



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**
федеральное государственное бюджетное образовательное учреждение
высшего образования
«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ФГБОУ ВО «ИГУ»

Факультет бизнес-коммуникаций и информатики
Кафедра естественнонаучных дисциплин

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине Б1.В.03 Анализ данных

направление подготовки 09.03.03 Прикладная информатика

направленность (профиль) Прикладная информатика (разработка программного обеспечения)

Одобрено
УМК факультета бизнес-коммуникаций
и информатики

Разработан в соответствии с ФГОС ВО

с учетом требований проф. стандарта

Председатель УМК

В.К. Карнаухова
ФИО, должность, ученая степень, звание

подпись, печать

Разработчики:

(подпись)

доцент
(занимаемая должность)

О.И. Бернгардт
(инициалы, фамилия)

Цель фонда оценочных средств. Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Б1.В.03 Анализ данных». Перечень видов оценочных средств соответствует рабочей программе дисциплины.

Фонд оценочных средств включает контрольные материалы для проведения текущего контроля (в следующих формах: тест, устный опрос, практическое задание) и промежуточной аттестации в форме вопросов и заданий к зачету с оценкой.

Структура и содержание заданий – задания разработаны в соответствии с рабочей программой дисциплины «Б1.В.03 Анализ данных».

1. Паспорт фонда оценочных средств

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-1 Способность воспринимать математические, естественнонаучные, социально-экономические, инженерные знания, самостоятельно приобретать, развивать и применять их для формализации и решения задач разработки и модификации программного обеспечения	ПК-1.1	Знать специальные разделы математики, естественнонаучных и социально-экономических дисциплин необходимые для логического осмысления и обработки информации в профессиональной деятельности
	ПК-1.2	Уметь применять системный подход и математические методы в формализации решения прикладных задач разработки и модификации программных приложений
	ПК-1.3	Владеть навыками использования математических, естественнонаучных, социально-экономических, инженерных знаний в разработке компьютерных моделей и прототипов программного обеспечения для решения проектных и научно-исследовательских задач

2. Показатели и критерии оценивания компетенций, шкалы оценивания

2.1. Показатели и критерии оценивания компетенций

№ п/п	Раздел, тема	Код индикатора компетенции	Наименование ОС	
			ТК	ПА
1	Введение. Виды анализа. Типы данных	ПК-1.1	Тест, УО	Тест
2	Базовые техники визуализации экспериментальных данных	ПК-1.1, ПК-1.2	Тест, УО	Тест
3	Основы статистики и центральная предельная теорема	ПК-1.3	Тест, УО	Тест
4	Доверительные интервалы.	ПК-1.3	Тест, УО	Тест

№ п/п	Раздел, тема	Код индикатора компетенции	Наименование ОС	
			ТК	ПА
5	Проверка гипотез. Существование линейного тренда.	ПК-1.3, ПК-1.2	Тест, УО, Пз	Тест
6	Сравнение моделей. Метод максимального правдоподобия	ПК-1.3, ПК-1.1	Тест, УО	Тест
7	Корреляционный анализ. Линейная и ранговая корреляция.	ПК-1.1, ПК-1.2	Тест, УО, Пз	Тест
8	DBSCAN. Смесь гауссовых функций. Е-М алгоритм. Агломеративные и дивизивные методы.	ПК-1.2, ПК-1.1	Тест, УО	Тест
9	Понижение размерности данных. Одномерные и многомерные методы. Линейные и нелинейные методы.	ПК-1.3, ПК-1.1	Тест, УО, Пз	Тест
10	Обнаружение аномалий. Борьба с выбросами	ПК-1.1	Тест, УО, Пз	Тест
11	Матричные преобразования и разложения	ПК-1.3, ПК-1.2	Тест, УО	Тест
12	Ортогональные преобразования	ПК-1.3, ПК-1.1	Тест, УО	Тест
13	Преобразование Фурье. DFT и FFT. Частотный анализ	ПК-1.2, ПК-1.1	Тест, УО, Пз	Тест
14	Артефакты преобразования. Окно преобразования. MUSIC	ПК-1.3, ПК-1.1	Тест, УО	Тест
15	Использование преобразования Фурье для анализа звука. Спектрограммы	ПК-1.3, ПК-1.1	Тест, УО, Пз	Тест
16	Детерминированные сигналы	ПК-1.1	Тест, УО, Пз	Тест
17	Случайные сигналы	ПК-1.3, ПК-1.1	Тест, УО, Пз	Тест
18	Стационарные сигналы	ПК-1.1	Тест, УО	Тест
19	Нестационарные сигналы	ПК-1.3, ПК-1.1	Тест, УО	Тест
20	Методы устранения нестационарностей	ПК-1.3, ПК-1.1	Тест, УО	Тест
21	Линейная регрессия	ПК-1.3, ПК-1.1	Тест, УО	Тест
22	Модели ARIMA и SARIMAX	ПК-1.2, ПК-1.1	Тест, УО	Тест
23	Авторегрессия	ПК-1.3, ПК-1.1	Тест, УО, Пз	Тест
24	Виды статистических распределений	ПК-1.1	Тест, УО, Пз	Тест
25	Параметры случайных величин	ПК-1.1	Тест, УО	Тест

2.2. Критерии оценивания результатов обучения для текущего контроля успеваемости и промежуточной аттестации

Оценочное средство	Критерии оценивания	Шкала оценивания
Тест	Студентом даны правильные ответы на 91-100% заданий	Отлично
	Студентом даны правильные ответы на 81-90% заданий	Хорошо
	Студентом даны правильные ответы на 71-80% заданий	Удовлетворительно
	Студентом даны правильные ответы менее чем на 70% заданий	Неудовлетворительно

Оценочное средство	Критерии оценивания	Шкала оценивания
Устный опрос	Ответ соответствует поставленной теме и содержит ответы на поставленные задачи, имеет четкую структуру, логически сопоставляемую с поставленными вопросами. Ответ демонстрирует способность анализировать и обобщать информацию, опираясь на знания, полученные в ходе изучения темы, а также демонстрировать самостоятельность автора в решении поставленных задач. Ответ содержит качественную речь и аргументацию, которая убедительно подтверждает выводы и ответы на поставленные вопросы	Отлично
	Ответ должен быть направлен на ответ на поставленные вопросы и соответствовать поставленной теме, иметь логическую цепочку рассуждений и четко демонстрировать связь между поставленными вопросами. Ответ выдержан в четкой форме, быть грамотно и без ошибок озвучен, выделены ключевые термины. Ответ должен демонстрировать способность анализировать и критически оценивать информацию, выбирая ключевые аспекты и выделяя главные выводы	Хорошо
	Ответ должен соответствовать поставленной теме и содержать ответы на поставленные вопросы, должен содержать существенную информацию, ясно передавать ответы и идеи. Ответ должен содержать достаточное количество аргументов и примеров, связанных с темой работы и позволяющих изложить свою точку зрения. Ответ должен быть грамотно сформулирован	Удовлетворительно
	Ответ не соответствует поставленной теме или не содержит ответов на поставленные задачи, содержит недостаточно аргументации и примеров, которые подтверждают высказанные в ответе идеи и выводы. Ответ не соответствует логической цепочке рассуждений и не выполняет требования логической последовательности высказывания, затрудняющей понимание ответа. Ответ содержит грубые ошибки, что затрудняет понимание высказывания	Неудовлетворительно

Оценочное средство	Критерии оценивания	Шкала оценивания
Практическое задание	Задание выполнено верно. Выбран оптимальный путь решения. Присутствует развернутое описание алгоритма решения	Отлично
	Задание выполнено верно. Допущены негрубые логические ошибки при описании алгоритма решения. Отсутствуют пояснения к решению задания	Хорошо
	Ход решения задания верный, но допущены ошибки приведшие к неправильному ответу	Удовлетворительно
	В работе получен неверный ответ, связанный с грубыми ошибками допущенными в ходе решения, либо решение отсутствует полностью	Неудовлетворительно

2.3. Оценочные средства для текущего контроля (примеры)

2.3.1. Материалы для компьютерного тестирования обучающихся

Общие критерии оценивания

Процент правильных ответов	Оценка
91% – 100%	5 (отлично)
81% – 90%	4 (хорошо)
71% – 80%	3 (удовлетворительно)
Менее 70%	2 (неудовлетворительно)

Соответствие вопросов теста индикаторам формируемых и оцениваемых компетенций

№ вопроса в тесте	Код индикатора компетенции
1	ПК-1.3
2	ПК-1.3
3	ПК-1.3
4	ПК-1.1
5	ПК-1.3
6	ПК-1.1
7	ПК-1.1
8	ПК-1.1
9	ПК-1.1
10	ПК-1.3
11	ПК-1.3
12	ПК-1.3
13	ПК-1.2
14	ПК-1.3
15	ПК-1.2

№ вопроса в тесте	Код индикатора компетенции
16	ПК-1.3
17	ПК-1.1
18	ПК-1.3
19	ПК-1.3
20	ПК-1.2
21	ПК-1.3
22	ПК-1.3
23	ПК-1.1
24	ПК-1.3
25	ПК-1.1

Ключ ответов

№ вопроса в тесте	Номер ответа (или ответ, или соответствие)
1	c
2	a
3	d
4	c
5	a
6	b, c, d
7	c
8	a, c
9	b
10	a
11	c
12	b
13	1a, 2b
14	b
15	a
16	b
17	b
18	b
19	a
20	b
21	c
22	b
23	b
24	a
25	d

Перечень тестовых вопросов

№ 1. Задание с единичным выбором. Выберите один правильный ответ.

Что из перечисленного не является матричным преобразованием?

- a. Сингулярное матричное разложение (PCA/SVD)
- b. Неотрицательная матричная факторизация (NMF)
- c. Кластеризация методом DBSCAN

№ 2. Задание с единичным выбором. Выберите один правильный ответ.

В чем состоит принцип сжатия изображений с помощью ортогональных преобразований?

- a. Убрать паттерны наименьшей амплитуды, паттерны ищутся ортогональным преобразованием
- b. Убрать паттерны наибольшей амплитуды, паттерны ищутся ортогональным преобразованием

№ 3. Задание с множественным выбором. Выберите 1 правильный ответ.

В каких задачах можно использовать проверку гипотез?

- a. Проверка равенства медианы заданной величине
- b. Проверка равенства среднего заданной величине
- c. Проверка наличия линейного тренда
- d. Все перечисленное

№ 4. Задание с единичным выбором. Выберите один правильный ответ.

Что не является техникой визуализации экспериментальных данных?

- a. столбчатая диаграмма
- b. график
- c. все являются
- d. круговая диаграмма

№ 5. Задание с единичным выбором. Выберите один правильный ответ.

Какая формулировка ближе всего к центральной предельной теореме?

- a. Сумма большого числа независимых случайных величин имеет нормальное распределение
- b. Сумма большого числа независимых случайных величин имеет биномиальное распределение
- c. Сумма большого числа независимых случайных величин имеет распределение Пуассона

№ 6. Задание с множественным выбором. Выберите 3 правильных ответа.

Перечислите параметры распределений случайных величин.

- a. Наименее вероятное значение
- b. Дисперсия
- c. Среднее значение
- d. Наиболее вероятное значение

№ 7. Задание с единичным выбором. Выберите один правильный ответ.

Наиболее близкое определение корреляции

- a. Уровень определенности
- b. Уровень непохожести
- c. Уровень похожести
- d. Уровень значимости

№ 8. Задание с множественным выбором. Выберите 2 правильных ответа.

На какие типы обычно можно разделить данные в Data Science?

- a. упорядоченные/неупорядоченные
- b. важные/неважные
- c. категориальные/числовые
- d. правильные/неправильные

№ 9. Задание с единичным выбором. Выберите один правильный ответ.

Что такое стационарный сигнал?

- a. Это сигнал, значение которого постоянно во времени
- b. Это сигнал, значение статистических характеристик (средняя амплитуда, дисперсия) которого не меняется во времени

№ 10. Задание с единичным выбором. Выберите один правильный ответ.

Какие ряды данных можно считать нестационарными?

- a. Ряд значений скорости автомобиля
- b. Звук хода механических часов

№ 11. Задание с единичным выбором. Выберите один правильный ответ.

Выберите наиболее близкое определение случайного сигнала

- a. Сигнал, для которого нам известно только его медианное значение
- b. Сигнал, для которого нам известно только его среднее значение
- c. Сигнал, который мы не можем точно предсказать

№ 12. Задание с единичным выбором. Выберите один правильный ответ.

Как строится Фурье-спектрограмма?

- a. От всей звукозаписи вычисляется Фурье-преобразование, берется от него модуль. Этот спектр и является спектрограммой
- b. Временное окно заданной формы перемещается по звукозаписи, по нему вычисляется Фурье-преобразование, берется от него модуль. Последовательность Фурье спектров и является спектрограммой

№ 13. Задание на соответствие. Соотнесите элементы двух списков.

Поставить в соответствие

- | | |
|---------------------------------|-----------------------------------|
| 1. Статистическая кластеризация | a. Смесь нормальных распределений |
| 2. Метрическая кластеризация | b. DBSCAN |

№ 14. Задание с единичным выбором. Выберите один правильный ответ.

Правдоподобие это...

- a. Диапазон значений случайной величины такой, что 100% всех значений этой переменной попадает в него
- b. Вероятность того, что все полученные экспериментальные наблюдения удовлетворяют заданной модели

№ 15. Задание с единичным выбором. Выберите один правильный ответ.

Почему модели ARIMA и SARIMAX редко используются при анализе больших данных?

а. Они требуют большого количества машинных ресурсов (времени и памяти) для вычислений

б. Они требуют сложного программного обеспечения для вычислений

№ 16. Задание с единичным выбором. Выберите один правильный ответ.

Для какой задачи можно использовать авторегрессию?

а. Для определения марки автомобиля по его изображению

б. Для решения задачи прогноза температуры на улице по ее предыдущим значениям

№ 17. Задание с единичным выбором. Выберите один правильный ответ.

Какая формулировка ближе всего к определению аномалии?

а. Существенное отклонение величины от нуля

б. Существенное отклонение величины от ее модели

с. Существенное отклонение величины от ее среднего значения

№ 18. Задание с единичным выбором. Выберите один правильный ответ.

Какой метод обычно не используют для понижения размерности данных?

а. Метод многомерного шкалирования (MDS)

б. Кластеризация на основе плотности (DBSCAN)

с. Метод стохастической кодировки соседей (TSNE)

д. Метод главных компонент (PCA)

№ 19. Задание с единичным выбором. Выберите один правильный ответ.

Какие методы устранения нестационарностей ряда вы знаете?

а. Все перечисленное

б. Создание нового ряда с помощью вычисления производной (разности значений) предыдущего ряда

с. Создание нового ряда с помощью вычитания тренда из предыдущего ряда

№ 20. Задание с единичным выбором. Выберите один правильный ответ.

Для чего можно использовать преобразование Фурье?

а. Для определения размерности данных

б. Для поиска периодичностей в данных и их амплитуд

с. Для поиска пропусков в данных и их положений

№ 21. Задание с единичным выбором. Выберите один правильный ответ.

В каких задачах можно использовать линейную регрессию?

а. В задачах кластеризации

б. В задачах классификации

с. В задачах прогноза

№ 22. Задание с единичным выбором. Выберите один правильный ответ.

Для чего нужен доверительный интервал?

а. Для оценки диапазона значений, в который попадает 100% значений случайной величины

б. Для оценки диапазона значений, в который попадает заданный процент значений

случайной величины

№ 23. Задание с единичным выбором. Выберите один правильный ответ.

Выберите детерминированный сигнал

- a. Звуки за окном
- b. Звук ноты ля первой октавы

№ 24. Задание с единичным выбором. Выберите один правильный ответ.

Для чего используют различные типы окон в преобразованиях Фурье?

- a. Чтобы более точно определять количество и амплитуду периодических компонент в сигнале
- b. Чтобы убрать значимые периодические компоненты из спектра сигнала

№ 25. Задание с единичным выбором. Выберите один правильный ответ.

Какое из этих распределений непрерывное?

- a. Бернулли
- b. Биномиальное
- c. Пуассона
- d. Нормальное

2.3.2. Практические задания для оценки компетции «ПК-1.3»

№ 1. Проверка гипотез. Существование линейного тренда.

Проверить гипотезу о наличии тренда в заданных данных методом проверки разностей средних уровней.

№ 2. Корреляционный анализ. Линейная и ранговая корреляция.

Рассчитать коэффициенты корреляции Пирсона и Спирмена между тремя рядами случайных величин, проинтерпретировать результаты.

№ 3. Понижение размерности данных. Одномерные и многомерные методы. Линейные и нелинейные методы.

Отобразить заданный датасет многомерных данных на двумерном графике, понизив размерность данных методами PCA, MDS и TSNE

№ 4. Обнаружение аномалий. Борьба с выбросами.

Обнаружить выбросы значений в датасете случайной величины, удовлетворяющей нормальному распределению.

№ 5. Использование преобразования Фурье для анализа звука. Спектрограммы.

При помощи `scipy.signal.spectrogram` нарисовать спектрограмму звуков кита (пример разбирается в классе)

№ 6. Авторегрессия.

С помощью класса `sklearn.linear_model.LinearRegression` обучить программу авторегрессионного прогноза данных (аналогичный пример разбирается в классе)

№ 7. Виды статистических распределений.

Определить тип статистического распределения по его описанию в практической задаче (нормальное, равномерное, Бернулли, биномиальное, Пуассона)

2.3.3. Практические задания для оценки компетции «ПК-1.2»

№ 8. Преобразование Фурье. DFT и FFT. Частотный анализ.

Выполнить преобразование Фурье с заданным числовым рядом, пользуясь

библиотекой scipy.fft. Определить количество наиболее интенсивных гармоник.

№ 9. Детерминированные сигналы.

Предложить для каждого ряда данных его детерминированную модель и приближенную формулу.

№ 10. Случайные сигналы.

Определить по данным, какие сигналы можно считать случайными, а какие - детерминированными

3. Промежуточная аттестация

3.1. Методические материалы, определяющие процедуру оценивания знаний, умений, навыков и опыта деятельности

Зачет является заключительным этапом процесса формирования компетенций обучающегося при изучении дисциплины и имеет целью проверку и оценку знаний обучающегося по теории, и применению полученных знаний, умений и навыков при решении практических задач.

Зачет проводится по расписанию, сформированному учебно-методическим управлением, в сроки, предусмотренные календарным учебным графиком. Зачет принимается преподавателем, ведущим лекционные занятия.

Зачет проводится только при предъявлении обучающимся зачетной книжки и при условии выполнения всех контрольных мероприятий, предусмотренных учебным планом и рабочей программой дисциплины. Обучающимся на зачете представляется право выбрать один из билетов. Время подготовки к ответу составляет 30 минут. По истечении установленного времени обучающийся должен ответить на вопросы билета. Результаты зачета оцениваются по четырехбалльной системе и заносятся в зачетно-экзаменационную ведомость и зачетную книжку. В зачетную книжку заносятся только положительные оценки. Подписанный преподавателем экземпляр ведомости сдается не позднее следующего дня в деканат.

В случае неявки обучающегося на зачет в зачетно-экзаменационную ведомость делается отметка «не явка». Обучающиеся, не прошедшие промежуточную аттестацию по дисциплине, должны ликвидировать академическую задолженность в установленном локальными нормативными актами порядке.

3.2. Вопросы к зачету с оценкой

№	Вопрос	Код компетенции
1.	На какие типы обычно можно разделить данные в Data Science?	ПК-1.1
2.	Какие базовые техники визуализации экспериментальных данных вы знаете?	ПК-1.1
3.	Как формулируется центральная предельная теорема?	ПК-1.3
4.	Для чего нужен доверительный интервал?	ПК-1.3
5.	В каких задачах можно использовать проверку гипотез?	ПК-1.1
6.	Для чего используется метод максимального правдоподобия?	ПК-1.3
7.	Какие виды линейной и ранговой корреляции вы знаете?	ПК-1.2
8.	Что такое кластеризация данных? Какие виды кластеризации вы знаете?	ПК-1.3
9.	Какие методы понижения размерности данных вы знаете?	ПК-1.3
10.	Что такое аномалия и выброс?	ПК-1.1

№	Вопрос	Код компетенции
11.	Какие матричные преобразования вы знаете? Описать принципы каждого.	ПК-1.3
12.	В чем состоит принцип сжатия изображений с помощью ортогональных преобразований?	ПК-1.3
13.	Для чего используют преобразование Фурье и в чем его принцип?	ПК-1.2
14.	Для чего используют различные типы окон в преобразованиях Фурье? Какие окна вы знаете?	ПК-1.2
15.	Для чего используется спектрограмма, как она строится?	ПК-1.3
16.	Какие сигналы можно считать детерминированными? Приведите примеры.	ПК-1.1
17.	Какие сигналы можно считать случайными?	ПК-1.1
18.	Что такое стационарный сигнал? Как можно оценить, стационарен сигнал или нет?	ПК-1.1
19.	Какие ряды данных можно считать нестационарными? Привести примеры. Как можно это проверить?	ПК-1.2
20.	Какие методы устранения нестационарностей ряда вы знаете?	ПК-1.1
21.	В каких задачах можно использовать линейную регрессию? Как она строится?	ПК-1.2
22.	В чем основное отличие моделей ARIMA и SARIMAX от линейной регрессии?	ПК-1.3
23.	В чем основное отличие авторегрессии от регрессии	ПК-1.3
24.	Чем отличаются непрерывные распределения случайных величин от дискретных? Привести примеры.	ПК-1.1
25.	Какие параметры распределений случайных величин вы знаете? Как их определить?	ПК-1.2

3.3. Тематика курсовых работ

По данной дисциплине выполнение курсовых проектов (работ) не предусматривается.

3.4. Материалы для компьютерного тестирования обучающихся

Общие критерии оценивания

Процент правильных ответов	Оценка
91% – 100%	5 (отлично)
81% – 90%	4 (хорошо)
71% – 80%	3 (удовлетворительно)
Менее 70%	2 (неудовлетворительно)

Соответствие вопросов теста индикаторам формируемых и оцениваемых компетенций

№ вопроса в тесте	Код индикатора компетенции
1	ПК-1.3
2	ПК-1.3

№ вопроса в тесте	Код индикатора компетенции
3	ПК-1.3
4	ПК-1.1
5	ПК-1.3
6	ПК-1.1
7	ПК-1.1
8	ПК-1.1
9	ПК-1.1
10	ПК-1.3
11	ПК-1.3
12	ПК-1.3
13	ПК-1.2
14	ПК-1.3
15	ПК-1.2
16	ПК-1.3
17	ПК-1.1
18	ПК-1.3
19	ПК-1.3
20	ПК-1.2
21	ПК-1.3
22	ПК-1.3
23	ПК-1.1
24	ПК-1.3
25	ПК-1.1

Ключ ответов

№ вопроса в тесте	Номер ответа (или ответ, или соответствие)
1	с
2	а
3	d
4	с
5	а
6	b, c, d
7	с
8	а, с
9	b
10	а
11	с
12	b
13	1а, 2b

№ вопроса в тесте	Номер ответа (или ответ, или соответствие)
14	b
15	a
16	b
17	b
18	b
19	a
20	b
21	c
22	b
23	b
24	a
25	d

Перечень тестовых вопросов

№ 1. Задание с единичным выбором. Выберите один правильный ответ.

Что из перечисленного не является матричным преобразованием?

- a. Сингулярное матричное разложение (PCA/SVD)
- b. Неотрицательная матричная факторизация (NMF)
- c. Кластеризация методом DBSCAN

№ 2. Задание с единичным выбором. Выберите один правильный ответ.

В чем состоит принцип сжатия изображений с помощью ортогональных преобразований?

- a. Убрать паттерны наименьшей амплитуды, паттерны ищутся ортогональным преобразованием
- b. Убрать паттерны наибольшей амплитуды, паттерны ищутся ортогональным преобразованием

№ 3. Задание с множественным выбором. Выберите 1 правильный ответ.

В каких задачах можно использовать проверку гипотез?

- a. Проверка равенства медианы заданной величине
- b. Проверка равенства среднего заданной величине
- c. Проверка наличия линейного тренда
- d. Все перечисленное

№ 4. Задание с единичным выбором. Выберите один правильный ответ.

Что не является техникой визуализации экспериментальных данных?

- a. столбчатая диаграмма
- b. график
- c. все являются
- d. круговая диаграмма

№ 5. Задание с единичным выбором. Выберите один правильный ответ.

Какая формулировка ближе всего к центральной предельной теореме?

а. Сумма большого числа независимых случайных величин имеет нормальное распределение

б. Сумма большого числа независимых случайных величин имеет биномиальное распределение

с. Сумма большого числа независимых случайных величин имеет распределение Пуассона

№ 6. Задание с множественным выбором. Выберите 3 правильных ответа.

Перечислите параметры распределений случайных величин.

а. Наименее вероятное значение

б. Дисперсия

с. Среднее значение

д. Наиболее вероятное значение

№ 7. Задание с единичным выбором. Выберите один правильный ответ.

Наиболее близкое определение корреляции

а. Уровень определенности

б. Уровень непохожести

с. Уровень похожести

д. Уровень значимости

№ 8. Задание с множественным выбором. Выберите 2 правильных ответа.

На какие типы обычно можно разделить данные в Data Science?

а. упорядоченные/неупорядоченные

б. важные/неважные

с. категориальные/числовые

д. правильные/неправильные

№ 9. Задание с единичным выбором. Выберите один правильный ответ.

Что такое стационарный сигнал?

а. Это сигнал, значение которого постоянно во времени

б. Это сигнал, значение статистических характеристик (средняя амплитуда, дисперсия) которого не меняется во времени

№ 10. Задание с единичным выбором. Выберите один правильный ответ.

Какие ряды данных можно считать нестационарными?

а. Ряд значений скорости автомобиля

б. Звук хода механических часов

№ 11. Задание с единичным выбором. Выберите один правильный ответ.

Выберите наиболее близкое определение случайного сигнала

а. Сигнал, для которого нам известно только его медианное значение

б. Сигнал, для которого нам известно только его среднее значение

с. Сигнал, который мы не можем точно предсказать

№ 12. Задание с единичным выбором. Выберите один правильный ответ.

Как строится Фурье-спектрограмма?

а. От всей звукозаписи вычисляется Фурье-преобразование, берется от него модуль.

Этот спектр и является спектрограммой

б. Временное окно заданной формы перемещается по звукозаписи, по нему вычисляется Фурье-преобразование, берется от него модуль. Последовательность Фурье спектров и является спектрограммой

№ 13. Задание на соответствие. Соотнесите элементы двух списков.

Поставить в соответствие

- | | |
|---------------------------------|-----------------------------------|
| 1. Статистическая кластеризация | а. Смесь нормальных распределений |
| 2. Метрическая кластеризация | б. DBSCAN |

№ 14. Задание с единичным выбором. Выберите один правильный ответ.

Правдоподобие это...

- а. Диапазон значений случайной величины такой, что 100% всех значений этой переменной попадает в него
- б. Вероятность того, что все полученные экспериментальные наблюдения удовлетворяют заданной модели

№ 15. Задание с единичным выбором. Выберите один правильный ответ.

Почему модели ARIMA и SARIMAX редко используются при анализе больших данных?

- а. Они требуют большого количества машинных ресурсов (времени и памяти) для вычислений
- б. Они требуют сложного программного обеспечения для вычислений

№ 16. Задание с единичным выбором. Выберите один правильный ответ.

Для какой задачи можно использовать авторегрессию?

- а. Для определения марки автомобиля по его изображению
- б. Для решения задачи прогноза температуры на улице по ее предыдущим значениям

№ 17. Задание с единичным выбором. Выберите один правильный ответ.

Какая формулировка ближе всего к определению аномалии?

- а. Существенное отклонение величины от нуля
- б. Существенное отклонение величины от ее модели
- с. Существенное отклонение величины от ее среднего значения

№ 18. Задание с единичным выбором. Выберите один правильный ответ.

Какой метод обычно не используют для понижения размерности данных?

- а. Метод многомерного шкалирования (MDS)
- б. Кластеризация на основе плотности (DBSCAN)
- с. Метод стохастической кодировки соседей (TSNE)
- д. Метод главных компонент (PCA)

№ 19. Задание с единичным выбором. Выберите один правильный ответ.

Какие методы устранения нестационарностей ряда вы знаете?

- а. Все перечисленное
- б. Создание нового ряда с помощью вычисления производной (разности значений) предыдущего ряда
- с. Создание нового ряда с помощью вычитания тренда из предыдущего ряда

№ 20. Задание с единичным выбором. Выберите один правильный ответ.

Для чего можно использовать преобразование Фурье?

- a. Для определения размерности данных
- b. Для поиска периодичностей в данных и их амплитуд
- c. Для поиска пропусков в данных и их положений

№ 21. Задание с единичным выбором. Выберите один правильный ответ.

В каких задачах можно использовать линейную регрессию?

- a. В задачах кластеризации
- b. В задачах классификации
- c. В задачах прогноза

№ 22. Задание с единичным выбором. Выберите один правильный ответ.

Для чего нужен доверительный интервал?

- a. Для оценки диапазона значений, в который попадает 100% значений случайной величины
- b. Для оценки диапазона значений, в который попадает заданный процент значений случайной величины

№ 23. Задание с единичным выбором. Выберите один правильный ответ.

Выберите детерминированный сигнал

- a. Звуки за окном
- b. Звук ноты ля первой октавы

№ 24. Задание с единичным выбором. Выберите один правильный ответ.

Для чего используют различные типы окон в преобразованиях Фурье?

- a. Чтобы более точно определять количество и амплитуду периодических компонент в сигнале
- b. Чтобы убрать значимые периодические компоненты из спектра сигнала

№ 25. Задание с единичным выбором. Выберите один правильный ответ.

Какое из этих распределений непрерывное?

- a. Бернулли
- b. Биномиальное
- c. Пуассона
- d. Нормальное