

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Бойко Елена Григорьевна
Должность: Ректор
Дата подписания: 01.03.2024 12:27:14
Уникальный программный ключ:
e69eb689122030af7d22cc354bf0eb9d453ecf8f


Министерство сельского хозяйства Российской Федерации
ФГБОУ ВО Государственный аграрный университет Северного Зауралья

Инженерно-технологический институт

Кафедра математики и информатики

Утверждаю»

И.о. заведующего кафедрой

 М.В. Виноградова

«06» октября 2020 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Алгоритмы обработки и анализа данных

для направления подготовки 35.03.06 Агроинженерия
Профили: Технические системы в агробизнесе, Технический сервис в
Агропромышленном комплексе

Уровень высшего образования – бакалавриат

Форма обучения: очная, заочная

Тюмень, 2020

При разработке рабочей программы учебной дисциплины в основу положены:

- 1) ФГОС ВО по направлению подготовки 35.03.06 Агроинженерия, утвержденный Министерством образования и науки РФ «23» августа 2017 г. Приказ № 813
- 2) Учебный план основной образовательные программы Технические системы в агробизнесе, Технический сервис в Агропромышленном комплексе, одобрены Ученым советом ФГБОУ ВО ГАУ Северного Зауралья от от «23» сентября 2020 г. Протокол №2

Рабочая программа учебной дисциплины одобрена на заседании кафедры математики и информатики от «06» октября 2020 г. Протокол № 2.1

И.о. заведующего кафедрой  М.В. Виноградова

Рабочая программа учебной дисциплины одобрена методической комиссией института от «24» октября 2020 г. Протокол № 2

Председатель методической комиссии  института О.А. Мелякова

Разработчики:

Виноградова М.В. старший преподаватель кафедры математики и информатики

Директор института:  Г.А. Дорн

4. Содержание дисциплины

4.1. Содержание разделов дисциплины

№ п/п	Наименование раздела дисциплины	Содержание раздела
1.	Алгоритмы и структуры данных	Алгоритмы: построение и анализ. Ресурсная эффективность алгоритмов. Методы оценки ресурсной эффективности алгоритмов. Математический аппарат анализа алгоритмов. Структуры данных. Методы анализа и обработки данных. Первичная обработка данных.
2.	Однофакторная регрессионная модель	Элементы корреляционного анализа. Уравнение парной линейной регрессии. Метод наименьших квадратов (МНК) для нахождения параметров уравнения регрессии. Связь между результативным признаком и признаком – фактором. Показатели качества уравнения регрессии. Проверка гипотез о значимости параметров уравнения, коэффициента корреляции и уравнения регрессии в целом. Нелинейная регрессия, ее виды. Корреляция для нелинейной регрессии. Показатели качества нелинейной регрессии.
3.	Множественная регрессия	Модель множественной регрессии. Традиционный МНК для множественной регрессии. Показатели тесноты связи фактора с результатом: коэффициенты частной эластичности, стандартизированные коэффициенты (β -коэффициенты). Частная корреляция. Оценка значимости множественного уравнения регрессии, его факторов и частных коэффициентов корреляции. Фиктивные переменные.
4.	Пакет Ms Excel «Анализ данных»	Использование пакета Ms Excel «Анализ данных» для нахождения: основных числовых характеристик исследуемых признаков; уравнений регрессии, зависимости результативного признака от одного или большего числа факторов; показателей тесноты связи; показателей качества уравнения регрессии.

4.2. Разделы дисциплины и виды занятий

Очная форма обучения

№ п/п	Наименование раздела дисциплины	Лекц. типа	Семинар. типа	СР	Всего, часов
1	2	3	4	5	6
1.	Алгоритмы и структуры данных	4	-	10	14
2.	Однофакторная регрессионная модель	8	8	18	34
3.	Множественная регрессия	6	8	18	32
4.	Пакет Ms Excel «Анализ данных»	6	8	14	28
	Итого:	24	24	60	108

Заочная форма обучения

№ п/п	Наименование раздела дисциплины	Лекц. типа	Семинар. типа	СР	Всего, часов
1	2	3	4	5	6
1.	Алгоритмы и структуры данных	1	-	18	19
2.	Однофакторная регрессионная модель	2	2	28	32
3.	Множественная регрессия	2	2	28	32
4.	Пакет Ms Excel «Анализ данных»	1	2	22	25
	Итого:	6	6	96	108

4.3. Занятия семинарского типа

№ п/п	№ раздела дисциплины	Тема	Трудоемкость (час)	
			Очная	Заочная
1	2	3	4	5
1.	2	Метод наименьших квадратов (МНК) для нахождения параметров уравнения регрессии.	4	1
2.	2	Проверка гипотез о значимости параметров уравнения, коэффициента корреляции и уравнения регрессии в целом.	2	-
3.	2	Нелинейная регрессия, ее виды. Корреляция для нелинейной регрессии.	2	1
4.	3	Традиционный МНК для множественной регрессии. Показатели тесноты связи фактора с результатом.	4	1
5.	3	Оценка значимости множественного уравнения регрессии, его факторов и частных коэффициентов корреляции.	4	1
6.	4	Использование пакета Ms Excel «Анализ данных».	8	2
Итого:			24	6

4.4. Примерная тематика курсовых проектов (работ) не предусмотрено ОПОП.

5. Организация самостоятельной работы обучающихся по дисциплине

5.1. Типы самостоятельной работы и её контроль

Тип самостоятельной работы	Форма обучения		Текущий контроль
	Очная	Заочная	
Проработка материала лекций, подготовка к занятиям	30	72	Собеседование, тестирование
Самостоятельное изучение тем	6		Собеседование, тестирование
Контрольные работы	24	24	Защита контрольной работы
Всего часов:	60	96	

5.2 Учебно-методические материалы для самостоятельной работы

1. Методы и модели эконометрики. Часть 2. Эконометрика пространственных данных: учебное пособие / О. И. Бантикова, В. И. Васянина, Ю. А. Жемчужникова [и др.]; под редакцией А. Г. Реннер. — Оренбург: Оренбургский государственный университет, ЭБС АСВ, 2015. — 435 с. — ISBN 978-5-7410-1260-4. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/52325.html> (дата обращения: 30.09.2020). — Режим доступа: для авторизир. пользователей.

5.3. Темы, выносимые на самостоятельное изучение

Раздел 1 Алгоритмы и структуры данных. Базовые алгоритмы обработки данных. Нейронные сети. Деревья решений.

Раздел 2 Однофакторная регрессионная модель. Анализ одной и двух нормальных выборок. Методы проверки нормальности.

Раздел 3 Множественная регрессия. Двухфакторный анализ. Связь задач двухфакторного и однофакторного анализа. Таблица двухфакторного анализа.

Раздел 4 Пакет Ms Excel «Анализ данных. Выявление корреляционных зависимостей. Построение и визуализация трендов и регрессионных зависимостей между наборами данных.

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

6.1 Перечень компетенций и оценочные средства индикатора достижения компетенций

Код компетенции	Индикатор достижения компетенции	Перечень планируемых результатов обучения по дисциплине	Наименование оценочного средства
УК-1	ИД-2 _{УК-1} Осуществляет анализ и оценку информации, с использованием различных информационных ресурсов для решения поставленных задач	Знать: - алгоритмы сбора, обработки и анализа данных, необходимых для решения поставленных задач; Уметь: - использовать для решения профессиональных задач интегрированные информационные системы; Владеть: - навыками использования полученных теоретических и практических знаний в профессиональной деятельности.	Тест

6.2. Шкалы оценивания

Шкала оценивания тестирования на зачете

% выполнения задания	Результат
50 – 100	Зачтено
Менее 50	Не зачтено

6.4. Типовые контрольные задания или иные материалы

Указаны в приложении 1.

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

а) основная литература

1. Кувайскова, Ю. Е. Эконометрика: учебное пособие / Ю. Е. Кувайскова. — Ульяновск: Ульяновский государственный технический университет, 2017. — 166 с. — ISBN 978-5-9795-1722-3. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/106132.html>. — Режим доступа: для авторизир. пользователей

2. .

2. Федин, Ф. О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Ф. О.Федин, Ф. Ф. Федин. — Москва: Московский городской педагогический университет, 2012. — 308с. — ISBN 2227-8397. — Текст: электронный // Электронно-библиотечная система IPR BOOKS:[сайт]. — URL: <https://www.iprbookshop.ru/26445.html>. — Режим доступа: для авторизир.пользователей.

б) дополнительная литература

3. Неделько, В. М. Основы статистических методов машинного обучения: учебное пособие / В. М. Неделько. — Новосибирск: Новосибирский государственный технический университет, 2010.— 72 с. — ISBN 978-5-7782-1385-2. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/45418.html>. — Режим доступа: для авторизир.пользователей.

4. Теория и практика машинного обучения: учебное пособие / В. В. Воронина, А. В. Михеев, Н. Г. Ярушкина, К. В. Святков. — Ульяновск: Ульяновский государственный технический университет, 2017. — 291 с. — ISBN 978-5-9795-1712-4. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/106120.html>. — Режим доступа: для авторизир. пользователей.

5. Методы и модели эконометрики. Часть 1. Анализ данных: Учебное пособие / О.И. Бантикова, В.И. Васянина, Ю.А. Жемчужникова, А.Г. Реннер, Е.Н. Седова, О.И. Стебунова, Л.М. Туктамышева, О.С. Чудинова / под ред. А.Г. Реннера; Оренбургский гос. ун-т. - Оренбург: ОГУ, 2015. – 574 с. . — Текст: электронный // Электронно-библиотечная система ЭБС Лань: [сайт]. — <https://e.lanbook.com/book/98113>. — Режим доступа: для авторизир. пользователей.

Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

1. www.mathnet.ru – Общероссийский математический портал
2. www.twirpx.com – Учебно-методическая и профессиональная литература для обучающихся и преподавателей технических, естественнонаучных и гуманитарных специальностей.

3. www.math-net.ru – Журнал «Математическое моделирование»

4. www.exponenta.ru – Образовательный математический сайт

5. www.elementy.ru – Научно-популярный сайт российской фундаментальной науки.

7. www.intuit.ru - видеокурсы «Теория вероятностей и математическая статистика», «Основы теории вероятностей», «Основы математической статистики»,

8. Уравнение нелинейной регрессии [Электронный ресурс] / math.semestr.ru – Режим доступа: <https://math.semestr.ru/corel/noncorel.php>, свободный.

9. Эконометрическая модель как основа эконометрического моделирования. Типы данных и виды переменных в эконометрических исследованиях экономических явлений [Электронный ресурс] / studopedia.ru – Режим доступа: https://studopedia.ru/2_34934_ekonometriceskaya-model-kakosnova-ekonometriceskogo-modelirovaniya-tipi-dannih-i-vidi-peremennih-vekonometriceskih-issledovaniyah-ekonomicheskikh-yavleniy.html, свободный.

10. Регрессия в Excel [Электронный ресурс]: уравнение, примеры. Линейная регрессия/ fb.ru – Режим доступа: <http://fb.ru/article/322644/regressiyav-excel-uravnenie-primeryi-lineynaya-regressiya>, свободный

9. Методические указания для обучающихся по освоению дисциплины

Алгоритмы обработки и анализа данных [Электронный ресурс]: Методические указания к практическим работам для направлений подготовки, реализуемых в ГАУ Северного Зауралья / Автор-сост. М.В. Виноградова. – Тюмень: ГАУ Северного Зауралья, 2022. - 60 с.

10. Перечень информационных технологий

Для выполнения контрольной работы обучающимся рекомендовано использовать Microsoft Office (электронные таблицы Microsoft Excel), ЭИОС Moodle.

11. Материально-техническое обеспечение дисциплины

Для проведения занятий лекционного типа используются специальные помещения - учебные аудитории, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации.

Для проведения практических занятий (занятий семинарского типа), групповых и индивидуальных консультаций используется специальное помещение, укомплектованное персональными компьютерами с операционной системой Windows, пакетом Microsoft Office и др.

Для текущего контроля и промежуточной аттестации используются специальные помещения - учебные аудитории, укомплектованные специализированной мебелью, и (или) компьютерные классы, оснащенные компьютерной техникой с возможностью подключения к сети Интернет и обеспечением доступа в электронную информационно-образовательную среду университета.

Для выполнения самостоятельной работы обучающиеся могут пользоваться читальными залами библиотеки ФГБОУ ВО ГАУ Северного Зауралья, в том числе оснащённых компьютерами с локальной сетью и выходом в интернет.

12. Особенности освоения дисциплины для инвалидов и лиц с ограниченными возможностями здоровья

Обучение обучающихся с ограниченными возможностями здоровья при необходимости осуществляется на основе адаптированной рабочей программы с использованием специальных методов обучения и дидактических материалов, составленных с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся (обучающегося).

В целях освоения учебной программы дисциплины инвалидами и лицами с ограниченными возможностями здоровья обеспечивается:

- для инвалидов и лиц с ограниченными возможностями здоровья по зрению: размещение в доступных для обучающихся, являющихся слепыми или слабовидящими, местах и в адаптированной форме справочной информации о расписании учебных занятий; присутствие ассистента, оказывающего обучающемуся необходимую помощь; выпуск альтернативных форматов методических материалов (крупный шрифт или аудиофайлы), использование версии сайта для слабовидящих Э БС IPR BOOKS и специального мобильного приложения IPR BOOKS WV-Reader (программы не визуального доступа к информации, предназначенной для мобильных устройств, работающих на операционной системе Android и iOS, которая не требует специально обученного ассистента, т.к. люди с ОВЗ по зрению работают со своим устройством привычным способом, используя специальные штатные программы для незрячих людей, с которыми IPR BOOKS WV-Reader имеет полную совместимость);

- для инвалидов и лиц с ограниченными возможностями здоровья по слуху: надлежащими звуковыми средствами воспроизведение информации;

- для инвалидов и лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата: возможность беспрепятственного доступа обучающихся в учебные помещения, туалетные комнаты и другие помещения кафедры, а также пребывание в указанных помещениях.

Образование обучающихся с ограниченными возможностями здоровья может быть организовано как совместно с другими обучающимися, так и в отдельных группах или в отдельных организациях.

Приложение 1

Министерство сельского хозяйства Российской Федерации
ФГБОУ ВО Государственный аграрный университет Северного Зауралья
Инженерно-технологический институт
Кафедра математики и информатики

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по учебной дисциплине **Алгоритмы обработки и анализа данных**

для направления подготовки 35.03.06 Агроинженерия

Образовательные программы Технические системы в агробизнесе,
Технический сервис в Агропромышленном комплексе

Уровень высшего образования – бакалавриат

Разработчик: Виноградова М.В., старший преподаватель кафедры
математики и информатики

Утверждено на заседании кафедры
протокол № 2.1 от «06» октября 2020г.

И.о. заведующего кафедрой  М.В. Виноградова

Тюмень, 2020

**КОНТРОЛЬНЫЕ ЗАДАНИЯ И ИНЫЕ МАТЕРИАЛЫ ОЦЕНКИ
знаний, умений, навыков и (или) опыта деятельности, характеризующие этапы
формирования компетенций в процессе освоения дисциплины**

АЛГОРИТМЫ ОБРАБОТКИ И АНАЛИЗА ДАННЫХ

Вопросы к зачёту

	Вопросы
<p align="center">УК-1</p> <p>Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</p>	<ol style="list-style-type: none"> 1. Алгоритмы: построение и анализ. 2. Базовые алгоритмы обработки данных. 3. Нейронные сети. 4. Деревья решений. 5. Ресурсная эффективность алгоритмов. Методы оценки ресурсной эффективности алгоритмов. 6. Математический аппарат анализа алгоритмов. 7. Структуры данных. 8. Методы анализа и обработки данных. 9. Первичная обработка данных. 10. Анализ одной и двух нормальных выборок. 11. Элементы корреляционного анализа. Уравнение парной линейной регрессии. 12. Метод наименьших квадратов (МНК) для нахождения параметров уравнения регрессии. Связь между результативным признаком и признаком – фактором. 13. Показатели качества уравнения регрессии. 14. Проверка гипотез о значимости параметров уравнения, коэффициента корреляции и уравнения регрессии в целом. 15. Методы проверки нормальности. 16. Нелинейная регрессия, ее виды. Корреляция для нелинейной регрессии. Показатели качества нелинейной регрессии. 17. Однофакторный анализ. Непараметрические критерии проверки однородности. 18. Оценивание эффектов обработки (непараметрический подход). 19. Двухфакторный анализ. Связь задач двухфакторного и однофакторного анализа. 20. Таблица двухфакторного анализа. 21. Двухфакторный дисперсионный анализ. 22. Выявление корреляционных зависимостей. 23. Построение и визуализация трендов и регрессионных зависимостей между наборами данных. 24. Модель множественной регрессии. Традиционный МНК для множественной регрессии. 25. Показатели тесноты связи фактора с результатом: коэффициенты частной эластичности, стандартизированные коэффициенты (β-коэффициенты). 26. Частная корреляция. 27. Оценка значимости множественного уравнения регрессии, его факторов и частных коэффициентов корреляции. 28. Фиктивные переменные. 29. Использование пакета Ms Excel «Анализ данных» для нахождения: <ul style="list-style-type: none"> • основных числовых характеристик исследуемых признаков; • уравнений регрессии, зависимости результативного признака от одного или большего числа факторов; • показателей тесноты связи; • показателей качества уравнения регрессии.

Процедура оценивания зачета

Вопросы к зачету обучающиеся получают в течение первой недели начала изучения дисциплины.

Обучающийся допускается к промежуточной аттестации (зачету) по дисциплине в форме тестирования при условии посещения занятий и успешного выполнения текущего контроля в течение семестра, которое включает:

- положительные оценки за промежуточные контрольные работы (очная форма обучения);
- успешное собеседование по темам, выносимым на самостоятельное изучение;

В противном случае обучающийся не допускается к прохождению тестовых заданий, до полной ликвидации всех задолженностей.

Тест в системе ЭИОС Moodle включает 30 вопросов, в случайном порядке выбранных из банка вопросов. Обучающемуся предоставляется 2 попытки продолжительностью 45 минут каждая.

Критерии оценки зачета

Если зачет проводится в форме тестирования в ЭИОС Moodle, то оценка «Зачтено/Не зачтено» выставляется системой автоматически согласно шкале оценивания тестирования на зачете.

Если зачет проводится в форме итогового собеседования, то:

Оценка «Зачтено» выставляется, если обучающийся посещал занятия, успешно выполнил текущий контроль и успешно прошел итоговое собеседование.

Оценка «Не зачтено» выставляется, если обучающийся не был допущен к промежуточной аттестации или не прошел итоговое собеседование.

Комплект заданий для контрольной работы (очная и заочная формы обучения)

Тема «Алгоритмы и структуры данных»

Вариант 1. Известно количество операций, выполняемых двумя алгоритмами, решающими одну задачу:

$$T_1(n) = 90n^2 + 201n + 2000$$

$$T_2(n) = 2n^3 + 3$$

Какой из алгоритмов предпочтительнее использовать на практике? Представить графическое обоснование.

Вариант 2. Показать, что функция временной сложности $f(n) = \frac{1}{2}n^2 - 3n = \Theta(n^2)$

Вариант 3. Показать, что функция временной сложности $f(n) = 3n^2 + 5n \log(n) + 7n = \Theta(n^2)$

Вариант 4. Правило суммирования используется для вычисления времени последовательного выполнения программных фрагментов. Пусть имеется три фрагмента $\Theta(n^2)$, $\Theta(n^3)$, $\Theta(n \oplus \log(n))$. Найти $T_1(n) + T_2(n) + T_3(n)$.

Вариант 5. Показать, что функция $T(n) = 3 \cdot n^3 + 2 \cdot n^2$ имеет верхнюю оценку $\Theta(n^3)$.

Вариант 6. Пусть имеются два фрагмента со временем выполнения $\Theta(f(n))$ и $\Theta(g(n))$, где

$$f(n) = \begin{cases} n^4, & \text{если } n \text{ — четно} \\ n^2, & \text{если } n \text{ — нечетно} \end{cases} \quad g(n) = \begin{cases} n^2, & \text{если } n \text{ — четно} \\ n^3, & \text{если } n \text{ — нечетно} \end{cases}$$

Найти $T(n)$.

Вариант 7. Даны следующие функции от n :

$$f_1(n) = n^2$$

$$f_2(n) = n^2 + 1000n$$

$$f_3(n) = \begin{cases} n, & \text{если } n - \text{нечетно} \\ n^3, & \text{если } n - \text{четно} \end{cases}$$

$$f_4(n) = \begin{cases} n, & \text{если } n > 100 \\ n^3, & \text{если } n \leq 100 \end{cases}$$

Указать для каждой пары функций, когда $f_i(n)$ имеет порядок роста $\Theta(f_j(n))$ и когда $f_i(n)$ есть $\Omega(f_j(n))$.

Вариант 8. Найти порядок $T(n)$, если $T(1) = 1$:

- ✓ $T(n) = T(n/2) + n$;
- ✓ $T(n) = T(n-1) + n^2$;
- ✓ $T(n) = n \cdot T(n-1)$.

Вариант 9. Оцените количество операций для алгоритмов:

- поиска всех делителей числа;
- нахождения минимального и максимального элементов массива;
- определения количества положительных элементов массива;
- проверки числа на простоту.

В каждом случае опишите набор используемых элементарных операций. Определите асимптотическую сложность этих алгоритмов.

Вариант 10. Выполнить анализ временной сложности следующих алгоритмов:

- линейного поиска с барьером;
- сортировка методом простого обмена;
- сортировка методом простого выбора.

Тема «Однофакторная регрессионная модель».

Задание 1. По 10 сельскохозяйственным предприятиям имеются данные о себестоимости молока и средней продуктивности молока (табл. 1 и 2).

Таблица 1

Себестоимость молока, руб./л

Сельскохозяйственное предприятие	Предпоследняя цифра номера зачетной книжки									
	0	1	2	3	4	5	6	7	8	9
1	7,5	6,0	5,2	8,3	5,8	6,9	7,8	7,0	5,9	8,0
2	6,0	5,2	8,3	5,8	6,9	7,8	7,0	5,9	8,0	7,5
3	5,2	8,3	5,8	6,9	7,8	7,0	5,9	8,0	7,5	6,0
4	8,3	5,8	6,9	7,8	7,0	5,9	8,0	7,5	6,0	5,2
5	5,8	6,9	7,8	7,0	5,9	8,0	7,5	6,0	5,2	8,3
6	6,9	7,8	7,0	5,9	8,0	7,5	6,0	5,2	8,3	5,8
7	7,8	7,0	5,9	8,0	7,5	6,0	5,2	8,3	5,8	6,9
8	7,0	5,9	8,0	7,5	6,0	5,2	8,3	5,8	6,9	7,8
9	5,9	8,0	7,5	6,0	5,2	8,3	5,8	6,9	7,8	7,0
10	8,0	7,5	6,0	5,2	8,3	5,8	6,9	7,8	7,0	5,9

Таблица 2

Средняя продуктивность молока, кг

Сельскохозяйственное предприятие	Последняя цифра номера зачетной книжки									
	0	1	2	3	4	5	6	7	8	9
1	187	153	121	246	139	175	225	206	141	229
2	158	120	220	138	174	222	205	140	228	192
3	125	244	137	173	201	202	139	227	191	163
4	249	171	172	220	221	136	226	161	162	130
5	141	136	219	200	135	223	189	190	129	254
6	176	218	199	134	222	157	160	128	253	146
7	223	198	133	221	185	186	143	252	145	181
8	203	132	245	184	156	124	251	144	207	208
9	137	219	183	122	123	248	127	179	227	228
10	224	182	154	155	247	140	178	226	180	142

Требуется:

1. Рассчитать параметры уравнения парной линейной регрессии зависимости себестоимости молока от средней продуктивности.
2. Оценить качество уравнения с помощью средней ошибки аппроксимации.
3. Найти средний (обобщающий) коэффициент эластичности.
4. Оценить тесноту связи с помощью показателей корреляции и детерминации.
5. Оценить значимость коэффициента корреляции через t-критерий Стьюдента при $\alpha = 0,05$.
6. Оценить статистическую надежность результатов регрессионного анализа с помощью F-критерия Фишера при $\alpha = 0,05$.
7. Рассчитать прогнозное значение результата, если прогнозное значение фактора увеличится на 10 % от его среднего уровня.

Задание 2. В течение шести лет использовались три различных технологии по выращиванию сельскохозяйственной культуры. Данные по эксперименту (в ц/га) приведены в таблице. При уровне значимости $\alpha = 0,05$ методом дисперсионного анализа установить влияние различных технологий на урожайность культуры.

Вариант 1.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,28	0,32	0,36
2	0,34	0,36	0,38
3	0,30	0,29	0,31
4	0,36	0,35	0,34
5	0,34	0,32	0,35
6	0,32	0,30	0,33

Вариант 2.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,38	0,32	0,36
2	0,44	0,41	0,45
3	0,32	0,33	0,40
4	0,31	0,36	0,38
5	0,44	0,42	0,37
6	0,36	0,34	0,33

Вариант 3.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,18	0,24	0,36
2	0,28	0,36	0,12
3	0,12	0,28	0,22
4	0,14	0,40	0,45
5	0,32	0,16	0,40
6	0,25	0,30	0,28

Вариант 4.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,38	0,20	0,21
2	0,36	0,24	0,22
3	0,35	0,26	0,31
4	0,31	0,30	0,34
5	0,29	0,28	0,30
6	0,30	0,31	0,29

Вариант 5.

Вариант 6.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,10	0,17	0,14
2	0,15	0,15	0,18
3	0,14	0,25	0,30
4	0,18	0,22	0,27
5	0,20	0,30	0,34
6	0,16	0,28	0,40

Вариант 7.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,32	0,26	0,29
2	0,31	0,31	0,26
3	0,31	0,27	0,30
4	0,28	0,29	0,31
5	0,33	0,27	0,30
6	0,30	0,26	0,28

Вариант 9.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,36	0,34	0,30
2	0,38	0,35	0,32
3	0,34	0,32	0,36
4	0,38	0,37	0,29
5	0,35	0,29	0,31
6	0,36	0,31	0,34

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,20	0,16	0,14
2	0,25	0,15	0,20
3	0,25	0,26	0,28
4	0,28	0,22	0,25
5	0,20	0,30	0,32
6	0,19	0,25	0,38

Вариант 8.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,36	0,30	0,29
2	0,34	0,34	0,28
3	0,32	0,35	0,30
4	0,33	0,32	0,28
5	0,35	0,29	0,31
6	0,33	0,31	0,35

Вариант 10.

Год	Технология (фактор)		
	F ₁	F ₂	F ₃
1	0,27	0,31	0,34
2	0,34	0,35	0,33
3	0,31	0,29	0,31
4	0,36	0,34	0,33
5	0,33	0,32	0,31
6	0,32	0,30	0,32

Тема «Множественная регрессия».

Задание. По 30 сельскохозяйственным предприятиям имеются данные о средних значениях и вариации урожайности зерновых, количестве внесенных удобрений и насыщенности севооборота зерновыми, а также о значениях коэффициентов парной корреляции между этими признаками (табл. 3 и 4).

Таблица 3

Урожайность зерновых, количество внесенных удобрений и продолжительность вегетационного периода

Предпоследняя цифра номера зачетной книжки	Урожайность зерновых, ц/га		Внесено органических удобрений, ц/га		Насыщенность севооборота зерновыми, %	
	Среднее значение	Среднее квадратическое отклонение	Среднее значение	Среднее квадратическое отклонение	Среднее значение	Среднее квадратическое отклонение
	\bar{y}	σ_y	\bar{x}_1	σ_{x_1}	\bar{x}_2	σ_{x_2}
0	12,5	5,0	10	1,5	45	8,0
1	13,5	5,4	12	1,7	48	8,5
2	15,0	5,8	14	1,9	51	9,0
3	16,0	6,2	16	2,1	54	9,5
4	16,5	6,6	18	2,3	57	10,0

5	17,0	7,0	20	2,5	60	10,5
6	17,5	7,4	22	2,7	63	11,0
7	18,5	7,8	24	2,9	66	11,5
8	21,5	8,2	26	3,1	69	12,0
9	22,0	8,6	28	3,3	72	12,5

Таблица 4

Линейные коэффициенты парной корреляции

Последняя цифра номера зачетной книжки	r_{yx_1}	r_{yx_2}	$r_{x_1x_2}$
0	0,33	0,48	0,04
1	0,34	0,46	0,05
2	0,35	0,44	0,06
3	0,36	0,42	0,07
4	0,37	0,40	0,08
5	0,38	0,38	0,09
6	0,39	0,36	0,10
7	0,40	0,34	0,11
8	0,41	0,32	0,12
9	0,42	0,30	0,13

Требуется:

1. Построить уравнение множественной линейной регрессии зависимости урожайности зерновых от количества внесенных удобрений и насыщенности севооборота зерновыми.
2. Определить линейный коэффициент множественной корреляции.
3. Рассчитать общий F-критерий Фишера при уровне значимости $\alpha = 0,05$.

Тема «Пакет Ms Excel «Анализ данных»».

Вариант 1.

По 19 предприятиям оптовой торговли изучается зависимость объема реализации (y) от размера торговой площади (x_1) и товарных запасов (x_2). Были получены следующие варианты уравнений регрессии:

1. $y = 25 + 15x_1$, $r^2 = 0,90$;
2. $y = 42 + 27x_1$, $r^2 = 0,84$;
3. $y = 30 + 10x_1 + 8x_2$, $R^2 = 0,92$; (2,5) (4,0)
4. $y = 21 + 14x_1 + 20x_2 + 0.6x_2^2$, $R^2 = 0,95$; (5,0) (12,0) (0,2)

В скобках указаны значения стандартных ошибок для коэффициентов регрессии.

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Проанализируйте тесноту связи результата с каждым из факторов.
2. Выберите наилучшее уравнение регрессии, обоснуйте принятое решение.

Вариант 2.

По 25 территориям страны изучается влияние климатических условий на урожайность зерновых y (ц/га). Для этого были отобраны две объясняющие переменные:

x_1 – количество осадков в период вегетации (мм);

x_2 – средняя температура воздуха (0 C).

Матрица парных коэффициентов корреляции этих показателей имеет следующий вид:

	y	x₁	x₂
y	1,0		
x₁	0,6	1,0	
x₂	-0,5	-0,9	1,0

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Определите частные коэффициенты корреляции результата с каждым из факторов. Прокомментируйте различие полученных парных и частных коэффициентов корреляции результатов

2. Исследователь, анализирующий данную зависимость, намерен определить на основе приведенной выше матрицы, какое уравнение регрессии лучше строить:

- А) парную линейную регрессию y на x_1 ;
- Б) парную линейную регрессию y на x_2 ;
- В) множественную линейную регрессию.

Как бы вы ответили на эти вопросы?

3. Постройте уравнение регрессии в стандартизованном масштабе и сделайте выводы.

Вариант 3.

По 30 наблюдениям матрица парных коэффициентов корреляции оказалась следующей:

	y	x₁	x₂	x₃
y	1,00			
x₁	0,30	1,00		
x₂	0,60	0,10	1,00	
x₃	0,40	0,15	0,80	1,00

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Постройте уравнение регрессии в стандартизованном виде и сделайте выводы.

2. Определите показатель множественной корреляции.

3. Оцените целесообразность включения переменной x_1 в модель после введения в нее переменных x_2 и x_3 .

Вариант 4.

По 30 предприятиям отрасли были получены следующие результаты регрессионного анализа зависимости объема выпуска продукции y (млн. руб.) от численности занятых на предприятии x_1 (чел.) и среднегодовой стоимости основных фондов x_2 (млн. руб.):

Коэффициент детерминации	???
Множественный коэффициент корреляции	0,85
Уравнение регрессии	$y = ??? + 0,48x_1 + 20x_2$
Стандартные ошибки параметров	2 0,06 ???
T-критерий для параметров	1,5 ??? 4

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Восстановите пропущенные характеристики.

2. С вероятностью 0,95 построьте доверительные интервалы для коэффициентов регрессии.

3. Проанализируйте результаты регрессионного анализа.

Вариант 5.

Зависимость потребления электроэнергии y (тыс. кВт×ч) от объемов производства продукции А – x_1 (тыс. ед.) и продукции Б – x_2 (тыс. ед.) характеризуется следующим образом:

Уравнение регрессии в стандартизованном виде	$t_y \approx 0,79t_{x_1} + 0,56t_{x_2}$
--	---

Коэффициент детерминации	0,95
Коэффициент вариации y	27%
Коэффициент вариации x_1	45%
Коэффициент вариации x_2	40%

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Сделайте выводы о силе влияния факторов на результат.
2. Учитывая значения коэффициентов вариации рассматриваемых признаков, определите частные коэффициенты эластичности, сделайте по ним выводы.
3. Оцените значимость уравнения регрессии, учитывая, что оно построено по 30 наблюдениям.

Вариант 6.

Имеется информация по 25 наблюдениям таблицы, приведенной ниже

Таблица

Признак	Среднее значение	Коэффициент вариации, %	Уравнение регрессии
y	35	20	$\hat{y} = 20 + x_1 - 2,0 \cdot x_2$
x_1	16	30	$\hat{y} = 9 + 1,1 \cdot x_1$
x_2	8	10	$\hat{y} = 4 - 4,1 \cdot x_2$

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Оцените значимость каждого уравнения регрессии, если известно, что $r_{x_1 x_2} = -0,35$.
2. Оцените значимость коэффициентов регрессии уравнения с двумя объясняющими переменными.
3. Определите показатели частной корреляции.
4. Найдите частные коэффициенты эластичности.

Вариант 7.

По совокупности 30 предприятий концерна изучается зависимость прибыли y (тыс. руб.) от выработки продукции на одного работника x_1 (ед.) и индекса цен на продукцию x_2 (%). Данные приведены в таблице

Таблица

Признак	Среднее значение	Среднее квадратическое отклонение	Парный коэффициент корреляции
y	250	38	$r_{yx_1} = 0,68$
x_1	47	12	$r_{yx_2} = 0,63$
x_2	112	21	$r_{x_1 x_2} = 0,42$

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Постройте линейные уравнения парной регрессии, оцените их значимость с помощью F -критерия Фишера.
2. Найдите уравнение множественной регрессии в стандартизованном и натуральном масштабе.
3. Рассчитайте множественный коэффициент корреляции, общий и частные критерии Фишера и сделайте выводы.

Вариант 8.

Изучается зависимость по 25 предприятиям концерна потребления материалов y (т) от энерговооруженности труда x_1 (кВт×ч на одного рабочего) и объема произведенной продукции x_2 (тыс. ед.). Данные приведены в таблице.

Таблица

Признак	Среднее значение	Среднее квадратическое отклонение	Парный коэффициент корреляции
y	12,0	2,0	$r_{yx_1} = 0,52$
x_1	4,3	0,5	$r_{yx_2} = 0,84$
x_2	10,0	1,8	$r_{x_1x_2} = 0,43$

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Постройте уравнение множественной регрессии и поясните экономический смысл его параметров.
2. Определите частные коэффициенты эластичности и стандартизованные коэффициенты регрессии.
3. Найдите частные и множественный коэффициенты корреляции.
4. Оцените значимость уравнения регрессии с помощью F -критерия Фишера.

Вариант 9.

По 50 семьям изучалось потребление мяса – y (кг на душу населения) от дохода – x_1 (руб. на одного члена семьи) и от потребления рыбы x_2 (кг на душу населения). Результаты оказались следующими:

Уравнение регрессии	$\hat{y} = -180 + 0,2 \cdot x_1 - 0,4 \cdot x_2$
Стандартные ошибки параметров	20 0,01 0,25
Множественный коэффициент корреляции	0,85

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Используя T -критерий Стьюдента, оцените значимость параметров уравнения.
2. Рассчитайте F -критерий Фишера.
3. Оцените по частным F -критериям Фишера целесообразность включения в модель:
А) фактора x_1 после фактора x_2 ;
Б) фактора x_2 после фактора x_1 .

Задача 10

По 40 предприятиям одой отрасли исследовалась зависимость производительности труда – y от уровня квалификации рабочих – x_1 и энерговооруженности их труда – x_2 . Результаты оказались следующими:

Уравнение регрессии	$\hat{y} = a + 10 \cdot x_1 + 2 \cdot x_2$
Стандартные ошибки параметров	0,5 2 ?
T -критерий для параметров	3 ? 5
Множественный коэффициент корреляции	0,85

Задание: с использованием средств Пакета Ms Excel «Анализ данных»:

1. Определите параметр a и заполните пропущенные значения.
2. Оцените значимость уравнения в целом, используя значения множественного коэффициента корреляции.
3. Какой из факторов оказывает более сильное воздействие на результат?

Процедура оценивания контрольных работ

Для каждой контрольной работы разработаны карточки по вариантам, с целью исключить возможность списывания и возможности объективно оценить работу обучающегося. При оценке уровня выполнения контрольной работы, в соответствии с поставленными целями и задачами для данного вида учебной деятельности установлены следующие критерии:

- умение самостоятельно осмысливать проблему на основе существующих методик;
- умение логично и грамотно излагать собственные умозаключения и выводы;
- умение анализировать и обобщать материал.

Критерии оценки

По результатам выполнения контрольных работ обучающемуся выставляются оценка:

«отлично», если работа выполнена полностью; в логических рассуждениях и обосновании решения нет пробелов и ошибок; в решении нет математических ошибок (возможна одна неточность, описка, которая не является следствием незнания или непонимания учебного материала).

«хорошо», если работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки); допущены одна ошибка или есть два – три недочёта в выкладках, рисунках, чертежах или графиках (если эти виды работ не являлись специальным объектом проверки).

«удовлетворительно», если допущено более одной ошибки или более двух – трех недочетов в выкладках, чертежах или графиках, но учащийся обладает обязательными умениями по проверяемой теме.

«неудовлетворительно», если допущены существенные ошибки, показавшие, что учащийся не обладает обязательными умениями по данной теме в полной мере.

Вопросы к собеседованию по темам, выносимым на самостоятельное изучение

Раздел 1 Алгоритмы и структуры данных.

1. Базовые алгоритмы обработки данных.
2. Нейронные сети.
3. Деревья решений.

Раздел 2 Однофакторная регрессионная модель.

1. Анализ одной и двух нормальных выборок.
2. Методы проверки нормальности.
3. Однофакторный анализ.
4. Непараметрические критерии проверки однородности.
5. Оценка эффектов обработки (непараметрический подход).

Раздел 3 Множественная регрессия.

1. Двухфакторный анализ. Связь задач двухфакторного и однофакторного анализа.
2. Таблица двухфакторного анализа.
3. Двухфакторный дисперсионный анализ.

Раздел 4 Пакет Ms Excel «Анализ данных».

1. Выявление корреляционных зависимостей.
2. Построение и визуализация трендов и регрессионных зависимостей между наборами данных.

Процедура оценивания собеседования

Собеседование проводится в форме индивидуального опроса для определения уровня освоенности обучающимися тем, выносимых на самостоятельное изучение. Как правило, собеседование проходит на консультации. Преподаватель заранее предупреждает обучающегося о сроках проведения собеседования, требованиях к подготовке материалов,

знакомит с вопросами к собеседованию.

Ответ на вопрос считается правильным, если по своему содержанию полностью соответствует заданному вопросу, содержит все необходимые теоретические факты и обоснованные выводы, а его изложение математически грамотны.

Критерии оценки

По результатам собеседования студенту выставляется «Зачтено», если он правильно, полно и аргументировано, отвечает на вопросы, демонстрирует знание вопроса и самостоятельность мышления, или «Не зачтено», если он не конкретно, слабо аргументировано и не убедительно, отвечает на вопросы, либо не отвечает совсем, демонстрирует незнание вопроса.