Discrete Structures

Worksheet 11.1

Spring 2021
March 17

---

1. **Warm up:** Answer the following questions.

   (a) From a deck of cards, how much more likely is it to choose two spades compared with choosing two queens?

   (b) If there is a 1/3 chance I am wearing blue socks and a 1/5 chance my socks have holes in them, what is the chance that I am wearing blue socks without holes?

   (c) True or false: If events $E_1, \ldots, E_n$ are all mutually independent, then they are pairwise independent.

2. Suppose that in a weekly lottery you have probability .002 of winning a prize with a single ticket. You buy 1 ticket per week for 52 weeks.

   (a) What is the probability that you win no prizes?

   (b) What is the probability that you win 3 or more prizes?

3. Suppose that there are 4 white and 2 black marbles. One way to arrange them is:

   ●○○○●○

   (a) List all 15 ways to arrange these marbles.

   (b) How many ways are there to arrange 4 white and n black marbles?

   (c) In which arrangements from part (a) is every white marble next to at least one other white marble?

   (d) List all the ways to arrange 2 black and 2 red marbles.

   (e) Explain why the lists in parts (c) and (d) have the same size.

   (f) How many ways are there to arrange 4 white and n black marbles so that every white marble is next to at least one other white marble?

   (g) If 21 black and 4 white marbles are arranged at random, what is the probability that every white marble is adjacent to at least one other white marble?

4. A florist stocks red roses and white roses. Of these, some have thorns and some do not. Let R be the event that a rose is red, and let T be the event that a rose has thorns. Suppose that the following data has been collected about the roses: 1/4 of the roses have thorns, 1/5 of the red roses do not have thorns, and 3/7 of roses with thorns are white. Consider the experiment of selecting a rose at random. Use the data to find the probability of each of the four events $R \cap T$, $\overline{R} \cap T$, $R \cap \overline{T}$, and $\overline{R} \cap \overline{T}$.

5. Consider the following algorithm, called the **Fermat primality test**, which takes as input a positive integer n to be tested if it is prime, and a positive integer k.

for $i = 1, 2, \ldots, k$:
  pick randomly $a \in \{1, \ldots, n-1\}$
  $t = a^{n-1} \pmod{n}$
  if $t \neq 1$:
    return “`not prime`”
  return “`probably prime`”

(a) Implement this algorithm in Python and run it for $k = 1000$ on:

i. the composite numbers $1048576 = 2^{20}$ and $28278749 = 7919 \cdot 3571$

ii. the prime numbers 104729 and 1299709

(b) An integer n is a *false positive* for the Fermat primality test if it is not prime and the primality test returns “`probably prime`”. Explain why $341 = 11 \cdot 31$ is a false positive.

(c) An integer n is a *false negative* for the Fermat primality test if it is prime and the primality test returns “`not prime`”. Explain why there do not exist any false negatives.

6. *Bernoulli trials.* Emulate a poll where $n = 100$ people are asked a Yes/No question (“Will you vote in the municipal election on June 5, 2021?” or similar). Assume that all answers are mutually independent; and with a probability $p = 2/3$ the answer is “Yes”.

(a) A single *Bernoulli trial* has outcomes $\{\text{Yes}, \text{No}\}$ (the answer to the sociologist’s question). 100 people answered the same question, so the Bernoulli trials can be enumerated by $t = 0, 1, \ldots, 99$. Define the Bernoulli random variable for any integer $t \in [0; 99]$:

$$X(t) = \begin{cases} 1, & \text{if the person } t \text{ answered "Yes" (probability } p = 2/3), \\ 0, & \text{if the person } t \text{ answered "No" (probability } p = 1/3). \end{cases} \quad (1)$$

Find the expected value $E(X)$ and the variance $V(X) = E\left((X - E(X))^2\right)$ of this random variable.

(b) Use a computer to generate 100 values of this random variable and compute their total. What are the totals that you get most frequently?

```
from scipy.stats import bernoulli
p = 2/3

# Find the expected value ('mean') and variance
mean, var = bernoulli.stats(p, moments='mv')
print('(E(X)={}, V(X)={}'.format(mean, var))

# Generate random responses (1=success, 0=failure)
responses = bernoulli.rvs(p, size=100)
# Count the successful Bernoulli trials
print('sum(responses) = {}'.format(sum(responses)))
```

(c) Imagine that you run this poll 1000 times: Every time a pollster asks 100 people and records the total of "Yes" responses. At the very end the pollster creates a histogram that shows the frequency of every sum.

```
import matplotlib.pyplot as plt
from scipy.stats import bernoulli

p = 2/3
# Initially all poll totals (values 0..100) are set to '0'
allPolls = [0] * 101

for i in range(0,1000):
    responses = bernoulli.rvs(p, size=100)
    allPolls[sum(responses)] += 1

plt.bar(list(range(0,101)), allPolls)
plt.show()
```

(Need to ensure that the body of **for** loop is indented and followed by a blank line.)

(d) Use the linearity property of the expected value to compute $E(Y)$ where the random variable Y is defined as

$$Y = X(0) + X(1) + \ldots + X(99),$$

where every $X(t)$ is a Bernoulli random variable as defined in (1).