

Selected Problems Chapter 1
Introduction to Probability for Data Science
Stanley Chan

Mustaf Ahmed

March 27, 2022

Problem 1.1. (a) Show that

$$\sum_{k=0}^n r^k = \frac{1 - r^{n+1}}{1 - r}$$

for any $0 < r < 1$. Evaluate $\sum_{k=0}^{\infty} r^k$.

(b) Using the result of (a), evaluate

$$1 + 2r + 3r^2 + \dots$$

(c) Evaluate the sums

$$\sum_{k=0}^{\infty} k \left(\frac{1}{3}\right)^{k+1}, \text{ and } \sum_{k=2}^{\infty} k \left(\frac{1}{4}\right)^{k-1}$$

Proof.

Part (a). The first part is true by Theorem 1.1. For the second part, we must evaluate $\sum_{k=0}^{\infty} r^k = \lim_{k \rightarrow \infty} \frac{1 - r^{k+1}}{1 - r}$. Since the limit of the denominator is not zero, we have

$$\begin{aligned} \sum_{k=0}^{\infty} r^k &= \lim_{k \rightarrow \infty} \frac{1 - r^{k+1}}{1 - r} \\ &= \frac{\lim_{k \rightarrow \infty} 1 - r^{k+1}}{\lim_{k \rightarrow \infty} 1 - r} \\ &= \frac{1}{1 - r}. \end{aligned}$$

Part (b). By part (a), we know that $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$. Taking the derivative of both sides gives us

$$1 + 2r + 3r^2 + \dots = \frac{1}{(1-r)^2}.$$

Part (c).

$$\begin{aligned} \sum_{k=0}^{\infty} k \left(\frac{1}{3}\right)^{k+1} &= \sum_{k=1}^{\infty} k \left(\frac{1}{3}\right)^{k+1} \\ &= \left(\frac{1}{3}\right)^2 \sum_{k=1}^{\infty} k \left(\frac{1}{3}\right)^{k-1} \\ &= \frac{1}{4} \end{aligned}$$

$$\begin{aligned} \sum_{k=2}^{\infty} k \left(\frac{1}{4}\right)^{k-1} &= -1 + \sum_{k=1}^{\infty} k \left(\frac{1}{4}\right)^{k-1} \\ &= \frac{7}{9} \end{aligned}$$

□

Problem 1.4. (a)

```
import numpy as np

A = np.array([[1,2,3],[4,5,6],[7,8,9]])
x = np.array([[1],[2],[3]])

b = A.dot(x) # answer [14,32,50]
```

(b)

```
import math
import matplotlib.pyplot as plt

inputs = []
start = -math.pi
change = ((2 * math.pi)/1000)

# add values in domain to inputs
for i in range(1000):
    inputs.append(start)
    start = start + change

outputs = [math.sin(inputs[i]) for i in range(len(inputs))]

plt.plot(inputs, outputs)
plt.show()
```

Problem 1.9. A collection of 26 English letters, a-z, is mixed in a jar. Two letters are drawn at random, one after the other.

(a) What is the probability of drawing a vowel (a,e,i,o,u) and a consonant in either order?

(b) Write a MATLAB / Python program to verify your answer in part (a). Randomly draw two letters without replacement and check whether one is a vowel and the other is a consonant. Compute the probability by repeating the experiment 10000 times.

(a). The size of the sample space is $26 * 25 = 650$. The total ways of getting a vowel first and a consonant second is $5 * 21 = 105$. The total ways of getting a vowel second and a consonant first is $5 * 21 = 105$. Thus, the probability is

$$P(A) = \frac{210}{650} \approx .32$$

(b).

```
import random

count = 0

for i in range(100000):
    num_alphabet = [n for n in range(1,27)]

    # get first alphabet letter
    first = random.choice(num_alphabet)

    # remove first choice
    num_alphabet.remove(first)

    # get second alphabet letter
    second = random.choice(num_alphabet)

    if (first <= 5 and second > 5) or (first > 5 and second <= 5):
        count = count + 1

probability = count/100000

print(probability) # approximately .32
```