

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ЛУГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ ВЛАДИМИРА ДАЛЯ»

Экономический институт
Кафедра экономической кибернетики и прикладной статистики
(наименование кафедры)

УТВЕРЖДАЮ

Директор

« 28 »



Тхор Е.С.

20

года

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
по учебной дисциплине (практике)**

Современные технологии обработки статистических данных

(наименование учебной дисциплины, практики)

38.03.05 Бизнес-информатика

(код и наименование направления подготовки (специальности))

Экономическая аналитика и бизнес-статистика

(наименование профиля подготовки (специальности, магистерской программы); при отсутствии ставится прочерк)

Разработчик (разработчики):

доцент

(подпись)

(подпись)

Спорняк С.А.

ФОС рассмотрен и одобрен на заседании кафедры экономической кибернетики и прикладной статистики от « 25 » 02 20 25 г., протокол № 25

Заведующий кафедрой экономической кибернетики и прикладной статистики

(подпись)

Велигура А.В.

Луганск 2025 г.

**Комплект оценочных материалов по дисциплине (практике)
«Современные технологии обработки статистических данных»**

Задания закрытого типа

Задания закрытого типа на выбор правильного ответа

1. Выберите один правильный ответ:

Какой метод используется для оценки значимости различий между двумя выборками в статистическом анализе?

- А) Корреляционный анализ
- Б) Дисперсионный анализ
- В) t-критерий Стьюдента
- Г) Регрессионный анализ

Правильный ответ: В.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Выберите один правильный ответ:

Какая из перечисленных функций используется в Excel для вычисления среднего значения набора данных?

- А) СУММ
- Б) СРЗНАЧ
- В) МЕДИАНА
- Г) СЧЁТ

Правильный ответ: Б.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. Выберите один правильный ответ:

Какой из следующих критериев НЕ относится к непараметрическим методам статистического анализа?

- А) Критерий знаков
- Б) Критерий Вилкоксона
- В) Критерий Колмогорова-Смирнова
- Г) t-критерий Стьюдента

Правильный ответ: Г.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. Выберите один правильный ответ:

Какой тип диаграммы лучше всего подходит для отображения долевого соотношения категорий в данных?

- А) Гистограмма
- Б) Круговая диаграмма
- В) Линейный график

Г) Коробчатая диаграмма

Правильный ответ: Б.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. Выберите один правильный ответ:

Какой из следующих показателей используется для измерения разброса данных вокруг среднего значения?

А) Медиана

Б) Мода

В) Стандартное отклонение

Г) Среднее арифметическое

Правильный ответ: В.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. Выберите один правильный ответ:

Как называется процесс проверки гипотезы о том, что средние значения двух групп не различаются?

А) Корреляционный анализ

Б) Дисперсионный анализ

В) Тестирование статистических гипотез

Г) Кластерный анализ

Правильный ответ: В.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания закрытого типа на установление соответствия

1. Установите соответствие между статистическими методами и их назначением:

- | | |
|---|---------------------------------------|
| 1) Оценка влияния независимых переменных на зависимую | А) Корреляционно-регрессионный анализ |
| 2) Выявление групп объектов на основе их сходства | Б) Дисперсионный анализ |
| 3) Оценка значимости различий между группами | В) Факторный анализ |
| 4) Выделение скрытых факторов, влияющих на переменные | Г) Кластерный анализ |

Правильный ответ: 1 – А, 2 – Г, 3 – Б, 4 – В.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Установите соответствие между программными средствами и их возможностями:

- | | |
|--|--------------------|
| 1) Визуализация данных, проведение регрессионного и дисперсионного анализа | А) OpenOffice CALC |
| 2) Работа с временными рядами и авторегрессионными моделями | Б) R |

3) Проведение многомерного анализа данных и В) EViews
статистических тестов

4) Табличные вычисления, анализ выборок и Г) SPSS
статистических гипотез

Правильный ответ: 1 – Г, 2 – В, 3 – Б, 4 – А.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. Установите соответствие между показателями описательной статистики и их назначением:

- | | |
|---|------------------------------|
| 1) Измерение центральной тенденции | А) Среднее арифметическое |
| 2) Измерение разброса значений | Б) Дисперсия |
| 3) Измерение степени отклонения
распределения от симметрии | В) Коэффициент
асимметрии |
| 4) Значение, которое делит выборку
пополам | Г) Медиана |

Правильный ответ: 1 – А, 2 – Б, 3 – В, 4 – Г.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. Установите соответствие между видами анализа и их применением:

- | | |
|--|---------------------------------------|
| 1) Определение статистической значимости
различий между группами | А) Описательная
статистика |
| 2) Разделение объектов на группы по сходству
признаков | Б) Проверка
статистических гипотез |
| 3) Анализ распределения данных, вычисление
средних и стандартных отклонений | В) Кластерный анализ |
| 4) Оценка влияния одной переменной на
другую | Г) Регрессионный анализ |

Правильный ответ: 1 – Б, 2 – В, 3 – А, 4 – Г.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. Установите соответствие между методами анализа данных и их характеристиками:

- | | |
|--|------------------------------|
| 1) Используется для предсказания значений
зависимой переменной | А) Регрессионный анализ |
| 2) Выделяет скрытые структуры в данных | Б) Дискриминантный
анализ |
| 3) Классифицирует объекты на основе
независимых переменных | В) Факторный анализ |
| 4) Анализирует частотное распределение двух
категориальных переменных | Г) Кросс-табуляция |

Правильный ответ: 1 – А, 2 – Б, 3 – Б, 4 – Г.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. Установите соответствие между видами статистического тестирования и их назначением:

- | | |
|---|---------------------------------|
| 1) Сравнение средних значений двух групп | А) Критерий Стьюдента (t-тест) |
| 2) Анализ связи между категориальными переменными | Б) Критерий χ^2 -квадрат |
| 3) Сравнение средних значений трех и более групп | В) ANOVA (дисперсионный анализ) |
| 4) Проверка значимости коэффициентов регрессии | Г) Критерий Фишера |

Правильный ответ: 1 – А, 2 – Б, 3 – В, 4 – Г.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания закрытого типа на установление правильной последовательности

1. Установите правильную последовательность этапов анализа данных в OpenOffice CALC:

- А) Интерпретация полученных результатов
- Б) Формирование выборки
- В) Построение гистограммы
- Г) Анализ вариационного ряда

Правильный ответ: Б, В, Г, А.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Установите последовательность проведения регрессионного анализа в OpenOffice CALC:

- А) Построение корреляционной матрицы
- Б) Подбор уравнения регрессии
- В) Анализ исходной информации
- Г) Проверка значимости уравнения регрессии

Правильный ответ: В, А, Б, Г.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. Установите правильную последовательность шагов при проведении дисперсионного анализа в OpenOffice CALC:

- А) Интерпретация результатов
- Б) Проведение проверки гипотез
- В) Построение блочной диаграммы
- Г) Анализ исходной информации

Правильный ответ: Г, В, Б, А.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. Установите последовательность выполнения кластерного анализа в R:

- А) Выбор метода объединения
- Б) Интерпретация результатов
- В) Формирование кластеров
- Г) Определение меры расстояния

Правильный ответ: Г, А, В, Б.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. *Расположите в правильном порядке этапы построения модели ARIMA в Eviews:*

- А) Построение авторегрессионной модели
- Б) Оценка качества модели
- В) Определение временного лага
- Г) Выбор параметров сезонных колебаний

Правильный ответ: В, Г, А, Б

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. *Установите правильную последовательность проведения факторного анализа в R:*

- А) Оценка результатов анализа
- Б) Анализ структуры данных
- В) Выделение главных компонент
- Г) Объединение и обобщение данных

Правильный ответ: Г, В, Б, А

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания открытого типа

Задания открытого типа на дополнение

1. *Напишите пропущенное слово или словосочетание.*

_____ — это надстройка, используемая в OpenOffice CALC для проведения статистического анализа.

Правильный ответ: Пакет анализа.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. *Напишите пропущенное слово или словосочетание.*

Основная цель корреляционно-регрессионного анализа — выявление и измерение _____ между переменными.

Правильный ответ: зависимости.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. *Напишите пропущенное слово или словосочетание.*

Метод главных компонент используется для _____ данных с целью уменьшения их размерности.

Правильный ответ: снижения размерности.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. *Напишите пропущенное слово или словосочетание.*

В анализе данных с использованием R для проверки гипотезы о нормальности распределения используется тест _____.

Правильный ответ: Шапиро-Уилкса.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. *Напишите пропущенное слово или словосочетание.*

Модель ARIMA в EViews применяется для _____ временных рядов.

Правильный ответ: прогнозирования.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. *Напишите пропущенное слово или словосочетание.*

В кластерном анализе метод К-средних используется для _____ объектов в группы на основе их сходства.

Правильный ответ: разделения.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания открытого типа с кратким свободным ответом

1. *Дайте ответ на вопрос:*

Статистическая программа, используемая для авторегрессионного моделирования и анализа временных рядов, называется _____.

Правильный ответ: EViews.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. *Дайте ответ на вопрос:*

Метод анализа данных, который используется для уменьшения размерности исходных переменных и выделения главных компонент, называется _____.

Правильный ответ: Метод главных компонент (PCA).

Компетенции (индикаторы): ПК-2 (ПК-2.2).

3. *Дайте ответ на вопрос:*

В системе R инструмент для анализа и обработки больших данных, связанный с концепцией Data Mining, называется _____.

Правильный ответ: caret.

Компетенции (индикаторы): ПК-2 (ПК-2.2).

4. *Дайте ответ на вопрос:*

Основной метод статистического анализа, используемый для выявления взаимосвязей между переменными, называется _____.

Правильный ответ: Корреляционно-регрессионный анализ.

Компетенции (индикаторы): ПК-2 (ПК-2.2).

5. Дайте ответ на вопрос:

Процесс, в котором данные разделяются на группы по схожим характеристикам, называется _____.

Правильный ответ: Кластерный анализ.

Компетенции (индикаторы): ПК-2 (ПК-2.2).

Задания открытого типа с развернутым ответом

1. Почитайте текст задания. Продумайте логику и полноту ответа.

Запишите развернутый и обоснованный ответ.

В результате регрессионного анализа получены следующие данные:

	A	B	C	D	E	F	G
1							
2							
3	ВЫВОД ИТОГОВ						
4							
5	Регрессионная статистика						
6	Множественный R	0,993100061					
7	R-квадрат	0,986247732					
8	Нормированный R-квадрат	0,984719702					
9	Стандартная ошибка	929,2289258					
10	Наблюдения	11					
11							
12	Дисперсионный анализ						
13		df	SS	MS	F	Значимость F	
14	Регрессия	1	557313597,2	557313597,2	14,54375	0,009124	
15	Остаток	9	7771197,569	863466,3966			
16	Итого	10	565084794,7				
17							
18		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
19	Y-пересечение	10794,53017	1000,300908	10,79128299	1,89E-06	8531,692	13057,37
20	Переменная X 1	0,52472663	0,020654087	25,40546212	0,009124	0,478004	0,571449
21							

Запишите уравнение регрессии и сделайте выводы о качестве полученного уравнения регрессии.

Время выполнения – 20 мин.

Ожидаемый результат:

В нашем примере значение $R^2 = 0,9862477$ свидетельствует о том, что изменения зависимой переменной Y на 98,6% можно объяснить изменениями включенных в модель объясняющих переменных – X1 (среднемесячная номинальная начисленная заработная плата). И на 1,4% среднедушевые доходы зависят от других неучтенных факторов. Такое значение свидетельствует об адекватности модели.

Нормированный R-квадрат – поправленный (скорректированный по числу степеней свободы) коэффициент детерминации.

Для принятия решения о значимости уравнения регрессии в целом необходимо сравнить расчетное значение F статистики с ее критическим значением ($F_{\text{критическое}}$). При этом, если расчетное значение $F > F_{\text{критическое}}$, то нулевая гипотеза о незначимости уравнения регрессии в целом отклоняется, и мы заключаем, что уравнение регрессии в целом значимо. А если значение $F_{\text{расч}} < F_{\text{критическое}}$, то нулевая гипотеза о незначимости уравнения регрессии в целом не может быть отклонена, и мы заключаем, что уравнение регрессии в целом незначимо. F статистика для проверки гипотезы о значимости уравнения в целом в результатах представлена в таблице «Дисперсионный анализ». Степени свободы статистики указаны в столбце df , значение $k - 1$ выведено в строке «Регрессия», $n - k$ в строке «Остаток». Для нашего примера расчетное значение статистики Фишера равно $F = 14,54$.

Критическое значение статистики Фишера в Excel можно получить, используя функцию FРАСПОБР . В нашем примере критическое значение статистики Фишера для уровня значимости $\alpha = 5\% = 0,05$ равно $F_{\text{критическое}} = \text{FРАСПОБР}(\alpha = 0,05; k - 1 = 2 - 1 = 1; n - k = 11 - 2 = 9) = \text{FРАСПОБР}(0,05; 1; 9) = 5,117355029$. (Можно найти табличное значение в таблице значений F -критерия Фишера при уровне значимости $\alpha = 0,05$ и степенях свободы $k_1 = 1$; $k_2 = 9$ и оно также будет равно 5,12). Таким образом, расчетная статистика равная $F(k-1, n-k) = F(1, 9) = 14,54 > F_{\text{критическое}} = 5,12$ и мы отклоняем нулевую гипотезу о незначимости в целом уравнения регрессии и делаем вывод о статистической значимости уравнения регрессии в целом на 5 %-ом уровне значимости.

Для проверки значимости уравнения регрессии в целом можно посмотреть Значимость F . На уровне значимости $\alpha = 0,05$ уравнение регрессии признается значимым в целом, если Значимость $F \leq 0,05$.

В нашем случае значимость F меньше 0,05 (0,00912). Таким образом, полученное уравнение в целом значимо. В последней таблице приведены значения параметров (коэффициентов) модели, их стандартные ошибки и расчетные значения t -критериев Стьюдента для оценки значимости отдельных параметров модели.

Анализ данной таблицы позволяет сделать вывод о том, что на уровне значимости $\alpha = 0,05$ значимым оказывается при факторе X_1 , так как лишь для него P -значение меньше 0,05.

Уравнение регрессии имеет вид:

$$y = 10794,5 + 0,52472663X_1.$$

Критерии оценивания:

- оценка значимости коэффициентов регрессионной модели и в целом уравнения;
- запись регрессионного уравнения.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Почитайте текст задания. Продумайте логику и полноту ответа. Запишите развернутый и обоснованный ответ.

Как методы снижения размерности данных помогают в статистическом анализе? Приведите примеры их применения.

Время выполнения – 20 мин.

Ожидаемый результат:

Методы снижения размерности данных играют важную роль в статистическом анализе, особенно когда приходится работать с большими объемами информации, содержащей множество переменных. Эти методы позволяют упростить данные, сохраняя при этом их ключевые характеристики, что облегчает анализ, визуализацию и интерпретацию результатов. Основная цель снижения размерности — уменьшить количество переменных, устранив избыточность и шум, при этом сохранив важную информацию.

Основные методы снижения размерности:

Метод главных компонент:

Преобразует исходные переменные в новый набор некоррелированных переменных (главных компонент), которые объясняют максимальную дисперсию данных. Первые несколько компонент обычно содержат большую часть информации, что позволяет сократить количество переменных.

Пример применения: В анализе потребительских расходов РСА может быть использован для выявления основных факторов, влияющих на структуру потребления (например, доход, возраст, регион проживания).

Факторный анализ:

Этот метод используется для выявления скрытых (латентных) факторов, которые объясняют корреляции между наблюдаемыми переменными. Факторный анализ помогает сократить количество переменных, группируя их в несколько факторов.

Пример применения: В социологических исследованиях факторный анализ может быть использован для выявления основных факторов, влияющих на уровень удовлетворенности жизнью (например, доход, здоровье, образование).

Критерии оценивания:

- понимание концепции снижения размерности данных;
- описание метода главных компонент и факторного анализа;
- примеры использования.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Экспертное заключение

Представленный фонд оценочных средств (далее – ФОС) по дисциплине «Современные технологии обработки статистических данных» соответствует требованиям ФГОС ВО.

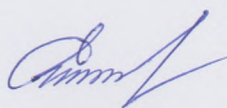
Предлагаемые формы и средства текущего и промежуточного контроля адекватны целям и задачам реализации основной профессиональной образовательной программы по направлению подготовки 38.03.05 Бизнес-информатика.

Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины представлены в полном объеме.

Виды оценочных средств, включенные в представленный фонд, отвечают основным принципам формирования ФОС.

Разработанный и представленный для экспертизы фонд оценочных средств рекомендуется к использованию в процессе подготовки обучающихся по указанному направлению.

Председатель учебно-методической
комиссии экономического института



Шаповалова Е.Н.

Лист изменения и дополнений

№ п/п	Виды дополнений и изменений	Дата и номер протокола заседания кафедры (кафедр, на котором были рассмотрены и одобрены изменения и дополнения	Подпись (с расшифровкой) заведующего кафедрой (заведующих кафедрами)