Discrete Structures

Worksheet 24

Spring 2022
April 1

1. This questions is about rolling dice. For each, find the probability p of success, and the probability q of failure.

   (a) A single die is rolled. Success is when the number rolled is divisible by two.

   (b) Two dice are rolled 5 times. Sucess is when exactly three of all the 10 numbers rolled is prime.

   (c) Three dice are rolled twice. Success is when the sum of each roll is larger than the previous roll.

2. In a surveyed population, 30% of those aged 50-60 have high blood pressure.

   (a) In a sample of 14 people aged 50-60, what is the probability that more than 6 of them have high blood pressure?

   (b) If the population has uniformly distributed age between 0 and 100, what is the probability that a sample of 40 people are aged 50-60 and have high blood pressure?

   (c) How many people (of the whole population) need to be sampled to guarantee a group of 20 people aged 50-60 with high blood pressure?

3. 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, ..., k:
  pick randomly a ∈ {1, ..., n − 1}
  t = a^(n−1) (mod n)
  if t ≠ 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.

4. A poll of $n = 100$ people asks a Yes/No question, with all anwers mutually independent. The probability that each persone answers "yes" is $p = 2/3$.

   (a) The poll can be enumerated, for $t = 0, 1, \ldots, 99$, by the Bernoulli random variable

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

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

   (b) 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 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 of responses is {}".format(sum(responses)))
```

   (c) Run this poll 1000 times and visualize the results as a histogram.

```
import matplotlib.pyplot as plt
from scipy.stats import bernoulli
p = 2/3

# Initially all poll totals are set to 0
allPolls = [0] * 101

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

# Visualize results
plt.bar(list(range(101)), allPolls)
plt.show()
```

   (d) Consider a related random variable Y, defined as

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

   where every $X(t)$ is a Bernoulli random variable as defined in (1). Use the linearity property of the expected value to compute $E(Y)$.

4. Complete the following tasks for next lab (Tuesday). They will be presented at the beginning of the lab.

(a) You are playing a game where you pick a number between 1 and 100 inclusive. If your number is the secret number, you win 74\$. If it is not, you lose 1\$. Let nX be the random variable representing the amount of money won or lost on n games.

i. Draw the graph of the distribution (probability function) of $1X$ and $2X$.

ii. What is the expected value and variance of $1X$ and $2X$?

(b) A Bernoulli trial with success probability p is repeated until the first success. Let X be the random variable representing the number of trials until the first sucess.

i. Draw the graph of the distribution (probability function) of X from $P(X = 1)$ to $P(X = 10)$, for $p = \frac{1}{10}$ and $p = \frac{9}{10}$.

ii. Find a formula for $P(X = k)$.

iii. What is the expected value and variance of X?

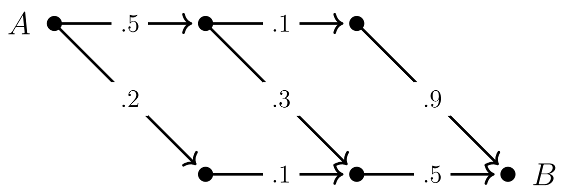
(c) Let `s` be a random binary string of length $n \in \mathbf{N}$, of characters `0` and `1`, with each character chosen uniformly and independently. Let k be an even number less than or equal to n. Let X be the random variable representing the number of length k substrings of `s` that are palindromes.

i. Compute the distribution (probability function) for $n = 3$ with $k = 2$, and for $n = 4$ with $k = 2, 4$.

ii. Compute the expected value of X for the same n, k pairs as above.

iii. Compute the expected value of X for any n, k.

(d) Let $G = (V, E)$ be the connected, directed graph below, with edge failure probabilities given on each edge. Edges fail independently of each other.



Compute the probability that the failed edges disconnect A from B.