

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 20.05.2024 15:58:19
Уникальный программный ключ:
ca953a0120d891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов имени Патриса Лумумбы»**

Инженерная академия

(наименование основного учебного подразделения (ОУП)-разработчика ОП ВО)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ДАННЫХ

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

01.04.02 ПРИКЛАДНАЯ МАТЕМАТИКА И ИНФОРМАТИКА

(код и наименование направления подготовки/специальности)

Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):

DATA SCIENCE И ЦИФРОВАЯ ТРАНСФОРМАЦИЯ

(наименование (профиль/специализация) ОП ВО)

2024 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Статистические методы анализа данных» входит в программу магистратуры «Data Science и цифровая трансформация» по направлению 01.04.02 «Прикладная математика и информатика» и изучается в 1 семестре 1 курса. Дисциплину реализует Кафедра механики и процессов управления. Дисциплина состоит из 8 разделов и 21 тема и направлена на изучение фундаментальных основ аппарата дескриптивного статистического анализа и доверительного оценивания; разбор основных методов решения типовых задач и знакомство с областью их применения в профессиональной деятельности.

Целью освоения дисциплины является формирование фундаментальных знаний и навыков применения методов решения задач, необходимых для профессиональной деятельности, повышение общего уровня математической грамотности студентов, чтобы научить их вычислять точечные оценки неизвестных параметров распределений и строить доверительные интервалы, дать студентам представление об основах корреляционного анализа.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Статистические методы анализа данных» направлено на формирование у обучающихся следующих компетенций (части компетенций):

Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
ОПК-1	Способен решать актуальные задачи фундаментальной и прикладной математики	ОПК-1.1 Знает основные законы, положения и методы в области решения актуальных задач фундаментальной и прикладной математики;; ОПК-1.2 Умеет выявлять сущность задач фундаментальной и прикладной математики;; ОПК-1.3 Владеет инструментами для решения задач задачи фундаментальной и прикладной математики.;
ОПК-2	Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Знает существующие математические методы решения прикладных задач;; ОПК-2.2 Владеет инструментами реализации новых математических методов решения прикладных задач;; ОПК-2.3 Владеет необходимыми знаниями, позволяющими совершенствовать существующие методы решения прикладных задач.;
ОПК-3	Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	ОПК-3.1 Знает методы и подходы для разработки математических моделей и анализа получаемых решений в области профессиональной деятельности;; ОПК-3.2 Умеет применять и выбирать наиболее приемлемые методы и подходы для разработки математических моделей и анализа получаемых решений;; ОПК-3.3 Владеет инструментами разработки и анализа математических моделей объектов профессиональной деятельности.;
ПК-1	Способен разрабатывать новые методики выполнения аналитических работ	ПК-1.1 Знает основные методы и подходы к анализу данных;; ПК-1.2 Умеет применять известные методы и подходы для проведения анализа данных;; ПК-1.3 Владеет алгоритмами по разработке методик проведения аналитических работ в профессиональной области.;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Статистические методы анализа данных» относится к обязательной части блока 1 «Дисциплины (модули)» образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Статистические методы анализа данных».

Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
ОПК-1	Способен решать актуальные задачи фундаментальной и прикладной математики		Преддипломная практика; Системы искусственного интеллекта; Машинное обучение и анализ больших данных; Geoinformation Systems and Applications;
ОПК-2	Способен совершенствовать и реализовывать новые математические методы решения прикладных задач		Искусственные нейронные сети (Глубокое обучение); Преддипломная практика;
ОПК-3	Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности		Преддипломная практика; Научно-исследовательская работа; Машинное обучение и анализ больших данных; Системы искусственного интеллекта; Проектирование автоматизированных систем управления;
ПК-1	Способен разрабатывать новые методики выполнения аналитических работ		Преддипломная практика; Научно-исследовательская работа; Системы искусственного интеллекта; Искусственные нейронные сети (Глубокое обучение); Машинное обучение и анализ больших данных; <i>Динамика и управление космическими системами**</i> ; <i>Dynamics and Control of Space Systems**</i> ; Geoinformation Systems and Applications;

* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Статистические методы анализа данных» составляет «4» зачетные единицы.

Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			1
<i>Контактная работа, ак.ч.</i>	34		34
Лекции (ЛК)	17		17
Лабораторные работы (ЛР)	17		17
Практические/семинарские занятия (СЗ)	0		0
<i>Самостоятельная работа обучающихся, ак.ч.</i>	83		83
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	27		27
Общая трудоемкость дисциплины	ак.ч.	144	144
	зач.ед.	4	4

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 1	Вероятностно-статистическая модель. Статистики. Выборочное среднее и дисперсия. Медиана и мода. Эмпирическая функция распределения. Гистограмма	1.1	Понятие выборки. Способы представления выборочных данных	ЛК, ЛР
		1.2	Основные выборочные характеристики: выборочное среднее, выборочная дисперсия, медиана, мода	ЛК, ЛР
		1.3	Эмпирическая функция распределения. Визуализация выборочных данных. Гистограмма. "Ящик с усами"	ЛК, ЛР
Раздел 2	Свойства статистик. Задача точечного оценивания. Метод максимального правдоподобия. Метод моментов	2.1	Понятие дескриптивной статистики	ЛК, ЛР
		2.2	Свойства выборочных оценок: состоятельность, несмещённость, эффективность	ЛК, ЛР
		2.3	Методы нахождения точечных оценок неизвестных параметров распределений: метод максимального правдоподобия, метод моментов	ЛК, ЛР
Раздел 3	Точечное и доверительное оценивание математического ожидания нормального распределения при известной и неизвестной дисперсии	3.1	Определение понятия доверительного множества	ЛК, ЛР
		3.2	Вывод формул для границ доверительных интервалов для математического ожидания нормального распределения при известной и неизвестной дисперсии	ЛК, ЛР
Раздел 4	Доверительный интервал для среднего квадратического отклонения нормального распределения при неизвестных параметрах. Распределение хи-квадрат. Распределение Стьюдента. Доверительный интервал для пропорции (доли) в случае больших выборок	4.1	Теоретическое обоснование формул для границ доверительного интервала для среднего квадратического отклонения нормального распределения (при неизвестных параметрах) и для пропорции (доли) в случае больших выборок	ЛК, ЛР
		4.2	Свойства распределений Хи-квадрат и Стьюдента	ЛК, ЛР
Раздел 5	Основные понятия теории проверки гипотез. Ошибки 1-го и 2-го рода. Р-значение. Проверка гипотез о параметрах нормального распределения.	5.1	Теоретические сведения о статистических гипотезах и релевантном понятийном аппарате.	ЛК, ЛР
		5.2	Критерии проверки гипотез о параметрах нормального распределения	ЛК, ЛР
Раздел 6	Сравнение двух групп	6.1	Возможные постановки задачи о сравнении выборок в форме статистических гипотез	ЛК, ЛР
		6.2	Критерии проверки гипотез о равенстве средних нормальных генеральных совокупностей при известных и неизвестных дисперсиях	ЛК, ЛР
		6.3	Критерий Стьюдента	ЛК, ЛР
Раздел 7	Критерий Хи-квадрат. Проверка независимости признаков	7.1	Возможные постановки задачи о независимости количественных и качественных признаков (компонент двумерных случайных величин)	ЛК, ЛР
		7.2	Теоретическое обоснование критерия Хи-квадрат	ЛК, ЛР
		7.3	Мозаичная диаграмма как средство визуальной оценки правдоподобности гипотезы о независимости признаков (в случае качественной группировки)	ЛК, ЛР

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 8	Непараметрические критерии проверки гипотез. Критерий Уилкоксона. Основы корреляционного анализа	8.1	Теоретическое обоснование применения непараметрических критериев проверки гипотез. Критерий Уилкоксона	ЛК, ЛР
		8.2	Основные понятия корреляционного анализа: ковариация, коэффициент корреляции, их выборочные оценки	ЛК, ЛР
		8.3	Применение выборочного коэффициента корреляции для оценки степени зависимости признаков (компонент двумерной нормальной случайной величины)	ЛК, ЛР

* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/семинарские занятия.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве [Параметр] шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	

* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики. - СПб.: Лань, 2011. - 256с
2. Ширяев А.Н., Эрлих И.Г., Яськов П.А. Вероятность в теоремах и задачах (с доказательствами и решениями). Книга 1. - М.: МЦНМО, 2013. - 648 с.
3. Статистический анализ данных, моделирование и исследование вероятностных

закономерностей. Компьютерный подход / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов и др. - М.: НИЦ ИНФРА-М, 2015. - 890 с.

Дополнительная литература:

1. Боровков А.А. Математическая статистика. - СПб.: Лань, 2010. - 704 с
2. Свешников А.А. Прикладные методы теории вероятностей. - СПб.: Лань, 2012. - 480 с.
3. Козлов А.Ю., Мхитарян В.С., Шишов В.Ф. Статистический анализ данных в MS Excel. - М.: ИНФРА-М, 2014. - 320 с.

Ресурсы информационно-телекоммуникационной сети «Интернет»:

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров

- Электронно-библиотечная система РУДН – ЭБС РУДН

<http://lib.rudn.ru/MegaPro/Web>

- ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>

- ЭБС Юрайт <http://www.biblio-online.ru>

- ЭБС «Консультант студента» www.studentlibrary.ru

- ЭБС «Троицкий мост»

2. Базы данных и поисковые системы

- электронный фонд правовой и нормативно-технической документации

<http://docs.cntd.ru/>

- поисковая система Яндекс <https://www.yandex.ru/>

- поисковая система Google <https://www.google.ru/>

- реферативная база данных SCOPUS

<http://www.elsevier.com/locate/scopus/>

Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля:*

1. Курс лекций по дисциплине «Статистические методы анализа данных».

* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ

Оценочные материалы и балльно-рейтинговая система* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Статистические методы анализа данных» представлены в Приложении к настоящей Рабочей программе дисциплины.

* - Ом и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.

РАЗРАБОТЧИКИ:

Доцент

Должность, БУП

Подпись

Салтыкова Ольга
Александровна

Фамилия И.О.

Доцент

Должность, БУП

Подпись

Демидов Александр
Сергеевич

Фамилия И.О.

РУКОВОДИТЕЛЬ БУП:

Заведующий кафедрой

Должность БУП

Подпись

Разумный Юрий
Николаевич

Фамилия И.О.

РУКОВОДИТЕЛЬ ОП ВО:

Доцент

Должность, БУП

Подпись

Салтыкова Ольга
Александровна

Фамилия И.О.