Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Новосибирский государственный технический университет»
Кафедра автоматизированных систем управления

Паспорт экзамена

по дисциплине «Математическое обеспечение систем обработки данных», 2 семестр

1. Методика оценки

Экзамен проводится в устной форме, по билетам. Билет формируется по следующему правилу: первый вопрос выбирается из диапазона вопросов 1–16, второй вопрос из диапазона вопросов 17–32 (список вопросов приведен ниже). В ходе экзамена преподаватель вправе задавать студенту дополнительные вопросы из общего перечня (п. 4).

Форма экзаменационного билета

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
Факультет АВТФ

Билет № _____
к экзамену по дисциплине «Математическое обеспечение систем обработки данных»

1. Разведочный анализ данных. Основные понятия, инструменты и методы разведочного анализа данных. Формулирование и проверка простейших гипотез (12 баллов)
2. Технология факторного анализа в компьютерных системах анализа данных (12 баллов)
3. Практическое задание. (16 баллов). Провести идентификацию заданного временного ряда.

Утверждаю: зав. кафедрой ____ _____ должность, ФИО
(подпись)
(дата)

2. Критерии оценки

- Ответ на экзаменационный билет (тест) считается **неудовлетворительным**, если студент при ответе на вопросы не дает определений основных понятий, не способен показать причинно-следственные связи явлений, при решении задачи допускает принципиальные ошибки, оценка составляет *0-10 баллов.*

- Ответ на экзаменационный билет (тест) засчитывается на **пороговом** уровне, если студент при ответе на вопросы дает определение основных понятий, может показать причинно-следственные связи явлений, при решении задачи допускает непринципиальные ошибки, например, вычислительные, оценка составляет *10-20 баллов.*
- Ответ на экзаменационный билет (тест) билет засчитывается на **базовом** уровне, если студент при ответе на вопросы формулирует основные понятия, законы, дает характеристику процессов, явлений, проводит анализ причин, условий, может представить качественные характеристики процессов, не допускает ошибок при решении задачи, оценка составляет *20-30 баллов.*
- Ответ на экзаменационный билет (тест) билет засчитывается на **продвинутом** уровне, если студент при ответе на вопросы проводит сравнительный анализ подходов, проводит комплексный анализ, выявляет проблемы, предлагает механизмы решения, способен представить количественные характеристики определенных процессов, приводит конкретные примеры из практики, не допускает ошибок и способен обосновать выбор метода решения задачи, оценка составляет 30-40 *баллов.*

3. Шкала оценки

В общей оценке по дисциплине экзаменационные баллы учитываются в соответствии с правилами балльно-рейтинговой системы, приведенными в рабочей программе дисциплины.

4. **Вопросы к** экзамену **по дисциплине «**Математическое обеспечение систем обработки данных**»**

1. Разведочный анализ данных. Основные понятия, инструменты и методы разведочного анализа данных. Формулирование и проверка простейших гипотез.
2. Общая схема проверки статистических гипотез. Основная и альтернативная гипотезы. Ошибка 1 и 2 рода. Статистический критерий. Мощность критерия. Процедура проверки гипотезы.
3. Проверка гипотез о виде распределения. Критерий Хи-квадрат и критерий Колмогорова. Основные виды распределений. Используемых в анализе данных.
4. t-критерий для одно- и двухвыборочных данных, для категоризованных данных.
5. Однофакторный и многофакторный дисперсионный анализ.
6. Математическая постановка задачи линейного регрессионного анализа. Возможности применения регрессионного анализа в экспериментальных исследованиях. Предположения регрессионного анализа.
7. МНК-оценки параметров модели. Свойства оценок (несмещенность, состоятельность, эффективность). Понятие функции потерь. Функция потерь для МНК.
8. Экспериментальная проверка предположений регрессинного анализа. Проверка гипотез об адекватности модели (значимость регрессии) и о значимости коэффициентов регрессии. Коэффициент детерминации. Анализ остатков регрессии.
9. Исследование предположений регрессионного анализа. Проблема мультиколлинеарности и гетероскедастичности. Анализ независимости остатков. Критерий Дарбина-Уотсона.
10. Множественная регрессия. Проблема мультиколлинеарности. Вычисление толерантности. Устранение избыточности.

11. Множественная регрессия. Понятие избыточного регрессора. Методы устранения избыточных регрессоров.
12. Технология линейного регрессионного анализа в компьютерных системах анализа данных.
13. Технология нелинейного регрессионного анализа в компьютерных системах анализа данных.
14. Центральная предельная теорема.
15. Основные статистики, характеризующие математическое ожидание выборки (среднее, медиана, мода), и формулы для их вычисления.
16. Основные статистики, характеризующие разброс выборки (дисперсия, с.к.о., размах), и формулы для их вычисления.
17. Асимметрия и эксцесс. Проверка симметричности и крутизны распределения.
18. Понятие квантили, процентили, квартили. Медиана. Использование квантилей при проверке гипотез.
19. Основные типы нелинейных регрессионных моделей. Возможности сведения нелинейных моделей к линейным. Существенно нелинейные модели.
20. Логит-регрессия и пробит-регрессия и проверка их адекватности.
21. Регрессионные модели с точками разрыва. Основные типы кусочно-линейной регрессии.
22. Построение нелинейной регрессионной модели и проверка ее адекватности. Общая модель роста и ее смысл. Основные методы оптимизации.
23. Математическая постановка задачи факторного анализа. Вращение факторов. Решение задачи классификации.
24. Факторный анализ как метод сокращения размерности.
25. Методы сглаживания временных рядов (скользящее среднее и скользящая медиана, метод prior).
26. Основные компоненты временного ряда. Методы выделения трендов.
27. Декомпозиция временных рядов (метод Census I).
28. Однопараметрическое экспоненциальное сглаживание.
29. Двухпараметрическое экспоненциальное сглаживание.
30. Трехпараметрическое экспоненциальное сглаживание.
31. Проверка качества модели сглаживания. Критерии качества. Прогнозирование временных рядов.
32. Технология факторного анализа в компьютерных системах анализа данных.