согласия хи-квадрат Пирсона. Критерий χ^2 Пирсона при неизвестных параметрах распределения.
15. Однофакторный дисперсионный анализ.
16. Двухфакторный дисперсионный анализ.
17. Таблицы сопряженности. Критерий хи-квадрат для проверки независимости компонент случайной величины. Критерий хи-квадрат для проверки однородности данных.
18. Непараметрические критерии. Примеры применения непараметрических критериев в экономике.
19. Ранговая корреляция. Примеры использования ранговой корреляции в экономике.
20. Задачи машинного обучения. Примеры задач машинного обучения в экономике, управлении и финансах.

21. Постановка задачи регрессионного анализа. Примеры задач регрессии в экономике.

22. Классификация с обучением. Постановка задачи классификации с обучением. Логистическая регрессия. Понятие о деревьях решений. Кредитный скоринг.

23. Кластерный анализ и поиск аномалий. Постановка задачи кластерного анализа. Метод К-средних. Сегментирование потребителей. Понятие о методах машинного обучения в задачах поиска аномалий.

Типовые задачи для подготовки к экзамену

1. Устройство состоит из пяти элементов, из которых два изношены. При включении устройства включаются случайным образом два элемента. Найти вероятность того, что включенными окажутся неизношенные элементы.

2. Для сигнализации об аварии установлены два независимо работающих сигнализатора. Вероятность того, что при аварии сигнализатор сработает, равна 0,95 для первого сигнализатора и 0,9 для второго. Найти вероятность того, что при аварии сработает только один сигнализатор.

3. Устройство состоит из трех независимо работающих основных элементов. Устройство отказывает, если откажет хотя бы один элемент. Вероятность отказа каждого элемента за время t равна 0,1. Найти вероятность безотказной работы устройства за время t , если: а) работают только основные элементы; б) включен один резервный элемент; в) включены два резервных элемента. Предполагается, что резервные элементы работают в том же режиме, что и основные, вероятность отказа каждого резервного элемента также равна 0,1 и устройство отказывает, если работает менее трех элементов.

4. Найти вероятность того, что событие A наступит ровно 70 раз в 243 испытаниях, если вероятность появления этого события в каждом испытании равна 0,25.

5. Отдел технического контроля проверяет на стандартность 900 деталей. Вероятность того, что деталь стандартна, равна 0,9. Найти с вероятностью 0,95 границы, в которых будет заключено число m стандартных деталей среди проверенных.

6. Устройство состоит из трех независимо работающих элементов. Вероятность отказа каждого элемента в одном опыте равна 0,1. Составить закон распределения числа отказавших элементов в одном опыте.

7. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение дискретной случайной величины X , заданной законом распределения:

а)

X	-4	6	10
P	0,2	0,3	0,5

б)

X	0,21	0,54	0,61
P	0,1	0,5	0,4

8. Случайная величина X задана функцией распределения

$$F(X) = \begin{cases} 0, & \text{при } x \leq -1, \\ \frac{3}{4} \cdot x + \frac{3}{4}, & \text{при } -1 < x \leq \frac{1}{3}, \\ 1, & \text{при } x > \frac{1}{3}. \end{cases}$$

9. Случайная величина X задана плотностью распределения $f(x) = 2x$ в интервале $(0,1)$; вне этого интервала $f(x) = 0$. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение величины X .

10. Математическое ожидание и среднее квадратическое отклонение нормально распределенной случайной величины X соответственно равны 20 и 5. Найти вероятность того, что в результате испытания X примет значение, заключенное в интервале $(15, 25)$.

11. По двум независимым выборкам, объемы которых $N_1=10$ и $N_2=15$, извлеченным из нормальных генеральных совокупностей X и Y , найдены исправленные выборочные дисперсии $S_x = 24,2$ и $S_y = 18,1$. При уровне значимости 0,01, проверить нулевую гипотезу $H_0: D(X) = D(Y)$ о равенстве генеральных дисперсий при конкурирующей гипотезе $H_1: D(X) > D(Y)$.

12. По выборке объема $n=80$, извлеченной из нормальной двумерной генеральной совокупности (X, Y) , найден выборочный коэффициент корреляции $r_b = 0,25$. Требуется при уровне значимости 0,01 проверить нулевую гипотезу о равенстве нулю генерального коэффициента корреляции при конкурирующей гипотезе $H_1: r_b \neq 0$.

13. Из нормальной генеральной совокупности извлечена выборка объема $n=18$ и по ней найдена исправленная выборочная дисперсия $s^2 = 0,36$. Требуется при уровне значимости 0,05 проверить нулевую гипотезу $H_0: \sigma^2 = \sigma_0^2 = 0,28$, приняв в качестве конкурирующей гипотезы $H_1: \sigma^2 > 0,28$.

14. По выборке объема $n=200$, извлеченной из нормальной двумерной генеральной совокупности (X, Y) , найден выборочный коэффициент корреляции $r_b = 0,3$. Требуется при уровне значимости 0,05 проверить нулевую гипотезу о равенстве нулю генерального коэффициента корреляции при конкурирующей гипотезе $H_1: r_b \neq 0$.

15. По результатам изучения 8 сельскохозяйственных предприятий получены следующие данные об объемах прибыли в тыс. руб.: 70; 65; 70; 80; 90; 75; 85; 80.

Найти выборочную среднюю и исправленную дисперсию.

16. По выборке объема $n=50$ найден средний вес $\bar{x} = 200$ г изделий, изготовленных на первом станке; по выборке объема $m=40$ найден средний вес $\bar{y} = 180$ г изделий, изготовленных на втором станке. Генеральные дисперсии известны: $D(X) = 80 \text{ г}^2$, $D(Y) = 100 \text{ г}^2$. Требуется, при уровне значимости 0,05, проверить нулевую гипотезу $H_0: M(X) = M(Y)$ при конкурирующей гипотезе $H_1: M(X) \neq M(Y)$. Предполагается, что случайные величины X и Y распределены нормально и выборки независимы.

17. Найти доверительный интервал для оценки с надежностью 0,99 неизвестного математического ожидания μ нормально распределенного признака X генеральной совокупности, если известны генеральное среднее квадратическое отклонение σ , выборочная средняя $\bar{x}_v$ и объем выборки n : $\sigma = 6$, $\bar{x}_v = 18,2$, $n = 30$.

18. Из нормальной генеральной совокупности с известным средним квадратическим отклонением $\sigma=30$ извлечена выборка объема $n = 60$ и по ней найдена выборочная средняя $\bar{x}=126,5$. Требуется при уровне значимости 0,01 проверить нулевую гипотезу $H_0: \mu = \mu_0 = 120$ при конкурирующей гипотезе $H_1: \mu \neq 120$.

19. Найти минимальный объем выборки, при котором с надежностью 0,975 точность оценки математического ожидания нормально распределенной генеральной совокупности по выборочной средней равна 0,6, если известно среднее квадратическое отклонение генеральной совокупности $\sigma = 2$.

20. Установлено, что средний вес таблетки лекарства сильного действия должен быть равен $\mu_0=0,50$ мг. Выборочная проверка 132 таблетки полученной партии лекарства показала, что средний вес таблетки этой партии $\bar{x}=0,54$ мг. Требуется при уровне значимости 0,01 проверить нулевую гипотезу $H_0: \mu = \mu_0 = 0,50$ при конкурирующей гипотезе $H_1: \mu > 0,50$. Многократными предварительными опытами по взвешиванию таблеток, поставляемых фармацевтическим заводом, было установлено, что вес таблеток распределен нормально со средним квадратическим отклонением $\sigma=0,12$ мг.

Пример экзаменационного билета по дисциплине

Федеральное государственное образовательное бюджетное учреждение высшего образования

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

Кафедра «Математика и информатика»

Дисциплина «Анализ данных»

Филиал Тульский

Семестр: 4 Направление: Менеджмент

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1

1. Определите схему последовательных испытаний. Какая формула используется для расчета вероятности в этом случае? (10 баллов)
2. Охарактеризовать виды статистических гипотез. (10 баллов)

Практикоориентированные задания

3. Вероятность наличия нужного покупателю товара в первом магазине равна 0,6, во втором – 0,7, в третьем – 0,8, в четвертом – 0,85. Покупатель в указанной последовательности посещает эти магазины до тех пор, пока не найдет нужный ему товар.

Составить закон распределения случайной величины X – числа магазинов, которые придется посетить покупателю.

Найти:

- 1) функцию распределения случайной величины X и построить ее график;
- 2) ее математическое ожидание и дисперсию. (20 баллов)
4. По выборке объема $n=200$, извлеченной из нормальной двумерной генеральной совокупности (X, Y) , найден выборочный коэффициент корреляции $r_b=0,3$. Требуется при уровне значимости 0,05 проверить нулевую гипотезу о равенстве нулю генерального коэффициента корреляции при конкурирующей гипотезе $H_1: r_b \neq 0$. (20 баллов)

7.2. Иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Порядок проведения промежуточной аттестации и система оценивания результатов промежуточной аттестации по итогам семестра (модуля) по программам бакалавриата и магистратуры в соответствии с учебными планами по направлениям подготовки высшего образования утверждена приказом ректора Финуниверситета по основной деятельности № 0557/о от 23.03.2017.

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

Основная

1. Соловьев В.И. Анализ данных в экономике: Теория вероятностей, прикладная статистика, обработка и анализ данных в Microsoft Excel [Электронный ресурс]: учебник. М.: КноРус, 2023. 497 с. ISBN 978-5-406-10701-0. <https://book.ru/book/946789> (дата обращения: 12.03.2024).

2. Кремер Н. Ш. Теория вероятностей и математическая статистика [Электронный ресурс]: учебник и практикум. 5-е изд., перераб. и доп. М.: Юрайт, 2024. 538 с. <https://urait.ru/bcode/541918> (дата обращения: 12.03.2024).

Дополнительная

3. Козлов А.Ю., Мхитарян В.С., Шишов В.Ф. Статистический анализ данных в MS Excel [Электронный ресурс]: учеб. пособие. М.: ИНФРА-М, 2023. 320 с. ISBN 978-5-16-004579-5. <https://znanium.com/catalog/product/1907518> (дата обращения: 12.03.2024).

4. Миркин Б. Г. Введение в анализ данных [Электронный ресурс]: учебник и практикум. М.: Юрайт, 2024. 174 с. <https://urait.ru/bcode/536117> (дата обращения: 12.03.2024).

5. Борисова Л. Р., Седых И. Ю., Хрипунова М. Б. Математика и анализ данных с поддержкой MS Excel и языка R. Практикум. Часть 1 [Электронный ресурс]: учебное пособие / под ред. И. Ю. Седых. М.: Прометей, 2024. 374 с. <https://znanium.ru/catalog/product/2144355> (дата обращения: 05.04.2024).

6. Борисова Л. Р., Светлова Н. А., Седых И. Ю. Математика и анализ данных с поддержкой MS Excel и языка R [Электронный ресурс]: учебное пособие / под ред. И. Ю. Седых. М.: Прометей, 2023. 728 с. <https://znanium.ru/catalog/product/2124863> (дата обращения: 05.04.2024).

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

1. <http://rts.micex.ru/>
2. <http://www.sks.ru/>
3. <http://www.cbr.ru/>
4. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>
5. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.biblio-online.ru/>

10. Методические указания для обучающихся по освоению дисциплины

Основная цель освоения содержания дисциплины получить теоретические знания в ходе лекционных занятий и семинаров, закрепить их в процессе систематической самостоятельной работы и сформировать знания, умения и владения в соответствии с компетенциями, содержащимися в рабочей программе по дисциплине.

Самостоятельная работа студента в процессе освоения дисциплины включает в себя:

- подготовку к семинарским занятиям;
- работу с электронными учебными ресурсами;
- изучение основной и дополнительной литературы по дисциплине;
- выполнение домашнего творческого задания, расчетно-аналитической работы;
- индивидуальные или групповые консультации по наиболее сложным вопросам;
- подготовка к экзамену.

Студенты, не выполнившие ДТЗ, РАР или не прошедшие собеседование по ней, к экзамену по дисциплине не допускаются.

Организация выполнения расчетно-аналитической работы включает несколько этапов. Прежде всего, студент должен изучить рекомендуемую учебную литературу и внимательно ознакомиться с методическими указаниями по выполнению работы.

Выполненная и надлежащим образом оформленная работа сдается на кафедру «Математика и информатика» в сроки, установленные учебным графиком. Если студент не допущен к защите, то работа должна быть доработана согласно замечаниям руководителя.

Критерии оценивания ДТЗ/РАР:

1) четкость структуры работы, представленной в содержании (рубрики дают однозначное понимание раскрываемой тематики и оформлены стилями заголовков соответствующего уровня);

2) наличие всех разделов, предложенных в методических указаниях, таблиц, иллюстраций (схем, рисунков и т.п.), ссылок на источники заимствованного материала;

3) наличие списка использованной литературы, оформленного в соответствии с правилами библиографического описания;

4) соответствие оформления расчетно-аналитической работы общим требованиям.

Расчетно-аналитическая работа оформляется в тетради или на компьютере с использованием текстового редактора:

- размер шрифта – 14;
- межстрочный интервал – полуторный;
- размеры полей: левого – 3 см, верхнего – 1,5, правого – 1, нижнего – 1,5 см;
- ориентация – книжная;
- форматирование основного текста и ссылок – по ширине;
- цвет шрифта – черный;
- абзацный отступ – 1,5 см.
- нумерация страниц – справа, внизу.

Каждую структурную часть расчетно-аналитической работы нужно начинать с нового листа. Точка в конце заголовка структурной части работы не ставится.

Необходимо стремиться к ясности, краткости и самостоятельности изложения материала.

Цитаты и заимствованный материал должны сопровождаться ссылками на источники, описание которых необходимо привести в списке использованной литературы (например: [2, с. 15]).

Сокращение слов и использование аббревиатур, за исключением общепринятых, в контрольной работе не допускаются. Общепринятые аббревиатуры необходимо расшифровать в тексте контрольной работы при их первом употреблении.

Цифровой материал целесообразно представлять в виде таблиц. Для этого в правом верхнем углу помещают надпись «Таблица» и указывают ее

порядковый номер. Таблицу снабжают тематическим заголовком, который располагают посередине страницы.

Приводимые в работе иллюстрации (схемы, диаграммы, графики, технические рисунки) должны быть выполнены четко, аккуратно и разборчиво. Рисунок должен иметь номер и подрисуночную подпись (например: Рис. 4. График функции $f(x)$).

На все таблицы и рисунки необходимо сделать ссылки в тексте работы.

На последней странице студент должен поставить подпись и дату сдачи работы на рецензирование.

Расчетно-аналитическая работа представляется на рецензирование в сброшюрованном виде (листы должны быть скреплены по левому краю).

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

11.1 Комплект лицензионного программного обеспечения

лицензионное отечественное:

1. Антивирусная защита Kaspersky Endpoint Security.
2. Astra Linux.
3. LibreOffice.

лицензионное импортное:

ПО для ТСО лиц с ограниченными возможностями здоровья
свободно распространяемое:
Архиватор KDE (Ark)

11.2 Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Гарант»
2. Информационно-правовая система «Консультант Плюс»
3. Информационно-образовательный портал Финансового университета. - URL: <http://www.fa.ru>. Доступ авторизованный.

11.3 Сертифицированные программные и аппаратные средства защиты информации

Не используются.

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

Образовательный процесс по учебной дисциплине осуществляется в учебных аудиториях для проведения учебных занятий. При освоении дисциплины используются технические средства мультимедийной техники аудиторий. Для проведения учебных занятий используются аудитории, оборудованные следующими техническими средствами: мультимедиа-проектором, экраном настенным или доской интерактивной с встроенной акустической

системой для воспроизведения аудио файла, персональным (-ными) компьютером (-рами) с доступом к Internet-ресурсам и электронно-образовательной среде Финансового университета.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения в сети «Интернет» и обеспечением доступа в электронно-образовательную среду Финансового университета.