МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ,
МИНИСТЕРСТВО НАУКИ, ВЫСШЕГО ОБРАЗОВАНИЯ И ИННОВАЦИЙ
КЫРГЫЗСКОЙ РЕСПУБЛИКИ

МОО ВО Кыргызско-Российский Славянский университет
имени первого Президента Российской Федерации Б.Н. Ельцина



УТВЕРЖДАЮ
директор Политехнического института
_____ Д.А.Коцев
3 сентября 2026 г.

Статистическая обработка данных

рабочая программа дисциплины (модуля)

Закреплена за **Высшая школа Цифровых технологий**

Учебный план g090404_26_1 пи_оаирп.plx
Направление подготовки 09.04.04 - РФ, 710400 - КР Программная инженерия
Магистерская программа "Основы анализа и разработки приложений с большими объемами распределенных данных"

Квалификация **магистр**

Форма обучения **очная**

Общая трудоемкость **3 ЗЕТ**

Часов по учебному 108

в том числе:

аудиторные занятия 38

самостоятельная 69,9

Виды контроля в семестрах:
зачет с оценкой 2

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	**2 (1.2)**		Итого	
Недель	18			
Вид занятий	УП	РП	УП	РП
Лекции	16	16	16	16
Практические	22	22	22	22
Контактная работа в период теоретического	0,1	0,1	0,1	0,1
В том числе инт.	8	8	8	8
В том числе в форме практ.подготовки	22	22	22	22
Итого ауд.	38	38	38	38
Контактная работа	38,1	38,1	38,1	38,1
Сам. работа	69,9	69,9	69,9	69,9
Итого	108	108	108	108

Программу составил(и):
*д.т.н., профессор , Лыченко Н.М., к.т.н., доцент, Осмонов М.С.*_____

Рецензент(ы):
д.т.н., доцент, Верзунов С.Н. _____

Согласовано РОП*: д.т.н., проф. Лыченко Н.М.*
_____

Рабочая программа дисциплины
Статистическая обработка данных

разработана в соответствии с ФГОС 3++:
Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 09.04.04 Программная инженерия (приказ Минобрнауки России от 19.09.2017 г. № 932)

составлена на основании учебного плана:

Направление подготовки 09.04.04 - РФ, 710400 - КР Программная инженерия
Магистерская программа "Основы анализа и разработки приложений с большими объемами распределенных данных"
утвержденного учёным советом вуза от 26.06.2026 протокол № 11.

Рабочая программа одобрена на заседании ВШ Цифровых технологий

Протокол от 02.09. 2026 г. № 1
Срок действия программы: 2026-2028 уч.г.
Руководитель подразделения: к.т.н., доцент Хмелева И.В.

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ
к рабочей программе дисциплины (модуля)

Статистическая обработка данных

_____ *наименование дисциплины*

№ п/п	Основание внесения изменения	Раздел/ пункт	Содержание изменения
1.			
2.			
3.			
4.			
5.			

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	
1.1	формирование у магистрантов углубленных компетенций в области сбора, математической систематизации, моделирования и интерпретации эмпирических данных с использованием современного программного обеспечения за счет:
1.2	- освоения теоретических основ теории вероятностей и математической статистики.
1.3	- развития навыков планирования статистического наблюдения и формирования выборочных совокупностей.
1.4	- обучения применению методов проверки статистических гипотез, корреляционного, регрессионного и дисперсионного анализа.
1.5	- освоения специализированных программных пакетов для автоматизации расчетов, в частности, RStudio.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП	
Цикл (раздел) ООП:	Б1.В
2.1	**Требования к предварительной подготовке обучающегося:**
2.1.1	Теория вероятностей и математическая статистика (бакалавриат)
2.1.2	Технология программирования (бакалавриат)
2.1.3	Б1.В.03. Наука о данных и аналитика больших объемов информации
2.2	**Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:**
2.2.1	Б1.В.05. Машинное обучение
2.2.2	Б2.В.01(У). Ознакомительная практика
2.2.3	Б1.О.03. Системы анализа больших данных
2.2.4	Б2.О.03(П). Научно-исследовательская работа (рассредоточенная)
2.2.5	Б2.О.02(Н). Научно-исследовательская работа (концентрированная)
2.2.6	Б3.О.02. Подготовка к сдаче и сдача государственного экзамена

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
ПК-4: Способен разрабатывать программное обеспечение разного уровня сложности и руководить его разработкой с использованием систем конфигурационного управления	
Знать:	
Уровень 1	методы создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов.
Уметь:	
Уровень 1	использовать методы создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов.
Владеть:	
Уровень 1	методами и средствами создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов с помощью современных компьютерных технологий.

В результате освоения дисциплины обучающийся должен

3.1	**Знать:**
3.1.1	- возможности анализа информации на основе методов статистической обработки данных (СОД) ;
3.1.2	- основные понятия, задачи, стадии и методы СОД,
3.1.3	- статистики и критерии для выявления процессов статистических характеристик случайных величин.
3.2	**Уметь:**
3.2.1	- выбирать методику статистического исследования экспериментальных данных; расчитывать интервалы для оценки характеристик случайных величин;
3.2.2	- определять степень полинома регрессионной зависимости в условиях неизвестного класса функций;
3.2.3	- расчитывать интерполяционные полиномы различными методами;

3.2.4	- проверять соответствие выдвигаемых гипотез с заданным уровнем значимости экспериментальным результатам;
3.2.5	- проводить дисперсионый и корреляционный анализы, анализ временных рядов.
3.3	**Владеть:**
3.3.1	- навыками систематизации, математического анализа и интерпретации количественной информации с помощью современных цифровых инструментов.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетен-ции	Литература	Инте ракт.	Пр. подг.	Примечание
	Раздел 1. Предмет и методы статистического анализа данных. Представление исходных данных							
1.1	Основные принципы и задачи анализа данных в фундаментальных и прикладных исследованиях. Общие сведения о разведочном анализе: описание данных, снижение размерности, классификация. Общие сведения о статистическом выводе: оценивание параметров и проверка статистических гипотез. /Лек/	2	2	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			
1.2	Изучение базовых типов, управляющих конструкций, функций, формул, пакетов расширений и графики языка R /Ср/	2	12	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			
1.3	Практическая работа №1. Знакомство со средой программирования RStudio. /Пр/	2	8	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3		8	
1.4	Проработка лекционного материала. Выполнение практической работы №1. /Ср/	2	12	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			
1.5	Типы исходных данных. Получение исходных данных для статистического анализа. Переменные и наблюдения. Представление исходных данных. Ряды распределения. Дискретные и непрерывные переменные. Таблицы частот. Группировка. Абсолютные и относительные частоты. Накопленные частоты. Эмпирическая функция распределения. Плотность распределения. Наглядное представление частот: диаграммы. Сводные характеристики распределения. /Лек/	2	2	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			
	Раздел 2. Статистические методы обработки и анализа экспериментальных данных							

2.1	Меры положения: минимум и максимум, квантили. Квантильная функция. Меры центральной тенденции: мода, медиана, среднее. Меры разброса: размах, межквартильный размах, стандартное отклонение, дисперсия, коэффициент вариации. Меры формы: коэффициенты асимметрии, эксцесса. Выявление выбросов. Робастность сводных характеристик. Преобразование данных. Преобразование порядковых данных: ранжирование. Метрические данные: центрирование, нормирование, стандартизация (перевод в z-шкалу). Нормальное распределение и его свойства. Нелинейная нормализация: степенное и логарифмическое преобразования; квантильная нормализация. Стандартизация измерений в психологии и стандартные шкалы. /Лек/	2	4	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3	2		Разбор реальных ситуаций (case study)
2.2	Практическая работа №2. Исследование законов распределения экспериментальных данных и проверка статистических гипотез. /Пр/	2	8	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3	2	8	совместная работа (в парах)
2.3	Изучение статистических методов анализа экспериментальных данных. /Ср/	2	4	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			
2.4	Проработка лекционного материала. Выполнение практической работы №2 /Ср/	2	12	ПК-4	Л1.1 Л1.2Л2.2 Л2.1 Л2.3 Э3			

2.5	Двумерные данные. Двумерное распределение. Функциональная и стохастическая связь. Меры связи. Коэффициенты сопряжѐнности. Парная корреляция. Двумерное нормальное распределение. Коэффициент линейной корреляции Пирсона. Коэффициенты ранговой корреляции Спирмена и Кэндалла. Наглядное представление парной связи. Модель дисперсионного анализа и коэффициент эта-квадрат. Регрессионная модель. Метод наименьших квадратов. Простая линейная регрессия и коэффициент детерминации. /Лек/	2	2	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			
	Раздел 3. Оценивание неизвестных параметров распределения. Статистическме критерии.							
3.1	Оценивание неизвестных параметров распределения. Статистический вывод. Сплошные и выборочные исследования. Генеральная совокупность и выборка. Задача оценивания параметра распределения генеральной совокупности. Параметры и статистики. Выборочные оценки. Метод моментов. Точечное и интервальное оценивание. Стандартная ошибка оценки. Доверительный интервал. /Лек/	2	2	ПК-4	Л1.1 Л1.2Л2.3 Л2.1 Л2.2 Э3	2		Разбор реальных ситуаций (case study)
3.2	Практическая работа №3. Регрессионный анализ /Пр/	2	6	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3	2	6	совместная работа (в парах)
3.3	Изучение методов классификации и прогнозирования. /Ср/	2	15,9	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			

3.4	Проверка статистических гипотез. Исследовательские (содержательные) и статистические гипотезы. Нулевая гипотеза. Размер эффекта и статистическая значимость. Принятие статистического решения. Ошибки первого и второго рода. Статистический критерий и его мощность. Статистические критерии. Критерии согласия, однородности и независимости. Критерии хи-квадрат Пирсона, D Колмогорова-Смирнова, W Шапиро-Уилка. Параметрические критерии. Одновыборочный и двухвыборочный t-критерий Пирсона. Критерий F Фишера. Непараметрические критерии. Медианный критерий. Критерий U Манна-Уитни. Схема повторных измерений. Критерии хи-квадрат МакНемара, t Стьюдента для зависимых выборок и Т Уилкоксона. /Лек/	2	4	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э3			
3.5	Проработка лекционного материала. Выполнение практической работы №3 /Ср/	2	14	ПК-4	Л1.1 Л1.2Л2.1 Л2.2 Л2.3 Э1 Э2 Э3			
3.6	/КрТО/	2	0,1	ПК-4	Э3			

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

5.1. Контрольные вопросы и задания

Контрольные вопросы для проверки ЗНАТЬ:
1. Классификация экспериментальных исследований. Ошибки исследования. Содержание экспериментальных исследований.
2. Определение генеральной совокупности. Определение выборочной совокупности. Вариационный и статистический ряды.
3. Эмпирическая функция распределения. Выборочное среднее. Асимметрия и эксцесс. Выборочная дисперсия
4. Изображение статистических рядов. Гистограмма. Полигон.
5. Точечное оценивание числовых характеристик и параметров распределений. Доверительные интервалы. Состоятельность, несмещенность, эффективность оценок, робастность оценок.
6. Проверка статистических гипотез. Постановка задачи проверки статистических гипотез. Критерий значимости. Односторонний и двусторонний критерии. Мощность критерия. Общая схема проверки статистических гипотез. Ошибки первого и второго рода. Проверка гипотез о законе распределения генеральной совокупности. Метод Пирсона. Критерий Смирнова.
7. Основные понятия, задачи интеллектуального анализа данных.
8. Классификация и прогнозирование. Обучение с учителем. Обучение без учителя. Дерево решений.
9. Метод опорных векторов.
10. Кластерный анализ. Иерархические и итерационные методы.
11. Анализ взаимосвязей между переменными, ассоциативные правила.

Контрольные вопросы для проверки УМЕТЬ и ВЛАДЕТЬ:
1. Построение эмпирической функции распределения, гистограммы, полигона абсолютных и относительных частот.
2. Реализация проверки статистических гипотез.

3. Реализация методов кластерного анализа.
4. Построение дерева решений.
5. Методы классификации и прогнозирования.
6. Построение ассоциативных правил.

5.2. Темы курсовых работ (проектов)

не предусмотрены

5.3. Фонд оценочных средств

Задания к практической работе №1. Знакомство с языком программирования R и средой программирования RStudio.
1. Изучить основы работы в системе R.
2. Изучить базовые типы и структуры данных: вектор, матрица, массив, список, фактор, фрейм.
3. Изучить управляющие конструкции языка R: if/else, for, while, repeat.
4. Изучить и научиться использовать функции, формулы.
5. Изучить наиболее распространённые пакеты расширений. Научиться подключать необходимые пакеты, просматривать список установленных пакетов.
6. Изучить функции работы с графикой, научиться использовать их.

Задания к практической работе №2. Исследование законов распределения экспериментальных данных и проверка статистических гипотез.
1. Анализ грубых ошибок измерения для заданной выборки.
* Реализовать на языке R программу, реализующую два способа выявления грубых ошибок измерения:
о сравнением нормированного значения случайной величины, подозреваемого как промах, с квантилем t -распределения для числа степеней свободы (- количество измерений) и вероятности . Уровень значимости α взять равным 5% (0,05).
о Сравнением нормированного значения случайной величины (математическое ожидание и дисперсия вычисляются без учета значения, подозреваемого как промах) со значением коэффициента K. K=4 для 6 < n <=100; K=4,5 для 100 < n <=1000 и K=5 для n > 1000.
2. Построение эмпирических функций распределения и плотности распределения. Программная реализация на языке R.
* Реализовать функции, определяющие минимальное и максимальное значения выборки, и осуществляющую построение вариационного ряда.
* Определить количество интервалов группировки.
* Построить гистограмму относительных и абсолютных частот.
* Построить полигон относительных и абсолютных частот.
* Построить эмпирическую функцию распределения.
3. Выполнить точечные оценки параметров распределения:
* определить оценки начальных моментов;
* вычислить смещенные и несмещенные оценки центральных моментов;
* определить:
о среднеквадратическое отклонение;
о коэффициент асимметрии;
о коэффициент эксцесса;
о стандартизованные переменные.
4. Проверка статистических гипотез * Провести проверку статистических гипотез о нормальном законе распределения с использованием критерия хи-квадрат К.Пирсона для двух способов расчета теоретических значений частот. 5. Подготовить в электронном виде отчет о проделанной работе.

Задания к Практической работе № 3. Регрессионный анализ.
1. Оценить параметры полиномиальной регрессии по экспериментальным данным:
* (I) На первом шаге представить функциональную связь между экспериментальными данными в виде полиномиальной регрессионной модели со степенью m=1.
* (II) Вычислить оценки параметров регрессии методом наименьших квадратов.
* (III) Проверить гипотезу об адекватности выбранной модели.
* (IV) Если гипотеза об адекватности не принимается, то необходимо увеличить степень полинома на единицу (m=m+1) и повторять пункты II, III и IV до тех пор, пока гипотеза об адекватности модели не станет справедливой (до m<= n-2).
2. Оценить параметры линейной регрессионной модели:
* (II) Вычислить оценки параметров регрессии
* (III) Проверить гипотезу об адекватности выбранной модели
* (IV) Проверить гипотезу о значимости параметров регрессии.
* (V) Незначимые параметры следует отбросить и вновь решить регрессионную задачу (см. пункт II), а затем следует выполнить пункт III. Если гипотеза об адекватности не подтверждается, то следует вернуться к предыдущей модели.
* (VI) Пункты IV и V повторять, пока все параметры не будут просмотрены на значимость.
4. Подготовить в электронном виде отчет о проделанной работе.

5.4. Перечень видов оценочных средств

Практические задания; Контрольные вопросы. Виды шкал оценивания представлены в Приложении 1

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год
Л1.1	Третьяк Л. Н., Воробьев А. Л.	Основы теории и практики обработки экспериментальных данных: Учебное пособие	Оренбург: Оренбургский государственный университет, ЭБС АСВ 2015
Л1.2	Васильев, А. А.	Теория вероятностей и математическая статистика : учебник и практикум для вузов / А. А. Васильев. — 2-е изд., испр. и доп.—224с. //	Москва : Издательство Юрайт, 2024 — ISBN 78-5-534-16714-6. Образовательная платформа Юрайт [сайт]. URL: https://urait.ru/bcode/538884

6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год
Л2.1	Stephane Tuffery	Data Mining and Statistics for Decision Making.	John Wiley & Sons, Ltd 2011
Л2.2	I.H.Witten, E.Frank,	Data Mining. Practical Machine Learning Tools and Techniques.	Elsevier Inc. 2011
Л2.3	Meta S. Brown	Data Mining for Dummies	John Wiley & Sons, Ltd 2014

6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	RStudio Desktop	https://www.rstudio.com/products/rstudio/download/
Э2	The Comprehensive R Archive Network	https://cran.rstudio.com/
Э3	The R Project for Statistical Computing.	https://www.r-project.org/

6.3. Перечень информационных и образовательных технологий

6.3.1 Компетентностно-ориентированные образовательные технологии

6.3.1.1	Изучение дисциплины студентами осуществляется в форме лекций, практических занятий в аудиторных условиях (компьютерные классы) и в процессе самостоятельной работы, контроля знаний.
6.3.1.2	Теоретическая информация представляется в виде компьютерных презентаций с использованием мультимедийных средств.
6.3.1.3	При проведении лекций используются интерактивные формы обучения (технологии типа «кейс-стади», т.е. в процессе лекции делается разбор часто встречающихся практических ситуаций с последующим опросом студентов на следующей лекции и организацией диалога «преподаватель-студент», «студент-студент» с целью выявления степени усвоения материала).
6.3.1.4	Практические занятия проводятся в компьютерных классах, оснащенных персональными компьютерами необходимой конфигурации и с установленным соответствующим программным обеспечением.
6.3.1.5	Практические занятия проводятся в интерактивной форме, в группах (используются технологии группового выполнения практических работ).
6.3.1.6	Защита практических работ проводится в виде собеседования с преподавателем по теории и программной реализации работы.

6.3.2 Перечень информационных справочных систем и программного обеспечения

6.3.2.1	MS Windows 10;
6.3.2.2	MS Office 2013;
6.3.2.3	RStudio, Интерпретатор языка R.

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1	Дисциплина должна быть обеспечена необходимым аппаратным (персональные компьютеры) и программным обеспечением (операционная система ОС Windows и пакет прикладных программ MS Office, RStudio, Интерпретатор языка R). Для проведения лекционных занятий используется мультимедийное оборудование (проектор, экран, интерактивная доска).

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

1. Технологическая карта дисциплины представлена в Приложении 2.
2. Текущий контроль осуществляется в течение семестра (оцениваются посещаемость и активность на занятиях, с/р студента).
3. Рубежный контроль осуществляется в течение семестра в виде защиты практических работ.
4. Промежуточный контроль проводится письменно. Письменное задание состоит из тестовых и открытых вопросов, а также включает вопросы практической направленности.
5. Все материалы по дисциплине (лекции, методические указания по выполнению практических работ, домашние задания, литература и др.) выкладываются в Google Classroom.

6. После выполнения каждой практической работы студент должен предоставить электронный отчет о ее выполнении.