

Подписано электронной подписью:
Вержицкий Данил Григорьевич
Должность: Директор КГПИ ФГБОУ ВО «КемГУ»
Дата и время: 2024-02-21 00:00:00
471086fad29a3b30e244c728abc3661ab35c9d50210dcf0e75e03a5b6fdf6436

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
«КЕМЕРОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Кузбасский гуманитарно-педагогический институт

Факультет информатики, математики и экономики

УТВЕРЖДАЮ
Декан
А.В. Фомина
«09» февраля 2023 г.

Рабочая программа дисциплины

К.М.07.05 Прикладная статистика и анализ данных

Направление подготовки

01.03.02 Прикладная математика и информатика

Направленность (профиль) подготовки

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

Программа бакалавриата

Квалификация выпускника

бакалавр

Форма обучения

Очная

Год набора 2023

Новокузнецк 2023

Оглавление

1 Цель дисциплины	3
Формируемые компетенции, индикаторы достижения компетенций, знания, умения, навыки	3
Место дисциплины.....	3
2 Объём и трудоёмкость дисциплины по видам учебных занятий. Формы промежуточной аттестации.	3
3. Учебно-тематический план и содержание дисциплины.....	4
3.1 Учебно-тематический план	4
4 Порядок оценивания успеваемости и сформированности компетенций обучающегося в текущей и промежуточной аттестации.....	4
5 Материально-техническое, программное и учебно-методическое обеспечение дисциплины.	5
5.1 Учебная литература	5
5.2 Материально-техническое и программное обеспечение дисциплины.....	6
5.3 Современные профессиональные базы данных и информационные справочные системы.....	7
6 Иные сведения и (или) материалы.....	7
6.1. Примерные вопросы и задания / задачи для промежуточной аттестации	7

1 Цель дисциплины.

В результате освоения данной дисциплины у обучающегося должны быть сформированы компетенции основной профессиональной образовательной программы бакалавриата (далее - ОПОП): ОПК-3.

Формируемые компетенции, индикаторы достижения компетенций, знания, умения, навыки

Таблица 1 – Индикаторы достижения компетенций, формируемые дисциплиной

Код и название компетенции	Индикаторы достижения компетенции по ОПОП	Знания, умения, навыки (ЗУВ), формируемые дисциплиной
ОПК-3 Способен применять и модифицировать математические модели для решения задач в области профессиональной деятельности	3.1 Применяет типовые математические модели для решения задач в области профессиональной деятельности 3.2 Адаптирует математические модели к конкретным профессиональным задачам	Знать: – Типовые математические модели при для решения прикладных задач математической статистики в профессиональной области. Уметь: – Применять математические модели при для решения прикладных задач математической статистики в профессиональной области. – Адаптировать математические модели прикладных задач математической статистики к конкретным профессиональным задачам. – выбирать и применять математические модели прикладной статистики необходимые для решения профессиональных задач Владеть: – Методами прикладной математической статистики для проектирования типовых математических моделей для решения задач в профессиональной области.

Место дисциплины

Дисциплина включена в модуль «Математическое моделирование в задачах профессиональной деятельности» ОПОП ВО. Дисциплина осваивается на 2 курсе в 3 семестре.

2 Объём и трудоёмкость дисциплины по видам учебных занятий.

Формы промежуточной аттестации.

Таблица 2 – Объем 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	108
2 Контактная работа обучающихся с преподавателем (по видам учебных занятий) (всего)	72
Аудиторная работа (всего):	72
в том числе:	
лекции	36
практические занятия, семинары	36
Внеаудиторная работа (всего):	
3 Самостоятельная работа обучающихся (всего)	36
4 Промежуточная аттестация обучающегося - зачет (3 семестр)	

3. Учебно-тематический план и содержание дисциплины.

3.1 Учебно-тематический план

Таблица 3 - 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 3						
1	Статистические гипотезы	20	8	6	6	Лабораторная работа 1,2
2	Дисперсионный анализ	12	4	4	4	Лабораторная работа 3,4
3	Непараметрические критерии проверки статистических гипотез	24	8	8	8	Лабораторная работа 5,6
4	Многомерный регрессионный анализ	16	4	6	6	Лабораторная работа 7,8
5	Факторный анализ	12	4	4	4	Лабораторная работа 9
6	Бинарная логистическая регрессия	12	4	4	4	Лабораторная работа 10
7	Дискриминантный анализ	12	4	4	4	Лабораторная работа 11
	Промежуточная аттестация					зачет
ИТОГО по семестру 3		108	36	36	36	

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 4.

Таблица 4 - Балльно-рейтинговая оценка результатов учебной работы обучающихся по видам (БРС)

Учебная работа (виды)	Сумма баллов	Виды и результаты учебной работы	Оценка в аттестации	Баллы
Текущая учебная работа в семестре (Посещение занятий по расписанию и выполнение заданий)	80	Лабораторная работа №1 (отчет о выполнении контрольной работы)	3 балла (пороговое значение) 5 баллов (максимальное значение)	3-5
		Лабораторная работа №2 (отчет о выполнении контрольной работы)	3 балла (пороговое значение) 5 баллов (максимальное значение)	3-5
		Лабораторная работа №3 (отчет о выполнении контрольной работы)	3 балла (пороговое значение) 5 баллов (максимальное значение)	3-5
		Лабораторная работа №4 (отчет о выполнении контрольной работы)	3 балла (пороговое значение) 5 баллов (максимальное значение)	3-5
		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)	5 балла (пороговое значение) 10 баллов (максимальное значение)	5-10
		Лабораторная работа №8 (отчет о выполнении контрольной работы)	5 балла (пороговое значение) 10 баллов (максимальное значение)	5-10
		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
		Лабораторная работа №11 (отчет о выполнении контрольной работы)	3 балла (пороговое значение) 8 баллов (максимальное значение)	3-8
Итого по текущей работе в семестре				41-80
Промежуточная аттестация (зачет)	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.	3 балла (пороговое значение) 5 баллов (максимальное значение)	3 - 5
		Решение задачи 2.	3 балла (пороговое значение) 5 баллов (максимальное значение)	3 - 5
Итого по промежуточной аттестации (зачет)				10-20
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 5)

Таблица 5 – 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. Кулаичев, А.П. Методы и средства комплексного анализа данных [Электронный ресурс]: учебн. пособие / А.П. Кулаичев. – Электрон.текстовые дан. – Москва: ИНФРА-М, 2016. – Режим доступа: <https://new.znaniyum.com/read?id=252560>

2. Козлов, А.Ю. Статистический анализ данных в MSExcel [Электронный ресурс]: учебн. пособие / А.Ю. Козлов, В.С. Мхитарян, В.Ф. Шишов. – Электрон.текстовые дан. – Москва: ИНФРА-М, 2012. – Режим доступа:

<http://www.znaniium.com/bookread.php?book=238654>

3. Горлач, Б.А. Теория вероятностей и математическая статистика [Электронный ресурс]: Учебник / Б.А. Горлач – Электрон. текстовые дан. – Москва: Лань, 2013. – Режим доступа: <http://e.lanbook.com/reader/book/4864/>

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. Хрущева, И.В. Теория вероятностей [Электронный ресурс]: Учебник / И.В. Хрущева – Электрон. текстовые дан. – Москва: Лань, 2009. – Режим доступа: <http://e.lanbook.com/reader/book/425/>

3. 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/>

4. 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 КГПИ ФГБОУ ВО «КемГУ»:

404 Учебная аудитория для проведения: - занятий лекционного типа; - групповых и индивидуальных консультаций; - текущего контроля и промежуточной аттестации. Специализированная (учебная) мебель: доска меловая, кафедра, столы, стулья. Оборудование: <i>переносное</i> - ноутбук, экран, проектор. Используемое программное обеспечение: MSWindows (MicrosoftImaginePremium 3 year по лицензионному договору № 1212/KMP от 12.12.2018 г. до 12.12.2021 г.), LibreOffice (свободно распространяемое ПО), Яндекс.Браузер (отечественное свободно распространяемое ПО). Интернет с обеспечением доступа в ЭИОС.	Учебный корпус №4. 654079, Кемеровская область, г. Новокузнецк, пр-кт Metallurgov, д. 19
502 Компьютерный класс. Учебная аудитория (мультимедийная) для проведения: - занятий лекционного типа; - занятий семинарского (практического) типа; - занятий лабораторного 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. Лабораторное оборудование: стационарное – компьютеры для обучающихся (16 шт.). Используемое программное обеспечение: MSWindows (MicrosoftImaginePremium 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), Яндекс.Браузер (отечественное свободно распространяемое ПО), MicrosoftVisualStudio (MicrosoftImaginePremium 3 year по лицензионному договору № 1212/KMP от 12.12.2018 г. до 12.12.2021 г.), Среда статистических вычислений Rv.4.0.2 (свободно распространяемое ПО). Интернет с обеспечением доступа в ЭИОС.	Учебный корпус №4. 654079, Кемеровская область, г. Новокузнецк, пр-кт Metallurgov, д. 19

5.3 Современные профессиональные базы данных и информационные справочные системы.

Перечень СПБД и ИСС по дисциплине

1. Общероссийский математический портал (информационная система) - <http://www.mathnet.ru/>
2. Научная электронная библиотека eLIBRARY.RU – крупнейший российский информационный портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты - www.elibrary.ru
3. База данных Science Direct (более 1500 журналов издательства Elsevier, среди них издания по математике и информатике), режим доступа :<https://www.sciencedirect.com>

6 Иные сведения и (или) материалы.

6.1. Примерные вопросы и задания / задачи для промежуточной аттестации

1. Основная и альтернативная гипотеза. Виды гипотез. Статистический критерий значимости, критическая область, ошибки 1-го и 2-го рода, уровень значимости и мощность критерия.
2. Отклонение выборочной средней от заданного значения с известной и неизвестной дисперсией. Исключение выбросов. Отклонение выборочной дисперсии от заданного значения.
3. Сравнение двух дисперсий нормальных генеральных совокупностей. Критерий Фишера. Критическое значение в зависимости от альтернативной гипотезы.
4. Сравнение нескольких дисперсий нормальных распределений. Критерий Бартлетта.
5. Сравнение двух выборочных средних нормальных генеральных совокупностей, дисперсии которых известны и равны; не равны; неизвестны, но предполагаются равными. Практический смысл задачи.
6. Сравнение более двух средних нормальных генеральных совокупностей, дисперсии которых неизвестны и одинаковы.
7. Проверка гипотезы о значимости выборочного коэффициента корреляции двумерной нормальной генеральной совокупности.
8. Проверка гипотезы о нормальном распределении генеральной совокупности. Критерий согласия Пирсона. Методика расчета теоретических частот для показательного, равномерного распределения, биномиального и пуассоновского распределения.
9. Проверка вероятностей, задающих полиномиальное распределение. Проверка независимости двух признаков по таблице сопряженности.
10. Методы проверки 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.
11. Использование приложения «Excel» для проверки статистических гипотез. Многофакторный дисперсионный анализ.
12. Функциональная, статистическая и корреляционная зависимость между двумя случайными величинами. Выборочное уравнение прямой линии регрессии. Выборочный коэффициент корреляции. Свойства. Методика расчета выборочного коэффициента корреляции.
13. Выборочное корреляционное отношение. Свойства. Выборочное корреляционное

- отношение как мера корреляционной связи. Достоинства и недостатки этой меры.
14. Ранговая корреляция Спирмена для двух выборок. Коэффициент конкордации Кендалла для нескольких выборок. Ранговая корреляция Кендалла.
 15. Криволинейная корреляция. Метод выравнивания нелинейной корреляционной зависимости.
 16. Понятие о множественной корреляции. Исследование множественной и нелинейной корреляции с помощью приложения «Excel». Исследование ранговой корреляции и конкордации.
 17. Метод статистических испытаний. Имитация выборок с помощью приложения «Excel».
 18. Многомерный анализ данных
 19. Метод главных компонент
 20. Дискриминантные функции и их классификация
 21. Расчет коэффициентов дискриминантных функций
 22. Классификация при наличии двух и более обучающих выборок.

Практические задания

1. Смоделировать выборку 80 значений 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 фут³/сек.

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
1	15	11	123	190	138	98	88	80	76	49	27	16
2	14	12	115	210	130	45	43	40	35	30	29	17

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. 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.

X	0	1	2	3	4	5
k	10	1	5	1	5	10

Используя критерий Колмогорова – Смирнова, при уровне значимости $\alpha = 0,05$ проверить гипотезу о том, что среди этих монет нет поддельных.

2. В таблице приведены ранги объектов X, Y .

X	1	2	3	4	5	6	7	8	9	10
Y	1	7	3	10	5	2	9	8	4	6

Найти выборочные коэффициенты ранговой корреляции Спирмена и Кендалла, а также проверить гипотезу об отсутствии ранговой корреляции при уровне значимости $\alpha = 0,05$.

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.

Член жюри	Номер пары					
	1	2	3	4	5	6
1	1	2	3	4	5	6
2	1	2	4	5	3	6

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.

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		

1. Смоделировать выборку 100 значений случайной величины X , распределенной нормально с параметрами (20, 10), проверив гипотезу о форме распределения.

2. В стоматологическую поликлинику прибывает простейший поток больных с интенсивностью $\lambda = 5$ чел/час в обычное время дня, в часы «пик» она возрастает до $\lambda_+ = 12$ чел/час, а в часы «спада» достигает величины $\lambda_- = 3$ чел/час. Средняя продолжительность сеанса лечения составляет $\overline{t_{об}} = 0,5$ час/чел. Определить вероятность образования очереди $P_{оч}$ и среднюю длину очереди $\overline{l_{оч}}$ в течение дня, а также необходимое число стоматологов в часы «пик» n_+ и «спада» n_- , обеспечивающих такую же длину очереди и вероятность ее образования, как в обычное время дня.

В. 14

1. В стоматологическую поликлинику прибывает простейший поток больных с интенсивностью $\lambda = 5$ чел/час. Средняя продолжительность сеанса лечения составляет $\overline{t_{об}} = 0,5$ час/чел. Уровень суммарных потерь связан с простым средним числом свободных стоматологов $n_{св}$ и пребыванием среднего числа больных в очереди $L_{оч}$. Построить график зависимости $f(n) = n_{св} + L_{оч}$, где n – количество стоматологов, и определить оптимальное количество стоматологов, при котором уровень суммарных потерь минимален.

2. Фирма «Надежный компьютер», которая реализует компьютеры, стремится оптимизировать затраты, связанные с управлением запасов. Пусть суточный спрос является равномерно распределенной величиной, которая принимает значения $q = 1, \dots, 10$, а вероятности $p(q) = 0,1$ при всех значениях q . Рассматриваются затраты с такими же параметрами, как в примере 1: $\pi = 10$; $c = 3$; $K = 8$; $h = 2$. Найти методом статистических испытаний оценку суточных суммарных затрат для оптимальной стратегии пополнения запасов и исследовать чувствительность этой оценки к параметрам стратегии.

Кейс-задания

Компетенция	Задания
-------------	---------

Компетенция	Задания																										
ОПК-3 Способен применять и модифицировать математические модели для решения задач в области профессиональной деятельности	Кейс-задание 1 Ситуация: Во время исследования двух швейных фабрик была получена следующая таблица:																										
	<table><tr><td>X</td><td>Y</td></tr><tr><td>715</td><td>660</td></tr><tr><td>685</td><td>640</td></tr><tr><td>680</td><td>645</td></tr><tr><td>665</td><td>620</td></tr><tr><td>660</td><td>600</td></tr><tr><td>655</td><td>620</td></tr><tr><td>645</td><td>580</td></tr><tr><td>630</td><td>600</td></tr><tr><td>625</td><td>565</td></tr><tr><td>620</td><td>600</td></tr><tr><td>610</td><td>580</td></tr><tr><td>595</td><td>550</td></tr></table>	X	Y	715	660	685	640	680	645	665	620	660	600	655	620	645	580	630	600	625	565	620	600	610	580	595	550
	X	Y																									
	715	660																									
	685	640																									
	680	645																									
	665	620																									
	660	600																									
	655	620																									
	645	580																									
	630	600																									
	625	565																									
	620	600																									
610	580																										
595	550																										
В ней указан объем производства за последние 12 месяцев.																											
Выясните:																											
1) отличается ли на уровне значимости 0,05 среднее значения производства в месяц на двух фабриках																											
2) отличается ли на уровне значимости 0,05 разброс объема производства																											
Кейс-задание 2 Ситуация: Во время исследования двух швейных фабрик была получена следующая таблица:																											
<table><tr><td>X</td><td>Y</td></tr><tr><td>715</td><td>660</td></tr><tr><td>685</td><td>640</td></tr><tr><td>680</td><td>645</td></tr><tr><td>665</td><td>620</td></tr><tr><td>660</td><td>600</td></tr><tr><td>655</td><td>620</td></tr><tr><td>645</td><td>580</td></tr><tr><td>630</td><td>600</td></tr><tr><td>625</td><td>565</td></tr><tr><td>620</td><td>600</td></tr><tr><td>610</td><td>580</td></tr></table>	X	Y	715	660	685	640	680	645	665	620	660	600	655	620	645	580	630	600	625	565	620	600	610	580			
X	Y																										
715	660																										
685	640																										
680	645																										
665	620																										
660	600																										
655	620																										
645	580																										
630	600																										
625	565																										
620	600																										
610	580																										

Компетенция	Задания		
	595	550	В ней указан объем производства за последние 12 месяцев. Выясните: 1) отличается ли на уровне значимости 0,05 среднее значения производства в месяц на двух фабриках 2) с помощью ранговой корреляции выясните существует ли зависимость объема производства на одной фабрике от другой.

Составитель (и): канд. технн. наук Вячкин Е.С.

(фамилия, инициалы и должность преподавателя (ей))