Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего профессионального образования

«Владимирский государственный университет

имени Александра Григорьевича и Николая Григорьевича Столетовых»

(ВлГУ)

Кафедра информационных систем и программной инженерии

Практическая работа № 1

по дисциплине "Математические основы анализа данных"

ТЕМА РАБОТЫ:

Основные понятия теории вероятностей. Условная вероятность

Выполнил:

студент гр. ПРИм-124

Парахин К.В.

Принял:

Доцентр кафедры ИСПИ

Курочкин С.В.

Владимир 2024 г.

Цель работы:

Получить практические навыки определения вероятности события различными методами и работе в среде разработки Jupiter Notebook.

Выполнение работы

Установим в Visual Studio Code расширение Jupiter, а также python сервер для работы с Jupiter Notebook. Также был установлен python и pip.

Задание:

Среди клиентов туристической фирмы 30% ездили в Турцию, 20% — в

Египет, 10% — в Грецию; в Турцию и Египет — 12%, в Египет и Грецию

— 5%, в Турцию и Грецию — 6%, во все три страны — 4%. Найти

вероятность того, что случайно выбранный клиент: а) ездил в Турцию или

Египет, б) ездил в Египет или Грецию, в) ездил в Турцию. Египет или

Грецию, г) не ездил ни в одну из перечисленных стран

**Решение**

Для решения задачи методом "дерево вероятностей" в Jupyter Notebook необходимо:

1. Построить дерево вероятностей:

    Рассмотрим клиентов, которые посещали некоторые страны Европы (Турцию, Египет или Грецию), посещавших их по отдельности, либо все вместе, либо непосещавших никакие из них.

    Конкретно в данной задаче нас интересуют группы: клиенты, ездившие в Турцию или Египет, езившие в Египет или Грецию, ездившие в одну из 3 стран, неездившие никуда.

    Для каждой группы у нас есть вероятность того, что случайный клиент попадает в эту группу (успех) или непопадает в группу (неудача).

2. Найти полную вероятность того, что случайный клиент попадет в каждую из этих групп, используя формулу полной вероятности.

3. Построить дерево вероятностей с помощью графических инструментов библиотеки matplotlib.

Стоит заметить, что полная вероятность (1) складывается из 2 основных исходов: случайный клиент ездил в какую то из 3 стран или не ездил никуда. Вероятность того, что клиент ездил хоть куда то - складывается из суммы вероятностей исходов: клиент ездил только в Турцию, клиент ездил только в Грецию, клиент ездил только в Египет, клиент ездил в Турцию и Египет (но не был в Греции), клиент ездил в Египет и Грецию (но не был в Турции), клиент ездил в Турцию и Грецию (но не был в Египте) и клиент был в 3 странах сразу

Вероятность того, что клиент ездил в Турцию или Египет будет складываться из суммы вероятностей: клиент ездил только в Турцию, клиент ездил только в Египет, клиент ездил в Турцию и Египет (но не был в Греции), клиент ездил в Египет и Грецию (но не был в Турции), клиент ездил в Турцию и Грецию (но не был в Египте) и клиент был в 3 странах сразу

Вероятность того, что клиент ездил в Египет или Грецию будет складываться из суммы вероятностей: клиент ездил только в Грецию, клиент ездил только в Египет, клиент ездил в Турцию и Египет (но не был в Греции), клиент ездил в Египет и Грецию (но не был в Турции), клиент ездил в Турцию и Грецию (но не был в Египте) и клиент был в 3 странах сразу

Листинг кода:

import matplotlib.pyplot as plt

# Вероятности

P\_T = 0.3    # Вероятность того, что клиент ездил в Турцию

P\_E = 0.2    # Вероятность того, что клиент ездил в Египет

P\_G = 0.1    # Вероятность того, что клиент ездил в Грецию

P\_TandE = 0.12  # Вероятность того, что клиент ездил в Турцию и Египет

P\_EandG = 0.05  # Вероятность того, что клиент ездил в Египет и Грецию

P\_TandG = 0.06  # Вероятность того, что клиент ездил в Турцию и Грецию

P\_TandEandG = 0.04 # Вероятность того, что клиент ездил в Турцию, Египет и Грецию

# Вероятность того, что клиент ездил в Турцию или Египет = 0.38

P\_TorE = P\_T + P\_E - P\_TandE

# Вероятность того, что клиент ездил в Египет или Грецию = 0.25

P\_EorG = P\_E + P\_G - P\_EandG

# Вероятность того, что клиент ездил в Турцию или Грецию (но не был в Египте) = 0.02

P\_TorGnotE = P\_TandG - P\_TandEandG

# Вероятность того, что клиент ездил в Турцию или Египте (но не был в Греции) = 0.08

P\_TorEnotG = P\_TandE - P\_TandEandG

# Вероятность того, что клиент ездил в Египет или Грецию (но не был в Турции) = 0.01

P\_EorGnotT = P\_EandG - P\_TandEandG

# Вероятность того, что клиент ездил только в Турцию = 0.3 - 0.02 - 0.08 - 0.04 = 0.16

P\_onlyT = P\_T - P\_TorGnotE - P\_TorEnotG - P\_TandEandG

# Вероятность того, что клиент ездил только в Египет = 0.2 - 0.08 - 0.01 - 0.04 = 0.07

P\_onlyE = P\_E - P\_TorEnotG - P\_EorGnotT - P\_TandEandG

# Вероятность того, что клиент ездил только в Грецию = 0.1 - 0.02 - 0.01 - 0.04 = 0.03

P\_onlyG = P\_G - P\_TorGnotE - P\_EorGnotT - P\_TandEandG

# Вероятность того, что клиент ездил в Турцию, Египет или Грецию = 0.16 + 0.03 + 0.07 + 0.02 + 0.08 + 0.01 + 0.04 = 0.41

P\_TorEorG = P\_onlyT + P\_onlyG + P\_onlyE + P\_TorGnotE + P\_TorEnotG + P\_EorGnotT + P\_TandEandG

# Вероятность того, что ни в одной стране не был = 1 - 0.41 = 0.59

P\_other = 1 - P\_TorEorG

# Построение дерева вероятностей

fig, ax = plt.subplots(figsize=(10, 6))

# Дерево вероятностей

# Первый уровень

ax.text(0.1, 0.9, 'Все клиенты', fontsize=12, ha='center')

# Второй уровень

ax.text(0.2, 0.7, 'Клиенты, ездившие\n в Турцию \n или Грецию \n или Египет\n(41%)', fontsize=11, ha='center')

ax.text(0.2, 0.5, 'Клиенты, не ездившие \n никуда\n(59%)', fontsize=11, ha='center')

# Третий уровень для тех, кто ездил в какую то страну

ax.text(0.5, 0.85, 'Только в Турцию\n(16%)', fontsize=10, ha='center')

ax.text(0.5, 0.75, 'Только в Грецию\n(3%)', fontsize=10, ha='center')

ax.text(0.5, 0.65, 'Только в Египет\n(7%)', fontsize=10, ha='center')

ax.text(0.5, 0.55, 'В Турции и Египте (не был в Греции)\n(8%)', fontsize=10, ha='center')

ax.text(0.5, 0.45, 'В Египте и Греции (не был в Турции) \n(1%)', fontsize=10, ha='center')

ax.text(0.5, 0.35, 'В Турции и Греции (не был в Египте)\n(2%)', fontsize=10, ha='center')

ax.text(0.5, 0.25, 'В каждой из стран\n(4%)', fontsize=10, ha='center')

# Четвертый уровень для тех, кто ездил в одну из двух стран (вопросы а и б)

ax.text(0.85, 0.75, 'В Турции или Египте (38%)', fontsize=10, ha='center')

ax.text(0.85, 0.35, 'В Египте или Греции (25%)', fontsize=10, ha='center')

# Соединительные линии

ax.plot([0.1, 0.2], [0.9, 0.7], 'k-')

ax.plot([0.1, 0.2], [0.9, 0.5], 'k-')

ax.plot([0.2, 0.5], [0.7, 0.85], 'k-')

ax.plot([0.2, 0.5], [0.7, 0.75], 'k-')

ax.plot([0.2, 0.5], [0.7, 0.65], 'k-')

ax.plot([0.2, 0.5], [0.7, 0.55], 'k-')

ax.plot([0.2, 0.5], [0.7, 0.45], 'k-')

ax.plot([0.2, 0.5], [0.7, 0.35], 'k-')

ax.plot([0.2, 0.5], [0.7, 0.25], 'k-')

ax.plot([0.5, 0.85], [0.85, 0.75], 'k-')

ax.plot([0.5, 0.85], [0.65, 0.75], 'k-')

ax.plot([0.5, 0.85], [0.55, 0.75], 'k-')

ax.plot([0.5, 0.85], [0.45, 0.75], 'k-')

ax.plot([0.5, 0.85], [0.35, 0.75], 'k-')

ax.plot([0.5, 0.85], [0.25, 0.75], 'k-')

ax.plot([0.5, 0.85], [0.75, 0.35], 'k-')

ax.plot([0.5, 0.85], [0.65, 0.35], 'k-')

ax.plot([0.5, 0.85], [0.55, 0.35], 'k-')

ax.plot([0.5, 0.85], [0.45, 0.35], 'k-')

ax.plot([0.5, 0.85], [0.35, 0.35], 'k-')

ax.plot([0.5, 0.85], [0.25, 0.35], 'k-')

ax.axis('off')

plt.show()

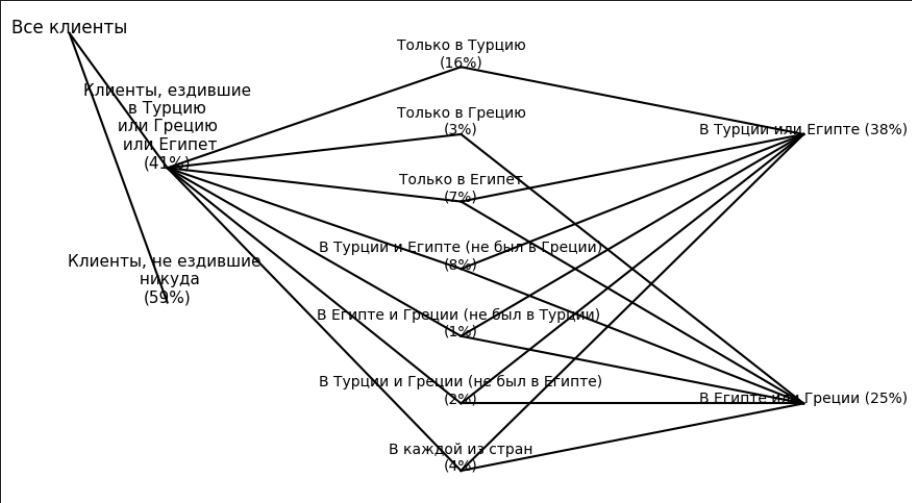


Рисунок 1. Визуализация скрипта по поиску отдельных вероятностей

Собственно, ответы на вопросы а и б находятся справа на графике, а ответы на вопросы в и г – слева на графике.

Решение задач методом геометрической вероятности

В урне 5 белых и 4 черных шара. Из урны наугад вынимают два шара. Какова вероятность того, что это будут: а) два белых шара; б) два черных шара; в) один черный и один белый.

Решение:

В данной задаче можно использовать комбинаторику и формулы поиска сочетаний без повторений для нахождения вероятностей вытаскивания шаров определенного цвета.

Шаг 1: Определение общего числа возможных исходов

Всего в урне 5 белых и 4 черных шара. Общее количество шаров в урне равно 5 + 4 = 9

Мы выбираем два шара из этих девяти. Число всех возможных C(9, 2) = 36

Шаг 2: Вычисление вероятностей

а) Вероятность вытащить два белых шара

Число способов выбрать 2 белых шара из 5 равно C(5, 2) = 10

Таким образом, вероятность вытащить два белых шара равна отношению C(5, 2) к C(9, 2) = 10/36 = 5/18

б) Вероятность вытащить два черных шара

Число способов выбрать 2 черных шара из 4 равна C(4, 2) = 6

Вероятность вытащить два черных шара равна отношению C(4, 2) к C(9, 2) = 6/36 = 1/6

в) Вероятность вытащить один белый и один черный шар

Число способов выбрать 1 белый шар из 5 и 1 черный из 4 равна произведению числа сочетаний C(5, 1) \* C(4, 1) = 5 \* 4 = 20

Вероятность вытащить один белый и один черный шар равна отношению полученного числа C(5, 1) \* C(4, 1) к полной вероятности C(9, 2): 20/36 = 5/9

Листинг кода:

import math

import matplotlib.pyplot as plt

# Общее число способов выбрать 2 шара из 9

total\_variants\_number = math.comb(9, 2)

# а) Вероятность вытащить два белых шара

two\_white\_variants\_number = math.comb(5, 2)

two\_white\_probability = two\_white\_variants\_number / total\_variants\_number

# б) Вероятность вытащить два черных шара

two\_black\_variants\_number = math.comb(4, 2)

two\_black\_probability = two\_black\_variants\_number / total\_variants\_number

# в) Вероятность вытащить один белый и один черный шар

one\_white\_one\_black\_variants\_number = math.comb(5, 1) \* math.comb(4, 1)

one\_white\_one\_black\_probability = one\_white\_one\_black\_variants\_number / total\_variants\_number

# Проверка полной вероятности: 10/36 + 6/36 + 20/36 = 1

# two\_white\_probability + two\_black\_probability + one\_white\_one\_black\_probability = 1

# Визуализация графика с полученными вероятностями

probabilities = [two\_white\_probability, two\_black\_probability, one\_white\_one\_black\_probability]

labels = ['Два белых шара - 5/18', 'Два черных шара - 1/6', 'Один белый и один черный шар - 5/9']

# Построение графика с кривыми

plt.figure(figsize=(8, 6))

# Данные для кривых

x = [1, 2, 3]  # Позиции для каждой вероятности на оси X

y = probabilities  # Вероятности для оси Y

# Построение графика кривой

plt.plot(x, y, marker='o', linestyle='-', color='b', label='Вероятности', linewidth=2, markersize=8)

# Настройки осей и графика

plt.xticks(x, labels)  # Заменяем метки на оси X

plt.title('Вероятности вытаскивания шаров разного цвета из урны')

plt.ylabel('Вероятность')

plt.ylim(0, 1)

plt.grid(axis='y', linestyle='--', alpha=0.7)

# Показать значения вероятностей на графике

for i, prob in enumerate(probabilities):

    plt.text(x[i], prob + 0.02, f'{prob:.2f}', ha='center', fontsize=12)

# Показать график

plt.show()

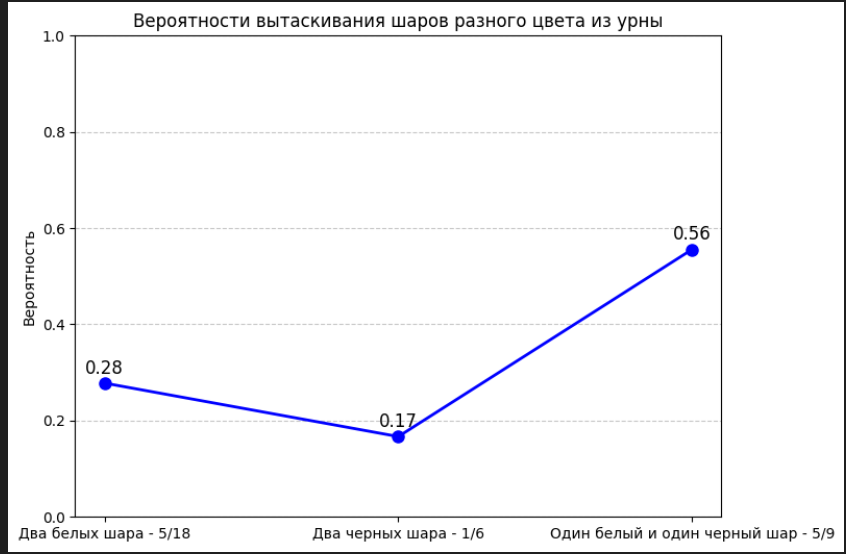


Рисунок 2. Визуализация вычисления комбинаторных вероятностей

Вывод

В результате выполнения работы, я получил практические навыки определения вероятности события различными методами и работе в среде разработки Jupiter Notebook.