

Демонстрационный вариант по направлению 01.04.02 «Прикладная математика и информатика» ОП «Анализ данных в девелопменте»

Время выполнения заданий – 240 минут.

Решения должны быть записаны четко по-русски или по-английски.

Каждая задача оценивается не более чем 10 баллами, максимальная сумма – 100 баллов.

Если будет решено более 10 задач, то засчитываются только 10 лучших решений (по числу баллов).

Задача 1.

Найдите значение предела:

а) $\lim_{x \rightarrow 0} \frac{\sin 3x^2}{(x^2+5x) \sin x}$
б) $\lim_{x \rightarrow 0} \frac{1 - \cos x}{\sqrt{x+1}-1}$

Задача 2.

Вычислите определённый интеграл:

$$\int_0^{\pi} (e^x - \cos x) dx$$

Задача 3.

Решите дифференциальное уравнение:

$$y'' + 6y' + 8y = \frac{4e^{-2x}}{2 + e^{2x}}$$

Задача 4.

Найдите определитель матрицы:

$$\begin{pmatrix} 1 & 5 & 2 & 2 \\ 2 & 4 & 4 & 6 \\ -3 & 8 & -5 & 3 \\ 2 & 3 & -6 & 1 \end{pmatrix}$$

Задача 5.

Решить систему линейных уравнений:

$$\begin{cases} x_1 + 2x_2 + x_3 = 0 \\ -x_1 - 2x_2 - 3x_3 = 0 \\ x_1 + 2x_2 + 2x_3 = 0 \end{cases}$$

Задача 6.

Найти собственные векторы линейного преобразования, заданного матрицей:

$$A = \begin{pmatrix} 1 & 5 & 2 \\ -3 & 8 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

Задача 7.

В группе 14 человек знают английский язык, 16 человек знают китайский язык, 20 человек знают арабский язык и 22 человека знает немецкий язык. В группе нет людей, знающих три языка, и 23 человека в группе знают ровно два языка из перечисленных. Сколько человек в группе знают ровно один язык из перечисленных?

Задача 8.

Сколько слагаемых получится после раскрытия скобок и приведения подобных слагаемых в выражении $(1 + x^2)^{100}(1 + x^3)^{100}$?

Задача 9.

Были отправлены посылки в три почтовых отделения. Вероятность своевременной доставки в первое отделение равна 0,95, во второе - 0,9, в третье - 0,8. Найти вероятность следующих событий:

- а) только одно отделение получит посылки вовремя;
- б) хотя бы одно отделение получит посылки с опозданием.

Задача 10.

Какова вероятность Вашей встречи с другом, если вы договорились встретиться в определенном месте, с 13.00 до 14.00 часов и ждете друг друга в течение 10 минут?

Задача 11.

По цели производится 4 выстрела. Вероятность попадания при этом растет так: 0,2, 0,3, 0,5, 0,7. Найти закон распределения случайной величины X – числа попаданий. Найти вероятность того, что $X \geq 1$.

Задача 12.

В нормально распределенной совокупности 15% значений меньше 12 и 40% значений больше 16.2. Найти среднее значение и стандартное отклонение данного распределения.

Для заданий 13 и 14 сдается .ipynb файл с пояснениями и кодом.

Задача 13.

Вам дан датасет Титаник со следующей информацией.

1. **PassengerId:** Уникальный индекс/номер строки. Начинается с 1 (для первой строки) и увеличивается на 1 для каждой следующей. Рассматриваем его как идентификатор строки и, что логично, идентификатор пассажира (т.к. для каждого пассажира в датасете представлена только одна строка).
2. **Survived:** Признак, показывающий был ли спасен данный пассажир или нет. 1 означает, что удалось выжить, и 0 - не удалось спастись.
3. **Pclass:** Класс билета. 1 - означает Первый класс билета. 2 - означает Второй класс билета. 3 - означает Третий класс билета.
4. **Name:** Имя пассажира. Имя также может содержать титулы и обращения. "Mr" для мужчин. "Mrs" для женщин. "Miss" для девушек (тут имеется в виду что для тех, кто не замужем, так было принято, да и сейчас тоже, говорить в западном обществе). "Master" для юношей.
5. **Sex:** Пол пассажира. Либо мужчины (=Male) либо женщины (=Female).
6. **Age:** Возраст пассажира. "NaN" значения в этой колонке означают, что возраст данного пассажира отсутствует/неизвестен/или не был записан в датасет.
7. **SibSp:** Количество братьев/сестер или супругов, путешествующих с каждым пассажиром.
8. **Parch:** Количество родителей детей (Number of parents of children travelling with each passenger).
9. **Ticket:** Номер билета.
10. **Fare:** Сумма, которую заплатил пассажир за путешествие.
11. **Cabin:** Номер каюты пассажира. "NaN" значения в этой колонке указывают на то, что номер каюты данного пассажира не был записан.
12. **Embarked:** Порт отправления данного пассажира.

Ответьте на вопросы:

- А) Есть ли зависимость между классом пассажира и выживаемостью? Каков процент выживших в каждом классе?
- Б) Какова вероятность выживания для пассажиров, путешествующих с семьей по сравнению с теми, у кого семьи не было на борту?
- В) Каков процент выживших среди пассажиров, у которых указан порт посадки "С" (Cherbourg)? Влияет ли порт посадки на выживаемость?
- Г) Существует ли связь между возрастом и классом билета у пассажиров? Какова вероятность выживания для пассажиров разных возрастных групп?

Задача 14.

По датасету из предыдущего задания построить следующие визуализации:

- А) Построить диаграмму(ы) показывающие распределение возрастов в разных классах.
- Б) Построить диаграмму(ы) показывающие процент выживших в каждом классе.