MATH 302 — Assignment 4

Stephanie Knill
54882113

Due: February 10, 2016

Question 1: Section 4.1 #32

(a) Let E be event that road 1 is clear and F be the event that road 2 is clear. Thus the probability of each road being passable is $P(E) = p$ and $P(F) = q$. When the system's components are in *series*, the reliability of them can be computed as follows:

$$P(W \to T) = P(E \cap F)$$

Since events E and F are independent, the *intersection* of the events is equal to the probability of each event occurring. Thus

$$\begin{aligned} P(E \cap F) &= P(E) \cdot P(F) \\ &= pq \end{aligned}$$

(b) In the case where the system's components are in *parallel*, the reliability of them is the *union* of the two events:

$$\begin{aligned} P(W \to T) &= P(E \cup F) \\ &= P(E) + P(F) - P(E \cap F) \\ &= p + q - pq \end{aligned}$$

(c) In this graph depicting the driving routes from Woodstock to Tunbridge, we now have 4 possible paths we can take such that we start at Woodstock and end in Tunbridge, and we do not use any edge (road) more than once. These will be represented by E_1, the path $W \to C \to T$, E_2, the path $W \to C \to D \to T$, E_3, the path $W \to D \to T$, and E_4, the path $W \to D \to C \to T$. Let $p = .8, q = .9$, and $r = .95$. Then we have $P(E_1) = P(E_3) = pq = .72$, $P(E_2) = p^2r = 0.608$, and $P(E_4) = q^2r = 0.7695$.

To compute the reliability of this graph, we have

$$\begin{aligned} P(W \to T) &= 1 - P(\text{no passable paths}) \\ &= 1 - P(\overline{E_1} \cap \overline{E_2} \cap \overline{E_3} \cap \overline{E_4}) \end{aligned}$$

Since each road being passable is independent, we have that E_i and E_j are independent for all $i \neq j$, $i, j \in \{1, 2, 3, 4\}$. Thus we can rewrite our equation as:

$$\begin{aligned} P(W \to T) &= 1 - P(\overline{E_1}) \cdot P(\overline{E_2}) \cdot P(\overline{E_3}) \cdot P(\overline{E_4}) \\ &= 1 - (1 - P(E_1)) \cdot (1 - P(E_2)) \cdot (1 - P(E_3)) \cdot (1 - P(E_4)) \\ &= 1 - (1 - .72) \cdot (1 - .0608) \cdot (1 - .72) \cdot (1 - .7695) \\ &\approx .9929 \end{aligned}$$

Question 2: Section 4.1 #43

Let Y be the random variable that gives the number of times the Yankees win a game in a series of length $n = 7$. Since the Yankees can win the world series by either winning 4, 5, 6 or 7 games, the probability of them winning the series is

$$P(Y \geq 4) = P(Y = 4) + P(Y = 5) + P(Y = 6) + P(Y = 7)$$

Computing each separately, we have

$$\begin{aligned} P(Y = 4) &= \binom{n}{k} \cdot p^k q^{n-k} \\ &= \binom{7}{4} (.6)^4 (.4)^3 \\ &\approx 0.2903 \end{aligned}$$

However, due to the greyscale of existence that is a student during midterm season, the remaining computations were done in Python (Figure 1). This gave a final value of $P(Y \geq 4) \approx 0.710208$.

Question 3: Section 4.1 #49

Let H be the event that a coin is tossed and is Heads, with probability $P(H) = p_i$, for urn $i \in \{1, 2\}$. Then in scenario (a) where we randomly select one of the two urns and flip calculator.py*

```python
# -*- coding: utf-8 -*-
"""
Created on Tue Feb  9 22:34:46 2016

@author: Stephanie
"""

from scipy.special import comb
import math
from math import log

# Question 2

n=7
sum = 0
for k in range(4,7+1):
    c = comb(n,k, exact=False)
    Y = c * (.6**k)*(.4**(n-k))
    sum += Y

```

Figure 1: Computation for $P(Y \geq 4)$ in Question 2.

both coins, we have

$$P(HH) = \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2 = \frac{1}{2}(p_1^2 + p_2^2)$$

and in scenario (b) where we flip one coin from each urn

$$P(HH) = p_1 \cdot p_2$$

Since $p_1 \neq p_2$, then

$$\begin{aligned}(p_1 - p_2)^2 &> 0 \\ p_1^2 - 2p_1p_2 + p_2^2 &> 0 \\ p_1^2 + p_2^2 &> 2p_1p_2 \\ \frac{1}{2}(p_1^2 + p_2^2) &> p_1p_2\end{aligned}$$

and we can conclude that the probability in scenario (a) is always greater than the probability in scenario (b) for all values of p_1 and p_2, thereby making scenario (a) the best choice.

Question 4: Section 5.1 #6

Let $X_1, \ldots, X_n$ be n mutually independent random variables uniformly distributed on integers 1 to k. Let Y denote the minimum of the X_i's. Then the distribution of Y can be expressed as

$$P(Y = j) = P(Y > j - 1) - P(Y > j), \text{ where } 1 < j < k$$

Since $min(X_1, \ldots, X_n) > j$ holds true exactly when $X_i > j$ for all $i \in \{1, \ldots, n\}$, we can calculate each component separately by

$$\begin{aligned} P(min(X_1, \ldots X_n) > j) &= P(X_1 > j) \cdot P(X_2 > j) \cdots P(X_n > j) \\ &= \frac{k-j}{k} \cdot \frac{k-j}{k} \cdots \frac{k-j}{k} \\ &= \Big(\frac{k-j}{k}\Big)^n \\ &= \frac{(k-j)^n}{k^n} \end{aligned}$$

and

$$\begin{aligned} P(min(X_1, \ldots X_n) > j - 1) &= P(X_1 > j - 1) \cdot P(X_2 > j - 1) \cdots P(X_n > j - 1) \\ &= \frac{k-(j-1)}{k} \cdot \frac{k-(j-1)}{k} \cdots \frac{k-(j-1)}{k} \\ &= \Big(\frac{k-(j-1)}{k}\Big)^n \\ &= \frac{(k-(j-1))^n}{k^n} \end{aligned}$$

Plugging this back into our original equation we have

$$P(Y = j) = \frac{(k-(j-1))^n}{k^n} - \frac{(k-j)^n}{k^n}$$

Question 5: Section 4.1 #15

Conjecture: For a geometric random variable X of parameter p and for non-negative integers m and n, we have that

$$P(X > n + m \mid X > m) = P(X > n)$$

Proof

$$\begin{aligned}
P(X > n+m \mid X > m) &= \frac{P((X > n+m) \cap (X > m))}{P(X > m)} \\
&= \frac{P(X > n+m)}{P(X > m)} \\
&= \frac{(1-p)^{n+m}}{(1-p)^m} \\
&= (1-p)^n \\
&= P(X > n)
\end{aligned}$$

■

Question 6: Section 5.1 #14

Let X be the number of people with a particular rare blood type, with parameter $p = 1/1000$.

(a) For $n = 10,000$, the probability of no person having this blood type is

$$\begin{aligned}
P(X = 0) &= \binom{n}{k} p^k q^{n-k} \\
&= \binom{10000}{0} (1/1000)^0 \cdot (999/1000)^{10000} \\
&= 4.517 \times 10^{-5}
\end{aligned}$$

(b) To find an n such that the probability of at least one person has a rare blood type in a population is greater than $1/2$ can be computed as follows:

$$\begin{aligned}
P(X \geq 1) &= 1 - P(X = 0) \\
&= 1 - \binom{n}{0} p^0 q^n \\
&= 1 - q^n
\end{aligned}$$

Since $P(X \geq 1) > 1/2$

$$\begin{aligned} 1 - q^n &> 1/2 \\ q^n &< 1/2 \\ \ln(q^n) &> \ln(1/2) \\ n &> \frac{\ln(1/2)}{\ln(q)} \\ n &> 692.8 \\ n &= 693 \end{aligned}$$

Question 7: Section 5.1 #35

Let S be the number of brass turnbuckles manufactured and D the number of those that are defective. A sample of s items is drawn without replacement. Let X be a random variable that gives the number of defective items in the sample. Let $p(d) = P(X = d)$.

(a) **Conjecture:**

$$p(d) = \frac{\binom{D}{d}\binom{S-D}{s-d}}{\binom{S}{s}}$$

Proof The probability that in a given sample s that exactly d items are defective can be expressed as

$$P(X = d) = \frac{\text{\# ways sample has } d \text{ defective}}{\text{total \# ways draw sample } s}$$

Computing the numerator, we have that the total number of ways to choose a sample s from a population of S items is $\binom{S}{s}$.

For the denominator, we must choose d defective items from a total of D items; in other words, we have $\binom{D}{d}$ possible combinations. Of the remaining $S - D$ items in our set that are not defective, we must choose a sample of $s - d$ items. This gives us $\binom{S-D}{s-d}$. Thus the denominator is given by $\binom{D}{d}\binom{S-D}{s-d}$.

Combining them all we have

$$P(X = d) = p(d) = \frac{\binom{D}{d}\binom{S-D}{s-d}}{\binom{S}{s}}$$

$\blacksquare$

Question 8

Proposition: Let X be a binomial random variable of parameters n, p. Then the value of k that maximizes $P(X = k)$ is

$$\lfloor (n+1)p \rfloor$$

Proof

$$\begin{aligned}\frac{P(X=k)}{P(X=k-1)} &= \frac{\binom{n}{k} \cdot p^k (1-p)^{n-k}}{\binom{n}{k-1} \cdot p^{k-1} (1-p)^{n-(k-1)}} \\ &= \frac{\binom{n}{k} \cdot p}{\binom{n}{k-1}(1-p)}\end{aligned}$$

Here

$$\begin{aligned}\frac{\binom{n}{k}}{\binom{n}{k-1}} &= \frac{n!}{k!(n-k)!} \cdot \frac{(k-1)!(n-k+1)!}{n!} \\ &= \frac{n+1}{k} - 1\end{aligned}$$

Substituting this back in

$$\frac{P(X=k)}{P(X=k-1)} = \left(\frac{n+1}{k} - 1\right) \cdot \frac{p}{(1-p)}$$

Since we want to find a k such that $\frac{P(X=k)}{P(X=k-1)} \geq 1$

$$\begin{aligned}\frac{n+1}{k} - 1 &\geq \frac{(1-p)}{p} \\ \frac{n+1}{k} &\geq \frac{1}{p} \\ k &\leq (n+1)p\end{aligned}$$

Since $k \in \mathbb{Z}$ and we wish to maximize $P(X = k)$, we can conclude that

$$k = \lfloor (n+1)p \rfloor$$

■