

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**федеральное государственное автономное
образовательное учреждение высшего образования_
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Институт информационных технологий, математики и механики

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Рабочая программа дисциплины

Математическое моделирование и прикладной анализ данных

Уровень высшего образования

Бакалавриат

Направление подготовки / специальность

01.03.02 - Прикладная математика и информатика

Направленность образовательной программы

Прикладная математика и информатика (общий профиль)

Форма обучения

очная

г. Нижний Новгород

2024 год начала подготовки

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.ДВ.11.01 Математическое моделирование и прикладной анализ данных относится к части, формируемой участниками образовательных отношений образовательной программы.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства	
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	Для текущего контроля успеваемости	Для промежуточной аттестации
ПК-4: Способен применять методы математического и компьютерного исследования при анализе задач на основе знаний фундаментальных математических и компьютерных наук	<p>ПК-4.1: Знает фундаментальные и теоретические основы, необходимые для исследования научных проблем</p> <p>ПК-4.2: Умеет самостоятельно применять полученные знания для анализа объекта исследования, определять цели и задачи исследования, а также выбирать корректный метод исследования научной проблемы</p> <p>ПК-4.3: Имеет практический опыт научно-исследовательской деятельности, а именно решения научных задач в соответствии с поставленной целью и выбранной методикой</p>	<p>ПК-4.1: Знает базовые математические модели регрессионного и факторного анализа данных. Знает понятийный аппарат, математические свойства (с доказательством) и методы отбора переменных. Знает спектр прикладных задач статистического вывода и примеры комбинированного применения методов.</p> <p>ПК-4.2: На примере модельных задач умеет формулировать цели исследования, строить регрессионные и факторные модели, подбирать модели статистического вывода, анализировать протоколы обработки данных в статистических пакетах, формулировать результаты. Имеет навыки оценки программных средств анализа данных, необходимые для развития таких программных средств и направлений их разработки.</p> <p>ПК-4.3: Владеет навыками самостоятельного выполнения проектов с</p>	<p>Практическое задание</p> <p>Проект</p> <p>Собеседование</p> <p>Тест</p>	<p>Зачёт:</p> <p>Контрольные вопросы</p>

		использованием пакетов прикладных программ, включая выбор цели исследования, построение математической модели, проверку ее качества, анализ результатов.		
--	--	--	--	--

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

	очная
Общая трудоемкость, з.е.	2
Часов по учебному плану	72
в том числе	
аудиторные занятия (контактная работа):	
- занятия лекционного типа	24
- занятия семинарского типа (практические занятия / лабораторные работы)	12
- КСР	1
самостоятельная работа	35
Промежуточная аттестация	0 Зачёт

3.2. Содержание дисциплины

(структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий)

Наименование разделов и тем дисциплины	Всего (часы)	в том числе			
		Контактная работа (работа во взаимодействии с преподавателем), часы из них			Самостоятельная работа обучающегося, часы
		Занятия лекционного типа	Занятия семинарского типа (практические занятия/лабораторные работы), часы	Всего	
	0 Ф 0	0 Ф 0	0 Ф 0	0 Ф 0	0 Ф 0
Тема 1. Примеры задач прогнозирования и анализа взаимосвязей. Инструменты описательной статистики. Модели снижения размерности.	4	2		2	2
Тема 2. Математическая модель факторного анализа. Аксиоматика, оценка параметров, оценка значений уровней факторов, регрессионный подход.	12	4	2	6	6
Тема 3. Основы и практика регрессионного анализа. Аксиоматика моделей простой линейной регрессии. Оценка параметров. Проверка значимости фактора. Построение доверительных интервалов, интервалов прогнозирования.	11	4	2	6	5
Тема 4. Аксиоматика моделей множественной линейной регрессии.	11	4	2	6	5

Оценка параметров. Проверка значимости факторов и значимости в целом. Проверка выполнения предположений.					
Тема 5. Статистические проверки выполнения предположений регрессионного анализа. Введение в проблемы построения моделей (избыточность, неоднородная дисперсия ошибок, автокорреляция). Виды и критерии пошаговых методов отбора переменных.	12	4	2	6	6
Тема 6. Модельные задачи статистического вывода, параметрические и непараметрические методы.	12	4	2	6	6
Тема 7. Логлинейные модели для анализа бинарных признаков. Logit-модели. Примеры методов анализа временных рядов. Таблицы дожития. Статистическое моделирование выживаемости.	9	2	2	4	5
Аттестация	0				
КСР	1			1	
Итого	72	24	12	37	35

Содержание разделов и тем дисциплины

Тема 1. Примеры задач прогнозирования и анализа взаимосвязей. Среднее, дисперсия и ковариация. Статистическая связь фактора и отклика. Принцип наименьших квадратов. Наклон и корреляция. Свойства остатков и оценочных значений. Правило разложения вариации, коэффициент детерминации и корреляция фактора с откликом. Качество модели, визуальный анализ остатков. Снижение размерности задачи методами факторного анализа, кластерного анализа, методом главных компонент.

Тема 2. Математическая модель факторного анализа. Свойства матриц факторных нагрузок. Свойства общностей и характеристик. Обоснование вращения факторов. Критерии качества решения. Схема оценки факторных нагрузок. Методы вращения. Интерпретация решения. Пример модельной задачи. Оценка уровней факторов, регрессионный подход.

Тема 3. Аксиоматика моделей простой линейной регрессии. Оценки, остатки и ошибки. Теорема Гаусса-Маркова. Оценка неизвестной дисперсии ошибок. Проверка значимости (t-статистика). Стандартизованные бета-коэффициенты. Доверительные интервалы и прогноз.

Тема 4. Аксиоматика моделей множественной линейной регрессии, запись модели в скалярной и векторной форме. Оценки, остатки и ошибки. Свойства оценочных коэффициентов. Оценка неизвестной дисперсии ошибок. Проверка значимости в целом (F-статистика), значимости факторов (t-статистика), доверительные интервалы и прогноз. Примеры модельных задач.

Тема 5. Мониторинг построения модели. Избыточность, ее последствия и диагностика. Влиятельность и значимость фактора. Метод исключения переменных. Метод включения переменных. Пошаговый метод с включением. Критические значения и вероятности. Проблемы построения моделей: гетероскедастичные ошибки и автокорреляция. Проверка Голдфелда-Квандта. Проверка Дарбина-Уотсона. Логарифмические преобразования, метод Кокрейна и взвешенный МНК. Примеры модельных задач.

Тема 6. Модельные задачи статистического вывода. Сравнение дисперсий двух совокупностей. Ранговый дисперсионный анализ и дисперсионный анализ числовых переменных (связанные и несвязанные выборки). Критерий независимости и критерий согласия "хи-квадрат" для номинальных переменных. Сравнение средних в двух совокупностях (параметрические и непараметрические методы), коэффициент ранговой корреляции Спирмена.

Тема 7. Логлинейные модели для анализа бинарных признаков. Logit-модели. Примеры методов анализа временных рядов. Приложения: Таблицы дожития. Статистическое моделирование выживаемости. Примеры модельных задач.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа обучающихся включает в себя подготовку к контрольным вопросам и заданиям для текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведенным в п. 5.

Для обеспечения самостоятельной работы обучающихся используются:

- электронный курс "Методы моделирования социально-экономических процессов: регрессионный анализ" (<https://e-learning.unn.ru/course/view.php?id=793>).

Иные учебно-методические материалы: 1. Садовничий В.А. Большие данные в современном мире. (15 ноября 2017 года ректор МГУ имени М.В. Ломоносова академик В.А. Садовничий прочитал перед студентами, аспирантами и сотрудниками первую лекцию в рамках нового межфакультетского учебного курса). Видеозапись <http://media.msu.ru/?p=16379>

2. Стронгина Н.Р., Марчева И.А. Регрессионный анализ в экономических приложениях: Учебное пособие. Н. Новгород: Изд-во ННГУ, 2002. – 214 с.

3. Балабанов А.С., Стронгина Н.Р. Анализ данных в экономических приложениях. Учебное пособие. Н. Новгород: Изд-во ННГУ, 2004. – 136 с.

4. Балабанов А.С., Власова О.В., Стронгина Н.Р. Статистический вывод в экономических приложениях. Н. Новгород: Изд-во ННГУ, 2002. - 67 с.

5. Кузнецова С.В., Стронгина Н.Р. Управление портфелем ценных бумаг: модели Марковица и Тобина. Учебное пособие. - Н. Новгород: Изд-во ННГУ, 2002. – 102 с.

6. Стронгина Н.Р. Элементы теории стабильного населения. Учебное пособие. Н. Новгород: Изд-во ННГУ, 2002. – 80 с.

7. Балабанов А.С., Стронгина Н.Р. Анализ данных в экономических приложениях. Компьютерный практикум в SPSS. Н. Новгород: Изд-во ННГУ, 2003. – 100 с.

8. Стронгина Н.Р. Методы моделирования социально-экономических процессов: анализ данных. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 814Е.14.08.
<https://e-learning.unn.ru/course/view.php?id=814>

5. Фонд оценочных средств для текущего контроля успеваемости и промежуточной аттестации по дисциплине (модулю)

5.1 Типовые задания, необходимые для оценки результатов обучения при проведении текущего контроля успеваемости с указанием критериев их оценивания:

5.1.1 Типовые задания (оценочное средство - Практическое задание) для оценки сформированности компетенции ПК-4:

Практические задания по дисциплине включают:

1. Стартовый практикум: Прикладной анализ данных на базе модели простой линейной регрессии (рекомендуется EXCEL).
2. Прикладной анализ данных на базе модели множественной регрессии (здесь и далее рекомендуются EXCEL, STATISTICA, SPSS).
3. Построение регрессионной модели в случае неоднородной дисперсии ошибок.
4. Подбор модели на основе процедур отбора переменных.
5. Построение регрессионной модели в случае корреляции ошибок.
6. Модельные задачи статистического вывода для номинальных, порядковых и числовых переменных. (рекомендуются EXCEL, STATISTICA, SPSS, по выбору студентов могут быть применены программные

средства на базе Python или R-проект)

7. Практические задания теоретического характера

Содержание и порядок выполнения заданий, а также файлы данных представлены на ресурсе:

Стронгина Н.Р. Методы моделирования социально-экономических процессов: анализ данных. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 793E.14.08. - Доступ требует

авторизации, <https://e-learning.unn.ru/enrol/index.php?id=793>

(см. Практикумы 1-5 и Практикумы 7-9 на ресурсе <https://e-learning.unn.ru/enrol/index.php?id=793>)

В ходе выполнения каждого практического задания

1. Опишите комплект данных модельной задачи.
2. Сформулируйте цель и задачу прикладного анализа данных.
3. Проведите первичный анализ данных.
4. Аргументируйте применение метода моделирования.
5. Опишите понятийный аппарат исследования, сформулируйте свойства метода.
6. Опираясь на математическую модель, объясните протокол обработки данных на компьютере.
7. Оцените качество построенной модели, сформулируйте результаты исследования.

Практические задания теоретического характера (примеры)

Сформулируйте определения и докажите утверждения:

1. Теорема о существовании и единственности МНК-приближения.
2. Теорема о свойствах оценочных значений отклика и остатков.
3. Теорема о статистических свойствах коэффициентов МНК-оценочного уравнения регрессии.
4. Теорема об оценке дисперсии ошибок (идея доказательства). Стандартная ошибка оценки.
5. Интервал доверия для $M(\hat{Y}_r)$, построенный на основе стандартной ошибки оценки.
6. Интервал прогнозирования для \hat{Y}_r , построенный на основе стандартной ошибки оценки.
7. Проверка значимости уравнения на основе доверительного интервала для коэффициента β_1 с использованием стандартной ошибки оценки.
8. Модель множественной линейной регрессии. Проверка значимости каждой переменной на основе доверительных интервалов с использованием стандартной ошибки оценки.
9. Модель множественной линейной регрессии. Точка наблюдения X_r . Замер Y_r . Прогнозное (оценочное) значение \hat{Y}_r . Интервал доверия и интервал прогнозирования.
10. Модель множественной линейной регрессии. Проверка значимости переменной. Статистика проверки. Решающее α -правило. Собственный уровень значимости (p). Решающее p -правило. Эквивалентность правил.
11. Модель множественной линейной регрессии. Проверка значимости в целом. Статистика проверки. Решающее α -правило. Решающее p -правило. Эквивалентность правил.
12. Постройте классификатор модельных задач статистического вывода, изученных в рамках дисциплины.

Критерии оценивания (оценочное средство - Практическое задание)

Оценка	Критерии оценивания
зачтено	Задание выполнено верно в полном объеме, либо выполнено в полном объеме с незначительными ошибками

Оценка	Критерии оценивания
не зачтено	Задание не выполнено, выполнено в существенно неполном объеме или допущены существенные ошибки

5.1.2 Типовые задания (оценочное средство - Проект) для оценки сформированности компетенции ПК-4:

Проект: Самостоятельный прикладной анализ данных

Тематика:

1. Прикладной анализ данных, построение факторной модели
2. Прикладной анализ данных, построение регрессионной модели
3. Прикладной анализ данных комбинированным методом

Индивидуальные и групповые задания проектов различаются комплектами (файлами) данных, тематикой прикладных задач и применением методов моделирования.

Комплекты (файлы) данных и методические указания представлены на ресурсах:

Стронгина Н.Р. Методы моделирования социально-экономических процессов: анализ данных. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 793Е.14.08. - Доступ требует авторизации, [; \(см. Практикум 4 на ресурсе <https://e-learning.unn.ru/enrol/index.php?id=793> \)](https://e-learning.unn.ru/enrol/index.php?id=793.)

Стронгина Н.Р. Методы моделирования социально-экономических процессов: анализ данных. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 814Е.14.08. - Доступ требует авторизации, [; \(см. Практикум 4 на ресурсе <https://e-learning.unn.ru/enrol/index.php?id=814> \)](https://e-learning.unn.ru/enrol/index.php?id=814.)

В ходе выполнения проекта:

1. Используйте данные в соответствии с индивидуальным заданием.
2. Определите цель исследования и массив исследуемых факторов.
3. Используйте методы первичного анализа, статистический вывод и методы снижения размерности.
4. Программные средства по выбору студента, рекомендуется статистический пакет.
5. В соответствии с методическими указаниями подготовьте отчет.

При подготовке отчета:

1. Опишите комплект данных проекта, укажите целесообразность их фильтрации и (или) группировки.
2. Сформулируйте цели и задачи прикладного анализа данных.
3. Сформулируйте результаты первичного анализа данных.
4. Аргументируйте выбор методов моделирования.
5. Опишите понятийный аппарат методов, сформулируйте основные свойства методов.
6. Опираясь на математическую модель и протокол обработки данных, опишите этапы исследования.
7. Структурируйте полученные результаты, проведите сравнительный анализ, сформулируйте выводы.
8. Оцените качество построенной модели.
9. Укажите инструменты исследования.

Критерии оценивания (оценочное средство - Проект)

Оценка	Критерии оценивания
зачтено	Проект выполнен верно и в полном объеме, либо выполнен в полном объеме с незначительными ошибками
не зачтено	Проект не выполнен, выполнен в неполном объеме или допущены существенные ошибки

5.1.3 Типовые задания (оценочное средство - Собеседование) для оценки сформированности компетенции ПК-4:

1. Принцип наименьших квадратов. Теорема о существовании и единственности МНК-приближения.
2. Запись прямой МНК-приближения, свойства коэффициентов и вывод формул для их вычисления.
3. Теорема о свойствах оценочных значений отклика и остатков.
4. Коэффициент детерминации R^2 . Диапазон изменения R^2 . Приведенный коэффициент детерминации.
5. Связь коэффициентов корреляции и детерминации. Диапазон и свойства коэффициента корреляции.
6. Уравнение простой линейной регрессии; множественной линейной регрессии; уравнения регрессии, линейные относительно параметров. Задачи регрессионного анализа.
7. Модель простой линейной регрессии. Ее предположения в канонической записи и их запись через математические ожидания. Интерпретация предположений в задачах обработки экспериментальных данных.
8. Модель простой линейной регрессии и МНК-оценочное уравнение регрессии, истинные и предсказанные значения коэффициентов, истинные и предсказанные значения отклика, ошибки и остатки: графический сравнительный анализ.
9. Теорема о дисперсии линейной комбинации независимых случайных величин. Теорема о линейных свойствах ковариации случайных величин. Теорема о дисперсии линейной комбинации двух признаков. Теорема о линейных свойствах ковариации признаков (формулировки).
10. Модель простой линейной регрессии. Теорема о статистических свойствах параметров МНК-оценочного уравнения регрессии. Влияние объема выборки, разброса точек наблюдения и дисперсии ошибок на дисперсию параметров МНК-уравнения.
11. Модель простой линейной регрессии. Теорема об оценке дисперсии ошибок (идея доказательства). Стандартная ошибка оценки.
12. Интервал $(1-\alpha) \times 100\%$ доверия для случайных величин, имеющих распределения: стандартное нормальное, нормальное, Стьюдента. Квантили распределений.
13. Модель простой линейной регрессии. Смысл величин Y_p , \hat{Y}_p , $Y_p - \hat{Y}_p$. Теоремы о статистических свойствах случайной величины \hat{Y}_p и случайной величины $Y_p - \hat{Y}_p$.
14. Модель простой линейной регрессии. Интервал доверия для $M(Y_p)$ на основе дисперсии ошибок. Интервал доверия для $M(Y_p)$ на основе стандартной ошибки оценки.
15. Модель простой линейной регрессии. Интервал прогнозирования для Y_p на основе стандартной ошибки оценки.
16. Модель простой линейной регрессии. Проверка значимости уравнения на основе доверительного интервала для коэффициента β_1 с использованием стандартной ошибки оценки.
17. Модель простой линейной регрессии. Проблема включения константы β_0 . Проверка на основе

доверительного интервала для β_0 на базе стандартной ошибки оценки.

18. Модель множественной линейной регрессии. Предположения модели. Запись МНК-оценочного уравнения регрессии. Проверка значимости каждой переменной на основе доверительных интервалов с использованием стандартной ошибки оценки (формулировки).
19. Модель множественной линейной регрессии. Точка наблюдения X_r . Замер Y_r . Прогнозное (оценочное) значение \hat{Y}_r . Интервал доверия и интервал прогнозирования (формулировки).
20. Основные термины теории статистических проверок (нулевая и альтернативная гипотезы, статистика проверки, ошибки первого и второго рода, уровень значимости проверки (α), мощность критерия (β)). Решающее правило проверки (α -правило). Формулировка результатов проверки.
21. Модель простой линейной регрессии. Проверка значимости уравнения. Статистика проверки. Решающее правило (α -правило). Собственный уровень значимости проверки (p). Решающее p -правило. Эквивалентность α -правила и p -правила.
22. Модель множественной линейной регрессии. Проверка значимости переменной. Статистика проверки. Решающее правило проверки (α -правило). Собственный уровень значимости проверки (p). Решающее p -правило. Эквивалентность α -правила и p -правила.
23. Модель множественной линейной регрессии. Проверка значимости в целом. Идея проверки. Статистика проверки. Решающее правило проверки (α -правило). Решающее p -правило. Эквивалентность правил.
24. Критерии качества регрессионной модели (остатки, значимость, надежность оценки параметров, объясняющая способность уравнения). Взаимосвязь критериев.
25. Краткая характеристика проблем избыточности, разнородной дисперсии ошибок и автокорреляции.
26. Проблема избыточных факторов (определение). Геометрическая интерпретация избыточности.
27. Влияние избыточности факторов на статистические свойства МНК-уравнения (примеры). Основной индикатор избыточности.
28. Задачи отбора переменных. Общие свойства методов отбора переменных, применяемых в пакетах статистических программ.
30. Факторная модель для стандартизованных признаков. Критерии качества решения.
31. Проблемы поиска факторных нагрузок. Теорема о «диагональных» условиях нормировки.
32. Принцип главных факторов, теоремы о ранге матричного произведения и спектральном разложении.
33. Обоснование ограничений на число общих факторов.
34. Оценка факторных нагрузок методом главных факторов: алгоритм и критерии останова.
35. Выбор числа общих факторов и начальная оценка общностей.
36. Методы оценки факторных значений: регрессионный способ. Свойства оценок для F , $V(F)$ и A .
37. Факторные значения, факторные нагрузки и вращение факторов.
38. Многомерная статистическая модель «Территория. Год. Показатели». Анализ структуры и динамики.
39. Модели поколения и населения. Показатели таблиц дожития.
40. Статистический анализ таблиц дожития.

Критерии оценивания (оценочное средство - Собеседование)

Оценка	Критерии оценивания
зачтено	Ответы на вопросы даны верно или имеют незначительные неточности
не зачтено	Ответов нет или допущены грубые ошибки

5.1.4 Типовые задания (оценочное средство - Тест) для оценки сформированности компетенции ПК-4:

Вопросы теста о математической модели метода (примеры)

Вопрос 1

Тип вопроса: сопоставление

Формулировка вопроса: **Метод наименьших квадратов (один фактор)**

Сопоставить символьной записи текстовую:

Ответ 1: X ... Верное значение – объясняющая переменная, фактор

Ответ 2: Y ... Верное значение – объясняемая переменная, отклик

Ответ 3: n ... Верное значение – число объектов (наблюдений)

Ответ 4: i ... Верное значение – индекс для объектов (наблюдений)

Ответ 5: X_i ... Верное значение – значение фактора на объекте

Ответ 6: Y_i ... Верное значение – значение отклика на объекте

Вопрос 7

Тип вопроса: классификация

Формулировка вопроса: **Простая линейная регрессия: свойства МНК-оценок коэффициентов**

Поведение дисперсии оценочных коэффициентов регрессии:

Ответ 1: при увеличении σ^2 (дисперсии ошибок) ...

Классифицирующий термин – дисперсии МНК-оценок коэффициентов регрессии возрастают

Ответ 2: при увеличении n (числа точек наблюдения) ...

Классифицирующий термин – дисперсии МНК-оценок коэффициентов регрессии убывают

Ответ 3: при увеличении «кучности» наблюдений (т.е. при снижении $\text{Var}(X)$) ...

Классифицирующий термин – дисперсии МНК-оценок коэффициентов регрессии возрастают

Вопрос 10

Тип вопроса: сопоставление

Формулировка вопроса: **Простая линейная регрессия: свойства МНК-оценок коэффициентов**

По набору значений (X_i, Y_i) , $i=1, \dots, n$...

Ответ 1: σ^2 (дисперсию ошибок) ... Верное значение – нельзя вычислить

Ответ 2: s^2 (оценку дисперсии ошибок) ... Верное значение – можно вычислить

Вопрос 12

Тип вопроса: одиночный выбор

Формулировка вопроса: **Простая линейная регрессия: проверка значимости фактора**

Распределение Стьюдента с числом степеней свободы $n-2$...

Варианты ответов:

* распределение статистики проверки

* распределение статистики проверки в случае, когда нулевая гипотеза верна (+)

Вопрос 15

Тип вопроса: классификация

Формулировка вопроса: **Простая линейная регрессия: проверка значимости фактора в пакете статистических программ**

Сопоставить символьной записи текстовую:

Ответ 1: Пусть $p=0.0002$. Тогда на уровне значимости $\alpha=0.05$...

Классифицирующий термин – нулевая гипотеза отклоняется, исследуемый фактор значим

Ответ 2: Пусть $p=0.0002$. Тогда на уровне значимости $\alpha=0.01$...

Классифицирующий термин – нулевая гипотеза отклоняется, исследуемый фактор значим
 Ответ 3: Пусть $p=0.14$. Тогда на уровне значимости $\alpha=0.05$...
 Классифицирующий термин – нулевая гипотеза принимается, исследуемый фактор не значим
 Ответ 4: Пусть $p=0.14$. Тогда на уровне значимости $\alpha=0.01$...
 Классифицирующий термин – нулевая гипотеза принимается, исследуемый фактор не значим
 Ответ 5: Пусть $p=0.03$. Тогда на уровне значимости $\alpha=0.05$..
 Классифицирующий термин – нулевая гипотеза отклоняется, исследуемый фактор значим
 Ответ 6: Пусть $p=0.03$. Тогда на уровне значимости $\alpha=0.01$...
 Классифицирующий термин – нулевая гипотеза принимается, исследуемый фактор не значим

Вопрос 19

Тип вопроса: классификация

Формулировка вопроса: **Простая линейная регрессия: прогнозирование**

Поведение границ интервала доверия

Ответ 1: при увеличении σ^2 (дисперсии ошибок) ...
 Классифицирующий термин – границы становятся шире
 Ответ 2: при увеличении n (числа точек наблюдения) ...
 Классифицирующий термин – границы становятся уже
 Ответ 3: при увеличении «кучности» наблюдений (т.е. при уменьшении $(\text{Var}(X))$) ...
 Классифицирующий термин – границы становятся шире
 Ответ 4: по мере приближения X_p к значению X' ...
 Классифицирующий термин – границы становятся уже
 Ответ 5: при увеличении уровня $1-\alpha$...
 Классифицирующий термин – границы становятся шире

Критерии оценивания (оценочное средство - Тест)

Оценка	Критерии оценивания
зачтено	Вопросы теста отвечены правильно, работа над ошибками проведена
не зачтено	Тест не отвечен, либо ответы частично правильны, работа над ошибками не проведена

5.2. Описание шкал оценивания результатов обучения по дисциплине при промежуточной аттестации

Шкала оценивания сформированности компетенций

Уровень сформированности компетенций (индикатор достижения компетенций)	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено		зачтено				
Знания	Отсутствие	Уровень	Минимальн	Уровень	Уровень	Уровень	Уровень

	знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	знаний ниже минимальных требований. Имели место грубые ошибки	о допустимый уровень знаний. Допущено много негрубых ошибок	знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок	знаний в объеме, соответствующем программе подготовки. Ошибок нет.	знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но некоторые с недочетами	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения. Решены все основные задачи с отдельным и несущественными недочетами, выполнены все задания в полном объеме	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие базовых навыков. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов	Продemonстрированы навыки при решении нестандартных задач без ошибок и недочетов	Продemonстрирован творческий подход к решению нестандартных задач

Шкала оценивания при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно», продемонстрированы знания, умения, владения по соответствующим компетенциям на уровне выше предусмотренного программой
	отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично».
	очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо»
	хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо».
	удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»

не зачтено	неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно».
	плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.3 Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения на промежуточной аттестации с указанием критериев их оценивания:

5.3.1 Типовые задания (оценочное средство - Контрольные вопросы) для оценки сформированности компетенции ПК-4

Контрольные вопросы по итогам выполнения практического задания теоретического характера

1. Приведите постановку задачи, сформулируйте метод моделирования, приведите определения и доказательство утверждений в соответствии с заданием.
2. Укажите место и роль доказанных утверждений в системе теоретического обоснования метода
3. Укажите место и роль доказанных утверждений в сфере применения метода.

Контрольные вопросы по итогам выполнения других практических заданий (практикум)

1. Опишите исходные данные практического задания (файл данных).
2. Сформулируйте цель и задачу исследования.
3. Проведите первичный анализ данных.
4. Аргументируйте применение предложенного в задании метода.
5. Опишите понятийный аппарат проведенного исследования, сформулируйте его основные свойства.
6. Комментируйте компьютерную диагностику построения решения.
7. Проверьте корректность применения метода (если нужно, используйте статистический вывод), и определите качество решения.
8. Сформулируйте результаты исследования.
9. Приведите характеристику использованных программных средств.

Контрольные вопросы по итогам выполнения проекта

1. Опишите исходные данные проекта, целесообразность их фильтрации и (или) группировки.
2. Опишите первоначальные и окончательные цели и задачи исследования (их может быть несколько).
3. Сформулируйте результаты первичного анализа данных.
4. Аргументируйте выбор методов для дальнейшего исследования. Опишите их понятийный аппарат.
5. Опираясь на компьютерную диагностику, опишите этапы применения методов.
6. Проверьте корректность применения методов (если нужно, используйте статистический вывод), определите качество предложенных решений.
7. Структурируйте полученные результаты, сформулируйте выводы.
8. Проведите сравнительный анализ использованных подходов и методов.
9. Приведите комплексную характеристику использованных программных средств.

Контрольные вопросы по итогам изучения теоретического материала

1. Принцип наименьших квадратов. Теорема о существовании и единственности МНК-приближения.
2. Запись прямой МНК-приближения, свойства коэффициентов и вывод формул для их вычисления.

3. Теорема о свойствах оценочных значений отклика и остатков.
4. Коэффициент детерминации R^2 . Диапазон изменения R^2 . Приведенный коэффициент детерминации.
5. Связь коэффициентов корреляции и детерминации. Диапазон и свойства коэффициента корреляции.
6. Уравнение простой линейной регрессии; множественной линейной регрессии; уравнения регрессии, линейные относительно параметров. Задачи регрессионного анализа.
7. Модель простой линейной регрессии. Ее предположения в канонической записи и их запись через математические ожидания. Интерпретация предположений в задачах обработки экспериментальных данных.
8. Модель простой линейной регрессии и МНК-оценочное уравнение регрессии, истинные и предсказанные значения коэффициентов, истинные и предсказанные значения отклика, ошибки и остатки: графический сравнительный анализ.
9. Теорема о дисперсии линейной комбинации независимых случайных величин. Теорема о линейных свойствах ковариации случайных величин. Теорема о дисперсии линейной комбинации двух признаков. Теорема о линейных свойствах ковариации признаков (формулировки).
10. Модель простой линейной регрессии. Теорема о статистических свойствах параметров МНК-оценочного уравнения регрессии. Влияние объема выборки, разброса точек наблюдения и дисперсии ошибок на дисперсию параметров МНК-уравнения.
11. Модель простой линейной регрессии. Теорема об оценке дисперсии ошибок (идея доказательства). Стандартная ошибка оценки.
12. Интервал $(1-\alpha) \times 100\%$ доверия для случайных величин, имеющих распределения: стандартное нормальное, нормальное, Стьюдента. Квантили распределений.
13. Модель простой линейной регрессии. Смысл величин \hat{Y}_p , \check{Y}_p , $Y_p - \check{Y}_p$. Теоремы о статистических свойствах случайной величины \check{Y}_p и случайной величины $Y_p - \check{Y}_p$.
14. Модель простой линейной регрессии. Интервал доверия для $M(\hat{Y}_p)$ на основе дисперсии ошибок. Интервал доверия для $M(\hat{Y}_p)$ на основе стандартной ошибки оценки.
15. Модель простой линейной регрессии. Интервал прогнозирования для \hat{Y}_p на основе стандартной ошибки оценки.
16. Модель простой линейной регрессии. Проверка значимости уравнения на основе доверительного интервала для коэффициента β_1 с использованием стандартной ошибки оценки.
17. Модель простой линейной регрессии. Проблема включения константы β_0 . Проверка на основе доверительного интервала для β_0 на базе стандартной ошибки оценки.
18. Модель множественной линейной регрессии. Предположения модели. Запись МНК-оценочного уравнения регрессии. Проверка значимости каждой переменной на основе доверительных интервалов с использованием стандартной ошибки оценки (формулировки).
19. Модель множественной линейной регрессии. Точка наблюдения X_p . Замер Y_p . Прогнозное (оценочное) значение \hat{Y}_p . Интервал доверия и интервал прогнозирования (формулировки).
20. Основные термины теории статистических проверок (нулевая и альтернативная гипотезы, статистика проверки, ошибки первого и второго рода, уровень значимости проверки (α), мощность критерия (β)). Решающее правило проверки (α -правило). Формулировка результатов проверки.
21. Модель простой линейной регрессии. Проверка значимости уравнения. Статистика проверки. Решающее правило (α -правило). Собственный уровень значимости проверки (p). Решающее p -правило. Эквивалентность α -правила и p -правила.
22. Модель множественной линейной регрессии. Проверка значимости переменной. Статистика проверки. Решающее правило проверки (α -правило). Собственный уровень значимости проверки (p). Решающее p -правило. Эквивалентность α -правила и p -правила.
23. Модель множественной линейной регрессии. Проверка значимости в целом. Идея проверки. Статистика проверки. Решающее правило проверки (α -правило). Решающее p -правило. Эквивалентность правил.
24. Критерии качества регрессионной модели (остатки, значимость, надежность оценки параметров,

объясняющая способность уравнения). Взаимосвязь критериев.

25. Краткая характеристика проблем избыточности, разнородной дисперсии ошибок и автокорреляции.

26. Проблема избыточных факторов (определение). Геометрическая интерпретация избыточности.

27. Влияние избыточности факторов на статистические свойства МНК-уравнения (примеры). Основной индикатор избыточности.

28. Задачи отбора переменных. Общие свойства методов отбора переменных, применяемых в пакетах статистических программ.

30. Факторная модель для стандартизованных признаков. Критерии качества решения.

31. Проблемы поиска факторных нагрузок. Теорема о «диагональных» условиях нормировки.

32. Принцип главных факторов, теоремы о ранге матричного произведения и спектральном разложении.

33. Обоснование ограничений на число общих факторов.

34. Оценка факторных нагрузок методом главных факторов: алгоритм и критерии остановки.

35. Выбор числа общих факторов и начальная оценка общностей.

36. Методы оценки факторных значений: регрессионный способ. Свойства оценок для F , $V(F)$ и A .

37. Факторные значения, факторные нагрузки и вращение факторов.

38. Многомерная статистическая модель «Территория. Год. Показатели». Анализ структуры и динамики.

39. Модели поколения и населения. Показатели таблиц дожития.

40. Статистический анализ таблиц дожития.

Критерии оценивания (оценочное средство - Контрольные вопросы)

Оценка	Критерии оценивания
зачтено	Контрольные вопросы отвечены верно и в полном объеме, могут быть допущены незначительные неточности
не зачтено	Ответов на контрольные вопросы нет, ответы даны не в полном объеме или допущены существенные ошибки

6. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Основная литература:

1. Гнеденко Борис Владимирович. Курс теории вероятностей : [учеб. для мех.-мат. специальностей ун-тов]. - 6-е изд., перераб. и доп. - М. : Наука, 1988. - 466, [1] с. : ил. - ISBN 5-02-013761-8 (в пер.) : 1.20., 483 экз.

2. Воскобойников Ю. Е. Регрессионный анализ данных в пакете MATHCAD / Воскобойников Ю. Е. - Санкт-Петербург : Лань, 2022. - 224 с. - Книга из коллекции Лань - Математика. - ISBN 978-5-8114-1096-5., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=799705&idb=0>.

3. Анализ данных : учебник / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. - Москва : Юрайт, 2023. - 490 с. - (Высшее образование). - ISBN 978-5-534-00616-2. - Текст : электронный // ЭБС "Юрайт"., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=847079&idb=0>.

Дополнительная литература:

1. Айвазян Сергей Арутюнович. Методы эконометрики : Учебник / Центральный экономико-математический институт Российской академии наук; Национальный исследовательский университет "Высшая школа экономики". - 1. - Москва : Издательство "Магистр", 2022. - 512 с. - ВО - Бакалавриат. - ISBN 978-5-9776-0153-5. - ISBN 978-5-16-110618-1. - ISBN 978-5-16-004050-9., <https://e-lib.unn.ru/MegaPro/UserEntry?Action=FindDocs&ids=834477&idb=0>.

Программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

Операционная система Windows (лицензия), SPSS или Statistica (лицензия), Microsoft Office (лицензия), текстовый редактор Word, табличный редактор EXCEL, расширение "Анализ данных" в пакете EXCEL, базовый комплект и комплект "Анализ данных" профессиональных статистических пакетов Statistica или SPSS, опционально программные средства анализа данных на базе Python, C++

Интернет-ресурсы:

1. Стронгина Н.Р. Методы моделирования социально-экономических процессов: анализ данных. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 814E.14.08. <https://e-learning.unn.ru/course/view.php?id=814> - Требуется регистрация.
2. Стронгина Н.Р. Методы моделирования социально-экономических процессов: регрессионный анализ. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Идентификационный номер 793E.14.08. <https://e-learning.unn.ru/course/view.php?id=793> - Требуется регистрация.
3. Российский статистический ежегодник. Росстат (ежегодное официальное издание). Сайт Федеральной службы государственной статистики. <https://rosstat.gov.ru/folder/210/document/12994>, http://www.gks.ru/bgd/regl/b16_13/Main.htm

7. Материально-техническое обеспечение дисциплины (модуля)

Учебные аудитории для проведения учебных занятий, предусмотренных образовательной программой, оснащены мультимедийным оборудованием (проектор, экран), техническими средствами обучения, компьютерами, специализированным оборудованием: Учебная аудитория для проведения лекций, оснащенная партами, стульями, обязательно учебной доской. Оборудование для показа презентаций (персональный компьютер, проектор, экран). Терминал-класс с лицензионным программным обеспечением. Учебная и научная литература, представленная в библиотечном фонде ННГУ, в электронных библиотеках и на кафедре ДУМЧА ИИТММ.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ОС ННГУ по направлению подготовки/специальности 01.03.02 - Прикладная математика и информатика.

Автор(ы): Стронгина Наталья Романовна, кандидат физико-математических наук, доцент.

Заведующий кафедрой: Калинин Алексей Вячеславович, доктор физико-математических наук.

Программа одобрена на заседании методической комиссии от 13.12.2023, протокол № 3.