

Подписано электронной подписью:
Вержицкий Данил Григорьевич
Должность: Директор КГПИ ФГБОУ ВО «КемГУ»
Дата и время: 2024-02-21 00:00:00

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Кемеровский государственный университет»
Кузбасский гуманитарно-педагогический институт
федерального государственного бюджетного образовательного учреждения
высшего образования
«Кемеровский государственный университет»
Факультет информатики, математики и экономики

УТВЕРЖДАЮ
Декан
А. В. Фомина
23 июня 2021 г.

Рабочая программа дисциплины

Б1.В.05 Прикладная статистика и анализ данных

Код, название дисциплины

Направление подготовки

**02.03.03 Математическое обеспечение и администрирование
информационных систем**

Код, название направления

Направленность (профиль) подготовки

Программное и математическое обеспечение информационных технологий

Программа бакалавриата

Квалификация выпускника
бакалавр

Форма обучения
очная

Год набора 2021

Новокузнецк 2021

Оглавление

1 Цель дисциплины	3
1.1 Формируемые компетенции	3
1.2 Индикаторы достижения компетенций.....	3
1.3 Знания, умения, навыки (ЗУВ) по дисциплине	4
2 Объём и трудоёмкость дисциплины по видам учебных занятий. Формы промежуточной аттестации.	5
3. Учебно-тематический план и содержание дисциплины.....	6
3.1 Учебно-тематический план	6
3.2. Содержание занятий по видам учебной работы.....	6
4 Порядок оценивания успеваемости и сформированности компетенций обучающегося в текущей и промежуточной аттестации.....	7
5 Материально-техническое, программное и учебно-методическое обеспечение дисциплины.	8
5.1 Учебная литература	8
5.2 Материально-техническое и программное обеспечение дисциплины.....	9
5.3 Современные профессиональные базы данных и информационные справочные системы.	10
6 Иные сведения и (или) материалы.....	10
6.1.Примерные темы письменных учебных работ	10
6.2. Примерные вопросы и задания / задачи для промежуточной аттестации	10

1 Цель дисциплины.

В результате освоения данной дисциплины у обучающегося должны быть сформированы компетенции основной профессиональной образовательной программы бакалавриата (далее - ОПОП):

ПК-1

Содержание компетенций как планируемых результатов обучения по дисциплине см. таблицы 1 и 2.

1.1 Формируемые компетенции

Таблица 1 - Формируемые дисциплиной компетенции

Наименование вида компетенции (универсальная, общепрофессиональная, профессиональная)	Наименование категории (группы) компетенций	Код и название компетенции
профессиональная		ПК-1 Способен применять математические методы с учетом допущений и ограничений, связанных с выбранным математическим материалом, и обосновывать выбор алгоритма решения задачи

1.2 Индикаторы достижения компетенций

Таблица 2 – Индикаторы достижения компетенций, формируемые дисциплиной

Код и название компетенции	Индикаторы достижения компетенции по ОПОП	Дисциплины и практики, формирующие компетенцию ОПОП
ПК-1 Способен применять математические методы с учетом допущений и ограничений, связанных с выбранным математическим материалом, и обосновывать выбор алгоритма решения задачи	1.1 Использует современные математические методы при разработке алгоритмов решения задач 1.2 Определяет допущения и ограничения математических методов при решении профессиональных задач 1.3 Использует математические материалы	Б1.В.05 Прикладная статистика и анализ данных Б1.В.06 Вычислительная математика Б1.В.ДВ.01.01 Математические модели и методы искусственного интеллекта Б1.В.ДВ.01.02 Разработка экспертных систем Б2.В.01(П) Производственная

Код и название компетенции	Индикаторы достижения компетенции по ОПОП	Дисциплины и практики, формирующие компетенцию ОПОП
	для разработки схем взаимодействия программы с другими программами 1.4 Выбирает математический материал для описания метода организации входных и выходных данных алгоритмов 1.5 Выбирает математический материал для разработки СИИ	практика. Профильная практика БЗ.01(Д) Подготовка к процедуре защиты и защита выпускной квалификационной работы

1.3 Знания, умения, навыки (ЗУВ) по дисциплине

Таблица 3 – Знания, умения, навыки, формируемые дисциплиной

Код и название компетенции	Индикаторы достижения компетенции, закрепленные за дисциплиной	Знания, умения, навыки (ЗУВ), формируемые дисциплиной
ПК-1 Способен применять математические методы с учетом допущений и ограничений, связанных с выбранным математическим материалом, и обосновывать выбор алгоритма решения задачи	1.1 Использует современные математические методы при разработке алгоритмов решения задач 1.2 Определяет допущения и ограничения математических методов при решении профессиональных задач	Знать: – Современные математические методы при разработке алгоритмов решения задач в области прикладной статистики и анализа данных. Уметь: – Определять допущения и ограничения математических методов при решении профессиональных задач в области прикладной статистики и анализа данных; – Использовать современные математические методы при решении задач прикладной статистики; – выбирать и применять математические методы необходимые для решения поставленных задач прикладной статистики Владеть: – математическими методами при разработке алгоритмов решения задач в области прикладной статистики и анализа данных.

**2 Объём и трудоёмкость дисциплины по видам учебных занятий.
Формы промежуточной аттестации.**

Таблица 4 – Объём и трудоёмкость дисциплины по видам учебных занятий

Общая трудоёмкость и виды учебной работы по дисциплине, проводимые в разных формах	Объём часов по формам обучения
	ОФО
1 Общая трудоёмкость дисциплины	144
2 Контактная работа обучающихся с преподавателем (по видам учебных занятий) (всего)	48
Аудиторная работа (всего):	48
в том числе:	
лекции	16
практические занятия, семинары	32
практикумы	
лабораторные работы	
в интерактивной форме	
в электронной форме	
Внеаудиторная работа (всего):	
в том числе, индивидуальная работа обучающихся с преподавателем	
подготовка курсовой работы /контактная работа	
групповая, индивидуальная консультация и иные виды учебной деятельности, предусматривающие групповую или индивидуальную работу обучающихся с преподавателем)	
творческая работа (эссе)	
3 Самостоятельная работа обучающихся (всего)	96
4 Промежуточная аттестация обучающегося:	Зачет с оценкой - 4 семестр

3. Учебно-тематический план и содержание дисциплины.

3.1 Учебно-тематический план

Таблица 5 - Учебно-тематический план очной формы обучения

№ недели п/п	Разделы и темы дисциплины по занятиям	Общая трудоёмкость (всего час.)	Трудоёмкость занятий (час.)			Формы текущего контроля и промежуточной аттестации успеваемости
			ОФО			
			Аудиторн. занятия		СРС	
			лекц.	практ.		
Семестр 4						
1	Статистические гипотезы	22	2	4	16	Лабораторная работа 1,2
2	Дисперсионный анализ	26	2	4	20	Лабораторная работа 3,4
3	Непараметрические критерии проверки статистических гипотез	32	4	8	20	Лабораторная работа 5,6
4	Корреляционный анализ	32	4	8	20	Лабораторная работа 7,8
5	Регрессионный анализ	32	4	8	20	Лабораторная работа 9,10
	Промежуточная аттестация					Зачет с оценкой
ИТОГО по семестру 4		144	16	32	96	
	Всего:	144	16	32	96	

3.2. Содержание занятий по видам учебной работы

Таблица 6 – Содержание дисциплины

Содержание лекционного курса

№ п/п	Наименование раздела и темы дисциплины	Содержание
1.1	Статистические гипотезы	Основные методы проверки статистических гипотез. Критерий Пирсона, Стьюдента, Фишера.
1.2	Дисперсионный анализ	Однофакторный и многофакторный дисперсионный анализ.
1.3	Непараметрические критерии проверки статистических гипотез	Отличие параметрических и непараметрических критериев для проверки статистических гипотез. Критерий Колмогорова-Смирнова, критерий Манна-Уитни, Вальда-Волфовица, Вилкоксона.
1.4	Корреляционный анализ	Вычисление параметрических коэффициентов корреляции. Вычисление непараметрических коэффициентов корреляции
1.5	Регрессионный анализ	Построение модели регрессии. Оценка адекватности регрессии. Анализ регрессионных остатков. Доверительный интервал для уравнения регрессии. Оценка дисперсии коэффициента регрессии и доверительных интервалов.

№ п/п	Наименование раздела и темы дисциплины	Содержание
1.1	Статистические гипотезы	Проверка статистических гипотез с помощью параметрических критериев. Гипотезы о значимости коэффициента корреляции, гипотеза о равенстве математических ожиданий, гипотеза о равенстве дисперсий.
1.2	Дисперсионный анализ	Применения однофакторного дисперсионного анализа для определения влияние фактора на отклик. Применения двухфакторного дисперсионного анализа для определения влияние фактора на отклик.
1.3	Непараметрические критерии проверки статистических гипотез	Проверка гипотез о равенстве функций распределения, расчет коэффициентов ранговой корреляции, проверка гипотезы о согласованности.
1.4	Корреляционный анализ	Вычислить коэффициент корреляции. Вычислить коэффициенты ранговой корреляции Спирмена, Кендела, коэффициент конкордации.
1.5	Регрессионный анализ	Построение модели регрессии. Оценка адекватности регрессии.

4 Порядок оценивания успеваемости и сформированности компетенций обучающегося в текущей и промежуточной аттестации.

Для положительной оценки по результатам освоения дисциплины обучающемуся необходимо выполнить все установленные виды учебной работы. Оценка результатов работы обучающегося в баллах (по видам) приведена в таблице 7.

Таблица 7 - Шкала и показатели оценивания результатов учебной работы обучающихся по видам в балльно-рейтинговой системе (БРС)

Учебная работа (виды)	Сумма баллов	Виды и результаты учебной работы	Оценка в аттестации (шкала и показатели оценивания)	Баллы
Текущая учебная работа в семестре (Посещение занятий по расписанию и выполнение заданий)	80	Лабораторная работа №1 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №2 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №3 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №4 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №5 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №6 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8

		Лабораторная работа №7 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №8 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №9 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
		Лабораторная работа №10 (отчет о выполнении контрольной работы)	4 балла (пороговое значение) 8 баллов (максимальное значение)	4-8
Итого по текущей работе в семестре				40-80
Промежуточная аттестация (зачет с оценкой)	20	Теоретический вопрос 1	2 балла (пороговое значение) 5 баллов (максимальное значение)	2 - 5
		Теоретический вопрос 2	2 балла (пороговое значение) 5 баллов (максимальное значение)	2 - 5
		Решение задачи 1.	2 балла (пороговое значение) 5 баллов (максимальное значение)	2 - 5
		Решение задачи 2.	2 балла (пороговое значение) 5 баллов (максимальное значение)	2 - 5
Итого по промежуточной аттестации (зачет с оценкой)				8-20
Суммарная оценка по дисциплине/ Сумма баллов текущей и промежуточной аттестации				51 – 100 б

В промежуточной аттестации оценка выставляется в ведомость в 100-балльной шкале и в буквенном эквиваленте (таблица 8)

Таблица 8 – Соотнесение 100-балльной шкалы и буквенного эквивалента оценки

Сумма набранных баллов	Уровни освоения дисциплины и компетенций	Экзамен		Зачет
		Оценка	Буквенный эквивалент	Буквенный эквивалент
86 - 100	Продвинутый	5	отлично	Зачтено
66 - 85	Повышенный	4	хорошо	
51 - 65	Пороговый	3	удовлетворительно	
0 - 50	Первый	2	неудовлетворительно	Не зачтено

5 Материально-техническое, программное и учебно-методическое обеспечение дисциплины.

5.1 Учебная литература

Основная учебная литература

1. Прикладная математическая статистика : учебное пособие / сост. А.А. Мицель. – Томск : ТУСУР, 2016. – 113 с. : ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=480889> (дата обращения: 21.01.2020). – Библиогр. в кн. – Текст : электронный.

2. Каган, Е.С. Прикладной статистический анализ данных : учебное пособие : [16+] / Е.С. Каган ; Кемеровский государственный университет. – Кемерово : Кемеровский государственный университет, 2018. – 235 с. : ил., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=573550> (дата обращения: 21.01.2020). – Библиогр.: с. 184-186. – ISBN 978-5-8353-2413-2. – Текст : электронный.

Дополнительная учебная литература

1. Кибзун, А.И. Теория вероятностей и математическая статистика [Электронный ресурс]. Базовый курс с примерами и задачами / А.И. Кибзун, Е.Р. Горяинова, А.В. Наумов – Электрон. текстовые дан. – Москва: Лань, 2005. – Режим доступа: <http://e.lanbook.com/reader/book/2198/>
2. Туганбаев, А.А. Теория вероятностей и математическая статистика [Электронный ресурс]: Учебник / А.А. Туганбаев, В.Г. Крупин – Электрон. текстовые дан. – Москва: Лань, 2011. – Режим доступа: <http://e.lanbook.com/reader/book/652/>
3. Бородин, А.Н. Элементарный курс теории вероятностей и математической статистики [Электронный ресурс]: Учебник / А.Н. Бородин – Электрон. текстовые дан. – Москва: Лань, 2011. – Режим доступа: <http://e.lanbook.com/reader/book/2026/>

5.2 Материально-техническое и программное обеспечение дисциплины.

Учебные занятия по дисциплине проводятся в учебных аудиториях НФИ КемГУ:

100 Учебная аудитория (мультимедийная) для проведения: -занятий лекционного типа; - групповых и индивидуальных консультаций. Специализированная (учебная) мебель: доска меловая, кафедра, столы, стулья, рабочее место для обучающегося с ОВЗ. Оборудование:<i>стационарное</i> - компьютер преподавателя, экран моторизованный, проектор, усилитель звука, колонки, микрофон преподавателя. Используемое программное обеспечение: Ubuntu Linux (свободно распространяемое ПО), LibreOffice (свободно распространяемое ПО), Яндекс.Браузер (отечественное свободно распространяемое ПО). Интернет с обеспечением доступа в ЭИОС.	654079, Кемеровская область, г. Новокузнецк, пр-кт Metallургов, д. 19
502 Компьютерный класс. Учебная аудитория (мультимедийная) для проведения: - занятий семинарского (практического) типа; - групповых и индивидуальных консультаций; - самостоятельной работы; - текущего контроля и промежуточной аттестации. Специализированная (учебная) мебель: доска меловая, столы компьютерные, стулья. Оборудование для презентации учебного материала: <i>стационарное</i> -компьютер, экран, проектор, наушники. Оборудование: <i>стационарное</i> – компьютеры для обучающихся (16 шт.). Используемое программное обеспечение: MS Windows (Microsoft Imagine Premium 3 year по лицензионному договору № 1212/KMP от 12.12.2018 г. до 12.12.2021 г.), LibreOffice (свободно распространяемое ПО), Firefox 14 (свободно распространяемое ПО),	654079, Кемеровская область, г. Новокузнецк, пр-кт Metallургов, д. 19

Яндекс.Браузер (отечественное свободно распространяемое ПО), Среда статистических вычислений R v.4.0.2 (свободно распространяемое ПО). Интернет с обеспечением доступа в ЭИОС.	
---	--

5.3 Современные профессиональные базы данных и информационные справочные системы.

Перечень СПБД и ИСС по дисциплине

1. База данных публикаций журнала Образование и общество, Федеральный портал Российское образование www.edu.ru, единое окно доступа к информационным ресурсам <http://window.edu.ru/resource/525/2525>

2. Общероссийский математический портал (информационная система) - <http://www.mathnet.ru/>

3. Научная электронная библиотека eLIBRARY.RU – крупнейший российский информационный портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты - www.elibrary.ru

6 Иные сведения и (или) материалы.

6.1. Примерные темы письменных учебных работ

6.2. Примерные вопросы и задания / задачи для промежуточной аттестации

Таблица 9 - Типовые (примерные) контрольные вопросы и задания

Разделы и темы	Примерные теоретические вопросы	Примерные практические задания и (или) задачи														
1. Название раздела																
Статистические гипотезы	Отклонение выборочной средней от заданного значения с известной и неизвестной дисперсией. Исключение выбросов. Отклонение выборочной дисперсии от заданного значения. Сравнение двух дисперсий нормальных генеральных совокупностей. Критерий Фишера. Критическое значение в зависимости от альтернативной гипотезы.	1. Произведено по 32 бросания пяти исследуемых монет и подсчитано количество k выпавших орлов $X = 0, 1, 2, 3, 4, 5$ в каждом испытании. В таблице приведены результаты. <table><tr><td>X</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>k</td><td>10</td><td>1</td><td>5</td><td>1</td><td>5</td><td>10</td></tr></table> Используя критерий Колмогорова – Смирнова, при уровне значимости $\alpha = 0,05$ проверить гипотезу о том, что среди этих монет нет поддельных.	X	0	1	2	3	4	5	k	10	1	5	1	5	10
X	0	1	2	3	4	5										
k	10	1	5	1	5	10										

	Сравнение нескольких дисперсий нормальных распределений. Критерий Бартлетта. Сравнение двух выборочных средних нормальных генеральных совокупностей, дисперсии которых известны и равны; не равны; неизвестны, но предполагаются равными. Практический смысл задачи. Сравнение более двух средних нормальных генеральных совокупностей, дисперсии которых неизвестны и одинаковы. Проверка гипотезы о значимости выборочного коэффициента корреляции двумерной нормальной генеральной совокупности.																																										
Дисперсионный анализ	Однофакторный дисперсионный анализ. Многофакторный дисперсионный анализ.	2. В таблице приведены данные об урожайности сельскохозяйственной культуры за 6 лет при разных технологиях обработки почвы. Выясним на уровне значимости $\alpha = 0,05$, зависит ли урожайность сельскохозяйственной культуры от технологии обработки почвы. Если зависит, то найти коэффициент детерминации технологии. <table><tr><th rowspan="2">Номер технологии</th><th colspan="6">Год наблюдения</th></tr><tr><th>1-й</th><th>2-й</th><th>3-й</th><th>4-й</th><th>5-й</th><th>6-й</th></tr><tr><td>1</td><td>140</td><td>141</td><td>140</td><td>141</td><td>142</td><td>145</td></tr><tr><td>2</td><td>150</td><td>149</td><td>150</td><td>147</td><td></td><td></td></tr><tr><td>3</td><td>147</td><td>147</td><td>145</td><td>150</td><td>150</td><td></td></tr><tr><td>4</td><td>144</td><td>147</td><td>142</td><td>146</td><td></td><td></td></tr></table>	Номер технологии	Год наблюдения						1-й	2-й	3-й	4-й	5-й	6-й	1	140	141	140	141	142	145	2	150	149	150	147			3	147	147	145	150	150		4	144	147	142	146		
Номер технологии	Год наблюдения																																										
	1-й	2-й	3-й	4-й	5-й	6-й																																					
1	140	141	140	141	142	145																																					
2	150	149	150	147																																							
3	147	147	145	150	150																																						
4	144	147	142	146																																							
Непараметрические	Методы проверки	1. Смоделировать выборку 80 значений																																									

критерии проверки статистических гипотез	гипотез, свободные от распределения. Критерии, основанные на знаках. Критерий Вальда-Вольфовица. Серии знаков для проверки гипотезы о случайности выборки. Критерии Вилкоксона, Манна-Уитни, Тьюки, Колмогорова-Смирнова.	случайной величины X , имеющей равномерное распределение в промежутке $[0, 4]$, проверив гипотезу о форме распределения. 2. Средний объем стока воды в реке фиксируется каждый месяц в течение двух лет. В таблице приведены наблюдаемые объемы стока в фут ³ /сек. Используя критерии Вальда – Вольфовица и Вилкоксона при уровне значимости $\alpha = 0,05$ проверить гипотезу об отсутствии систематических изменений объема стока из года в год. 1. Шесть пар участников соревнования по фигурному катанию на льду оценивались двумя членами жюри. В результате этой оценки у каждого члена жюри пара получала определенное место. Необходимо проверить с заданным уровнем значимости $\alpha = 0,05$ нулевую гипотезу $H_0 : W = 0$ об отсутствии согласованности оценок членов жюри. Результаты оценок приведены в таблице. <table><tr><th rowspan="2">Член жюри</th><th colspan="6">Номер пары</th></tr><tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th></tr><tr><td>1</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr><tr><td>2</td><td>1</td><td>2</td><td>4</td><td>5</td><td>3</td><td>6</td></tr></table>	Член жюри	Номер пары						1	2	3	4	5	6	1	1	2	3	4	5	6	2	1	2	4	5	3	6
Член жюри	Номер пары																												
	1	2	3	4	5	6																							
1	1	2	3	4	5	6																							
2	1	2	4	5	3	6																							
Корреляционный анализ	Коэффициент корреляции Исследование ранговой корреляции и конкордации. Ранговая корреляция Спирмена для двух выборок. Коэффициент конкордации Кендалла для нескольких выборок. Ранговая корреляция Кендалла.	Вычислить коэффициент множественной регрессии.																											
Регрессионный анализ	Построение модели регрессии. Оценка адекватности регрессии. Анализ регрессионных остатков. Доверительный интервал для уравнения регрессии.	Построить регрессионную модель на основе опытных данных.																											

	Оценка дисперсии коэффициента регрессии и доверительных интервалов. Нелинейная регрессионная модель. Множественная регрессионная модель.	
--	--	--

Составитель (и): Вячкин Е. С., доцент кафедры математики, физики и
математического моделирования

(фамилия, инициалы и должность преподавателя (ей))