

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ ТАДЖИКИСТАН
МЕЖГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКО-ТАДЖИКСКИЙ (СЛАВЯНСКИЙ) УНИВЕРСИТЕТ»

Естественнонаучный факультет

Кафедра математики и физики

«УТВЕРЖДАЮ»

« 28 » 08 2024 г.

Зав. кафедрой Гулбоев Б.Дж.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
по учебной дисциплине

Математические методы в исторических исследованиях

46.03.01 «История»

Душанбе 2024 г.

ПАСПОРТ ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине Математические методы в исторических исследованиях
В результате изучения данной дисциплины у обучающихся формируются следующие универсальные компетенции (элементы компетенций)

Таблица 1.

код	Формируемая компетенция	Содержание этапа формирования компетенции	Вид оценочного средства
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	ИУК-1.1. Демонстрирует знание особенностей системного и критического мышления и готовность к нему ИУК-1.2. Применяет логические формы и процедуры, способен к рефлексии по поводу собственной и чужой мыслительной деятельности ИУК-1.3. Анализирует источник информации с точки зрения временных и пространственных условий его возникновения ИУК-1.4. Анализирует ранее сложившиеся в науке оценки информации. ИУК-1.5. Сопоставляет разные источники информации с целью выявления их противоречий и поиска достоверных суждений ИУК-1.6. Аргументированно формирует собственное суждение и оценку информации, принимает обоснованное решение. ИУК-1.7. Определяет практические последствия предложенного решения задачи	Комплект тематик для дискуссии, полемики, устный опрос, тесты, контрольные работы, рефераты доклады, презентация

№ п/п	Контролируемые разделы, темы*	Формируемые компетенции*	Индикаторы достижения компетенции*	Оценочные средства*	
				Количество тестовых заданий/вопросов к экзамену/зачету /зачету (с оценкой)	Другие оценочные средства Вид
1.	Вариационные ряды и их характеристики	УК-1	ИУК-1.1. Демонстрирует знание особенностей системного и критического мышления и готовность к нему ИУК-1.2. Применяет логические формы и процедуры, способен к рефлексии по поводу собственной и чужой мыслительной деятельности ИУК-1.3. Анализирует источник информации с точки зрения временных и пространственных	5	Перечень вопросов для устного опроса, Задачи для самостоятельного решения

			<p>условий его возникновения ИУК-1.4. Анализирует ранее сложившиеся в науке оценки информации. ИУК-1.5. Сопоставляет разные источники информации с целью выявления их противоречий и поиска достоверных суждений ИУК-1.6. Аргументированно формирует собственное суждение и оценку информации, принимает обоснованное решение. ИУК-1.7. Определяет практические последствия предложенного решения задачи</p>		
2.	Основы математической теории выборочного метода	УК-1	<p>ИУК-1.1. ИУК-1.2. ИУК-1.3. ИУК-1.4. ИУК-1.5. ИУК-1.6. ИУК-1.7.</p>	2	Перечень вопросов для устного опроса, Задачи для самостоятельного решения
3.	Проверка статистических гипотез	УК-1	<p>ИУК-1.1. ИУК-1.2. ИУК-1.3. ИУК-1.4. ИУК-1.5. ИУК-1.6. ИУК-1.7.</p>	2	Перечень вопросов для устного опроса, Задачи для самостоятельного решения
4.	Дисперсионный анализ	УК-1	<p>ИУК-1.1. ИУК-1.2. ИУК-1.3. ИУК-1.4. ИУК-1.5. ИУК-1.6. ИУК-1.7.</p>	7	Перечень вопросов для устного опроса, Задачи для самостоятельного решения
5.	Корреляционный анализ	УК-1	<p>ИУК-1.1. ИУК-1.2. ИУК-1.3. ИУК-1.4. ИУК-1.5. ИУК-1.6. ИУК-1.7.</p>	7	Перечень вопросов для устного опроса, Задачи для самостоятельного решения
6.	Корреляционное отношение и индекс корреляции	УК-1	<p>ИУК-1.1. ИУК-1.2. ИУК-1.3. ИУК-1.4. ИУК-1.5. ИУК-1.6. ИУК-1.7.</p>	5	Перечень вопросов для устного опроса, Задачи для самостоятельного решения

7.	Регрессионный анализ	УК-1	ИУК-1.1. ИУК-1.2. ИУК-1.3. ИУК-1.4. ИУК-1.5. ИУК-1.6. ИУК-1.7.	7	Перечень вопросов для устного опроса, Задачи для самостоятельного решения
8.	Нелинейная регрессия	УК-1	ИУК-1.1. ИУК-1.2. ИУК-1.3. ИУК-1.4. ИУК-1.5. ИУК-1.6. ИУК-1.7.	5	Перечень вопросов для устного опроса, Задачи для самостоятельного решения
Всего:				61	

**МЕЖГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
 ВЫСШЕГО ОБРАЗОВАНИЯ
 «РОССИЙСКО-ТАДЖИКСКИЙ (СЛАВЯНСКИЙ) УНИВЕРСИТЕТ»**
Кафедра математики и физики
(наименование кафедры)

ПЕРЕЧЕНЬ ВОПРОСОВ ДЛЯ УСТНОГО ОПРОСА

1. Что такое вариационный ряд и как он строится?
2. Какова связь между вариационным рядом и частотным рядом?
3. Какие характеристики вариационного ряда вы можете выделить (например, медиана, мода, среднее)?
4. Как вычислить размах вариационного ряда и что он показывает?
5. В чем отличие между выборочным и генеральным вариационным рядом?
6. Что такое выборка и каковы основные типы выборок?
7. Каковы основные этапы процесса выборочного исследования?
8. Какова роль размера выборки в статистическом анализе?
9. Что такое стандартная ошибка выборки и как она рассчитывается?
10. Каковы основные принципы случайной выборки и их значение в статистике?
11. Что такое нулевая и альтернативная гипотеза?
12. Каковы основные этапы проверки статистической гипотезы?
13. Что такое уровень значимости и как он влияет на результаты проверки гипотез?
14. Как интерпретировать p -значение в контексте проверки гипотез?
15. Какие существуют ошибки первого и второго рода при проверке гипотез?
16. Что такое дисперсионный анализ и для чего он используется?
17. Каковы основные типы дисперсионного анализа (например, однофакторный, многофакторный)?
18. Какова роль F -статистики в дисперсионном анализе?
19. Как проводится пост-хок тестирование после дисперсионного анализа?

20. Какие предпосылки необходимо проверить перед проведением дисперсионного анализа?
21. Что такое корреляция и как она измеряется?
22. Какова разница между положительной и отрицательной корреляцией?
23. Что такое коэффициент корреляции Пирсона и как его интерпретировать?
24. Каковы ограничения использования корреляционного анализа?
25. Как можно визуализировать корреляцию между двумя переменными?
26. Что такое линейная регрессия и какова ее основная цель?
27. Каковы основные предпосылки линейной регрессии?
28. Как интерпретировать коэффициенты регрессии в модели?
29. Что такое множественная регрессия и как она отличается от простой регрессии?
30. Каковы методы оценки качества модели регрессии (например, R-квадрат)?

МЕЖГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКО-ТАДЖИКСКИЙ (СЛАВЯНСКИЙ) УНИВЕРСИТЕТ»

Естественнонаучный факультет
Кафедра математики и физики

Математические методы в исторических исследованиях
наименование дисциплины (модуля)

для 46.03.01 «История»
шифр/направление

Всеобщая и отечественная история
наименование профиля / специализации / программы

очная
форма обучения

ТЕСТОВЫЕ ЗАДАНИЯ ЗАКРЫТОГО ТИПА

1. **Какое из следующих определений является правильным для вариационного ряда?**
 - А) Упорядоченная последовательность наблюдений.
 - В) Сумма всех наблюдений.
 - С) Разница между максимальным и минимальным значениями.
 - D) Среднее арифметическое наблюдений.
 - **Правильный ответ: А**
2. **Как называется наименьшее значение в вариационном ряду?**
 - А) Мода
 - В) Медиана
 - С) Минимум
 - D) Максимум
 - **Правильный ответ: С**
3. **Что из следующего представляет собой размах вариационного ряда?**
 - А) Среднее значение
 - В) Разница между максимальным и минимальным значениями
 - С) Среднеквадратическое отклонение
 - D) Медиана
 - **Правильный ответ: В**
4. **Какой из следующих показателей характеризует степень разброса значений в вариационном ряду?**
 - А) Мода
 - В) Среднее арифметическое
 - С) Дисперсия

- D) Медиана
 - **Правильный ответ: С**
5. Если в вариационном ряде все значения равны, то его дисперсия:
- A) Больше нуля
 - B) Равна нулю
 - C) Бесконечна
 - D) Неопределена
 - **Правильный ответ: В**
6. Как называется значение, которое встречается в вариационном ряде наиболее часто?
- A) Мода
 - B) Медиана
 - C) Среднее арифметическое
 - D) Размах
 - **Правильный ответ: А**
7. Какой из следующих методов используется для нахождения медианы в неупорядоченном вариационном ряде?
- A) Сортировка значений
 - B) Вычисление среднего арифметического
 - C) Нахождение размаха
 - D) Нахождение моды
 - **Правильный ответ: А**
8. Какой из следующих показателей не является характеристикой вариационного ряда?
- A) Мода
 - B) Дисперсия
 - C) Стандартное отклонение
 - D) Коэффициент корреляции
 - **Правильный ответ: D**
9. При каком условии мода может быть единственной в вариационном ряде?
- A) Если ряд содержит только два значения
 - B) Если все значения различны
 - C) Если одно значение встречается чаще остальных
 - D) Если ряд содержит только одно значение
 - **Правильный ответ: С**
10. Какой из следующих методов позволяет визуально представить распределение значений в вариационном ряде?
- A) Гистограмма
 - B) Среднее арифметическое
 - C) Мода
 - D) Дисперсия
 - **Правильный ответ: А**
11. Что такое дисперсионный анализ?

- A) Метод для оценки корреляции между переменными
- B) Метод для проверки различий между средними значениями двух или более групп
- C) Метод для регрессии
- D) Метод для построения графиков
- **Правильный ответ: B**

12. Какой из следующих критериев используется в дисперсионном анализе для проверки гипотезы о равенстве средних?

- A) Критерий Стьюдента
- B) Критерий Фишера
- C) Критерий χ^2
- D) Критерий Манна-Уитни
- **Правильный ответ: B**

13. Какой тип дисперсионного анализа применяется, если у нас есть одна зависимая переменная и одна независимая категориальная переменная?

- A) Однофакторный дисперсионный анализ
- B) Многофакторный дисперсионный анализ
- C) Регрессионный анализ
- D) Корреляционный анализ
- **Правильный ответ: A**

14. Что измеряет дисперсия в контексте дисперсионного анализа?

- A) Среднее значение
- B) Разброс значений относительно среднего
- C) Корреляцию между переменными
- D) Плотность вероятности
- **Правильный ответ: B**

15. Какой из следующих методов используется для анализа дисперсии, когда у нас есть более двух групп?

- A) t-тест
- B) ANOVA
- C) Регрессионный анализ
- D) Корреляционный анализ
- **Правильный ответ: B**

16. Какой из следующих факторов влияет на мощность теста в дисперсионном анализе?

- A) Размер выборки
- B) Количество групп
- C) Уровень значимости
- D) Все вышеперечисленное
- **Правильный ответ: D**

17. Что означает нулевая гипотеза в контексте дисперсионного анализа?

- A) Средние значения всех групп равны

- В) Средние значения всех групп различны
- С) Данные нормально распределены
- D) Данные имеют одинаковую дисперсию
- **Правильный ответ: А**

18. Какой из следующих критериев используется для проверки предположения о равенстве дисперсий в дисперсионном анализе?

- А) Критерий Стьюдента
- В) Критерий Левена
- С) Критерий Фишера
- D) Критерий χ^2
- **Правильный ответ: В**

19. Какой из следующих методов может быть использован для пост-хок анализа после дисперсионного анализа?

- А) t-тест
- В) Метод Бонферрони
- С) Метод Шеффе
- D) Все вышеперечисленное
- **Правильный ответ: D**

20. Какой уровень значимости обычно используется в дисперсионном анализе для проверки гипотез?

- А) 0.01
- В) 0.05
- С) 0.10
- D) 0.20
- **Правильный ответ: В**

21. Какой коэффициент корреляции используется для измерения линейной зависимости между двумя количественными переменными?

- А) Коэффициент Пирсона
- В) Коэффициент Спирмена
- С) Коэффициент Кендалла
- D) Коэффициент вариации
- **Правильный ответ: А) Коэффициент Пирсона**

22. Какое значение коэффициента корреляции указывает на полную отрицательную линейную зависимость?

- А) 0
- В) 0.5
- С) -1
- D) 1

Правильный ответ: С) -1

23. Если коэффициент корреляции между переменными X и Y равен 0.85, то это означает, что:

- А) Переменные независимы
- В) Существует сильная положительная линейная зависимость

- C) Существует слабая отрицательная зависимость
 - D) Переменные не имеют корреляции
- Правильный ответ:** B) Существует сильная положительная линейная зависимость

24. Какой из следующих методов используется для оценки корреляции между ранговыми данными?

- A) Коэффициент корреляции Пирсона
- B) Коэффициент корреляции Спирмена
- C) Коэффициент корреляции Кендалла
- D) Обе B и C

Правильный ответ: D) Обе B и C

25. Что означает коэффициент корреляции, равный 0?

- A) Полная положительная зависимость
- B) Полная отрицательная зависимость
- C) Отсутствие линейной зависимости
- D) Слабая зависимость

Правильный ответ: C) Отсутствие линейной зависимости

26. Какой из следующих графиков лучше всего иллюстрирует положительную корреляцию между переменными?

- A) Точка, расположенная в центре
- B) Линия, наклоненная вниз
- C) Линия, наклоненная вверх
- D) Случайные точки без видимого порядка

Правильный ответ: C) Линия, наклоненная вверх

27. Какой из следующих факторов может повлиять на коэффициент корреляции?

- A) Размер выборки
- B) Наличие выбросов
- C) Нелинейные зависимости
- D) Все вышеперечисленное

Правильный ответ: D) Все вышеперечисленное

28. Если переменные X и Y имеют коэффициент корреляции 0.3, это означает, что:

- A) Существует сильная зависимость между переменными
- B) Существует слабая положительная зависимость
- C) Переменные независимы
- D) Существует сильная отрицательная зависимость

Правильный ответ: B) Существует слабая положительная зависимость

29. Какой из следующих методов может использоваться для визуализации корреляции между двумя переменными?

- A) Гистограмма
- B) Коробчатая диаграмма
- C) Диаграмма рассеяния

- D) Линейный график

Правильный ответ: C) Диаграмма рассеяния

30. Какой из следующих коэффициентов корреляции учитывает порядок данных?

- A) Коэффициент Пирсона
- B) Коэффициент Спирмена
- C) Коэффициент Кендалла
- D) Оба B и C

Правильный ответ: D) Оба B и C

31. Какой из следующих коэффициентов используется для оценки силы связи между переменными в линейной регрессии?

- A) Коэффициент корреляции
- B) Коэффициент детерминации
- C) Стандартная ошибка
- D) t-статистика

Правильный ответ: A) Коэффициент корреляции

32. Что обозначает коэффициент детерминации (R^2) в линейной регрессии?

- A) Долю вариации зависимой переменной, объясненную независимыми переменными
- B) Степень значимости независимой переменной
- C) Разницу между наблюдаемыми и предсказанными значениями
- D) Среднеквадратичную ошибку прогноза

Правильный ответ: A) Долю вариации зависимой переменной, объясненную независимыми переменными

33. Какой метод обычно используется для оценки параметров линейной регрессии?

- A) Метод максимального правдоподобия
- B) Метод наименьших квадратов
- C) Метод градиентного спуска
- D) Метод минимизации

Правильный ответ: B) Метод наименьших квадратов

34. Какой из следующих факторов может указывать на наличие мультиколлинеарности в модели регрессии?

- A) Высокий R^2
- B) Низкие p-значения для всех независимых переменных
- C) Высокие корреляции между независимыми переменными
- D) Наличие выбросов

Правильный ответ: C) Высокие корреляции между независимыми переменными

35. Что такое "остатки" в контексте линейной регрессии?

- A) Разница между наблюдаемыми и предсказанными значениями
- B) Среднее значение зависимой переменной
- C) Значения независимых переменных

- D) Оценки коэффициентов регрессии

Правильный ответ: А) Разница между наблюдаемыми и предсказанными значениями

36. Какой тест используется для проверки значимости коэффициента регрессии?

- А) Тест Фишера
- В) t-тест
- С) Z-тест
- D) Хи-квадрат тест

Правильный ответ: В) t-тест

37. Что означает "нормальность остатков" в модели линейной регрессии?

- А) Остатки должны следовать нормальному распределению
- В) Остатки должны быть равномерно распределены
- С) Остатки должны быть положительными
- D) Остатки должны быть независимыми

Правильный ответ: А) Остатки должны следовать нормальному распределению

38. Какой из следующих методов используется для проверки предположений о гомоскедастичности в регрессионном анализе?

- А) Тест Бреуша-Пагана
- В) t-тест
- С) ANOVA
- D) Тест Шапиро-Уилка

Правильный ответ: А) Тест Бреуша-Пагана

39. Какой вид регрессии используется, если зависимая переменная является категориальной?

- А) Линейная регрессия
- В) Полиномиальная регрессия
- С) Логистическая регрессия
- D) Множественная регрессия

Правильный ответ: С) Логистическая регрессия

40. Какой из следующих факторов не является предположением линейной регрессии?

- А) Линейность
- В) Нормальность остатков
- С) Независимость наблюдений
- D) Однородность выборки

Правильный ответ: D) Однородность выборки

ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Определите, что такое вариационный ряд и какова его основная цель в статистике. Приведите пример вариационного ряда на основе произвольного набора данных.

Ответ: Вариационный ряд — это упорядоченный набор наблюдаемых значений выборки, расположенных в порядке возрастания (или убывания). Основная цель вариационного ряда — проанализировать распределение данных, выявить их основные характеристики и тенденции. Пример: пусть у нас есть набор данных о росте 10 человек: 160, 165, 170, 175, 180, 155, 168, 172, 178, 162. Вариационный ряд будет: 155, 160, 162, 165, 168, 170, 172, 175, 178, 180.

2. Каковы основные характеристики вариационного ряда? Опишите, как они помогают в анализе данных.

Ответ: Основные характеристики вариационного ряда включают:

1. **Минимум** — наименьшее значение в ряду.
2. **Максимум** — наибольшее значение в ряду.
3. **Размах** — разница между максимальным и минимальным значениями.
4. **Мода** — значение, которое встречается наиболее часто.
5. **Медиана** — значение, которое делит ряд на две равные части.
6. **Среднее арифметическое** — сумма всех значений, деленная на их количество.

Эти характеристики помогают понять распределение данных, выявить отклонения и сделать выводы о тенденциях в выборке.

3. Объясните, что такое кумулятивный вариационный ряд и как он строится. Какую информацию он предоставляет?

Ответ: Кумулятивный вариационный ряд — это ряд, в котором каждое значение представляет собой сумму частот (или относительных частот) всех предыдущих значений вариационного ряда. Он строится путем накопления частот: для каждого значения вычисляется сумма частот всех значений, которые меньше или равны этому значению. Кумулятивный вариационный ряд позволяет анализировать, какая доля наблюдений находится ниже (или выше) определенного значения, и помогает визуализировать распределение данных, что особенно полезно для построения графиков и определения квантилей.

4. Объясните, что такое дисперсионный анализ (ANOVA) и в каких случаях его следует применять. Приведите пример.

Правильный ответ: Дисперсионный анализ (ANOVA) — это статистический метод, используемый для сравнения средних значений трех и

более групп для определения, существуют ли статистически значимые различия между ними. Он позволяет оценить влияние одного или нескольких факторов на зависимую переменную. ANOVA применяется, когда данные распределены нормально, и наблюдения в группах независимы. Например, ANOVA может быть использован для сравнения средних результатов экзаменов студентов, обучавшихся по разным методам (лекции, семинары, онлайн-курсы).

5. Какова основная гипотеза в дисперсионном анализе и как она формулируется? Что происходит, если гипотеза отвергается?

Правильный ответ: Основная гипотеза в дисперсионном анализе (H_0) утверждает, что все группы имеют одинаковые средние значения. Формулируется как "средние значения всех групп равны". Если гипотеза H_0 отвергается, это означает, что по крайней мере одна из групп имеет среднее значение, отличное от других, что указывает на наличие статистически значимых различий между группами.

6. Опишите процесс выполнения однофакторного дисперсионного анализа, включая основные этапы и необходимые предпосылки.

Правильный ответ: Процесс выполнения однофакторного дисперсионного анализа включает следующие этапы:

1. **Формулировка гипотез:** Определение нулевой гипотезы (H_0) и альтернативной гипотезы (H_1).
 2. **Сбор данных:** Сбор данных из разных групп, которые будут сравниваться.
 3. **Проверка предпосылок:** Убедитесь, что данные нормально распределены и имеют одинаковую дисперсию (гомоскедастичность).
 4. **Вычисление ANOVA:** Рассчитать F-статистику, используя формулы для межгрупповой и внутригрупповой дисперсии.
 5. **Сравнение с критическим значением:** Сравнить полученное значение F с критическим значением из таблицы F-распределения для определения значимости.
 6. **Интерпретация результатов:** Если F-статистика превышает критическое значение, отвергнуть H_0 и сделать вывод о наличии различий между группами.
- 7.** Объясните, что такое корреляция и как она измеряется. Какие значения коэффициента корреляции r указывают на сильную, умеренную и слабую корреляцию между двумя переменными?

Ответ: Корреляция — это статистическая мера, которая показывает степень и направление линейной зависимости между двумя переменными. Она измеряется с помощью коэффициента корреляции Пирсона r , который принимает значения от -1 до 1.

- Значения r близкие к 1 (например, $r > 0.7$) указывают на сильную положительную корреляцию.
- Значения r близкие к -1 (например, $r < -0.7$) указывают на сильную отрицательную корреляцию.

- Значения r около 0 (например, $-0.3 < r < 0.3$ – $0.3 < r < 0.3$) указывают на отсутствие или слабую корреляцию.
- **Вопрос:** В чем разница между корреляцией и причинно-следственной связью? Приведите примеры, чтобы проиллюстрировать ваше объяснение.
- **Ответ:** Корреляция указывает на наличие статистической зависимости между двумя переменными, но не подразумевает причинно-следственной связи. Например, существует корреляция между количеством мороженого, продаваемого летом, и количеством утоплений. Однако это не означает, что продажа мороженого вызывает утопления; скорее, обе переменные могут зависеть от третьей переменной — температуры. Причинно-следственная связь подразумевает, что изменение одной переменной вызывает изменение другой, что требует дополнительного анализа и экспериментов для подтверждения.

8. Каковы основные предпосылки для применения корреляционного анализа? Объясните, почему каждая из этих предпосылок важна.

Ответ: Основные предпосылки для применения корреляционного анализа включают:

1. **Линейность:** Предполагается, что связь между переменными является линейной. Это важно, потому что коэффициент корреляции Пирсона измеряет только линейные зависимости.
2. **Нормальность:** Данные должны быть нормально распределены, особенно для малых выборок. Это важно для корректного применения статистических тестов.
3. **Гомоскедастичность:** Дисперсия остатков должна быть постоянной при всех уровнях независимой переменной. Это важно для надежности результатов.
4. **Независимость наблюдений:** Наблюдения должны быть независимыми друг от друга, чтобы избежать искажения результатов.

9. Объясните, что такое линейная регрессия, и приведите пример, когда она может быть применена. Какие предпосылки необходимо проверить перед использованием линейной регрессии?

Ответ: Линейная регрессия — это статистический метод, используемый для моделирования зависимости между одной зависимой переменной и одной или несколькими независимыми переменными, предполагая линейную связь. Например, линейная регрессия может быть применена для предсказания цены дома на основе его площади и количества комнат. Перед использованием линейной регрессии необходимо проверить следующие предпосылки: линейность, независимость ошибок, нормальность распределения ошибок и гомоскедастичность (однородность дисперсии ошибок).

10. Проведите анализ остатков в модели линейной регрессии. Каковы основные шаги и почему это важно?

Ответ: Анализ остатков включает следующие основные шаги:

1. Вычисление остатков (разница между наблюдаемыми и предсказанными значениями).
2. Построение графика остатков против предсказанных значений для визуальной проверки на наличие паттернов.
3. Проверка нормальности распределения остатков с помощью тестов (например, тест Шапиро-Уилка).
4. Проверка гомоскедастичности с помощью графиков или тестов (например, тест Бреуша-Пагана).

Анализ остатков важен, так как помогает выявить проблемы в модели, такие как наличие нелинейных зависимостей или гетероскедастичности, что может привести к неверным выводам.

11. Опишите процесс выбора переменных в модели регрессии. Какие методы можно использовать для выбора наиболее значимых переменных?

Ответ: Процесс выбора переменных включает следующие этапы:

1. Определение всех возможных независимых переменных, которые могут влиять на зависимую переменную.
2. Использование методов, таких как:
 - Метод шаговой регрессии (вперед, назад, и обеим направлениям).
 - Критерий информации Акаике (AIC) или байесовский критерий информации (BIC) для оценки моделей.
 - Тестирование значимости коэффициентов регрессии (например, t-тест).
3. Оценка модели по критериям, таким как R^2 , скорректированный R^2 , и проверка предпосылок модели.

Критерии оценки:

- оценка «**зачтено**» выставляется студенту, если

Полное верное решение. В логическом рассуждении и решении нет ошибок, задача решена рациональным способом. Получен правильный ответ. Ясно описан способ решения.

- оценка «**не зачтено**»

Решение неверное или отсутствует

Разработчик: к.ф.-м.н., доцент Гулбоев Б.Дж.