Module 1 - Homework 1

Colt Thomas

March 6, 2021

Problem 1

Show that

(a) *If $P(A) = P(B) = P(A \cap B)$, then $P((A \cap B^c) \cup (B \cap A^c)) = 0$.*

Solve by substitution of A and B with $A \cap B$ in such a way we can associate and commute A^c and B^c into the substitution:

$P((A \cap B^c) \cup (B \cap A^c)) =$
$P(((A \cap B) \cap B^c) \cup ((A \cap B) \cap A^c)) =$
$P((A \cap (B \cap B^c)) \cup ((A \cap A^c) \cap B)$

The intersection of an event with its complement is 0. We can reduce the statement to:

$P((A \cap (0)) \cup ((0) \cap B) =$

An event intersected with the empty set is also 0

$P((0) \cup (0)) =$
$P(0) = 0$

(b) *If $P(A) = P(B) = 1$, then $P(A \cap B) = 1$.*

The probability of the universal set $P(S) = 1 = P(A) = P(B)$. We can substitute this in:

$P(A \cap B) = P(S \cap S) =$

The probability of anything intersected with itself is just itself, e.g. AA = A (see page 17 in the textbook)

$P(S \cap S) = P(S) = 1$.

Note: we could have simply substituted A for B, or vise versa instead of bringing in the universal set. I made this more convoluted than necessary.

Problem 2

(Boole's inequality or Union bound) Show that

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

Let n, i be integers such that $i = 1 \leq n$, and subset $A = [A_1, A_2, ...A_n - 1, A_n]$ be events such that $A \in S$

Consider the following three assumptions for boundary cases:

1. Let $n = 1$, and there is one element of subset A, A_1

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P(A_1), n = 1$$

$$\sum_{i=1}^{n} P(A_i) = P(A_1), n = 1$$

   Under these conditions, Boole's inequality holds true $P(A) \leq P(A)$

2. Let $n = k$, and there are k elements of subset A, such that $A = [A_1, A_2, ...A_k - 1, A_k]$. Assume all elements of subset A are mutually exclusive.

$$P\left(\bigcup_{i=1}^{n} A_i\right)$$

$$P(A_1 \cup A_2 \cup ... \cup A_{n-1} \cup A_n)$$

   By definition of mutual exclusion, the intersection of all elements of A are $\emptyset$, therefore we can summarize the expression as:

$$P(A_1) + P(A_2) + ... + P(A_{n-1}) + P(A_n) = \sum_{i=1}^{n} P(A_i)$$

3. Let $n = k$, and there are k elements of subset A, such that $A = [A_1, A_2, ...A_k - 1, A_k]$. Assume two elements of subset A, B_1 and B_2 are not mutually exclusive from each other.

   A union of two events/subsets that are not mutually exclusive is defined as:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B)$$

We can substitute our non-mutually exclusive pair of events B_1, B_2 in as follows:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

$$\Big(P(A_1 \cup A_2 \cup ... \cup A_{m-1} \cup A_m) - P(B_1 \cup B_2)\Big) + P(B_1 \cup B_2) \leq \sum_{i=1}^{n} P(A_i)$$

Where $m = n-2$ to represent the reduced subset A without B_1, B_2. Notice that $P(\cup_{i=1}^{m} A_i)$ has $P(B_1), P(B_2)$ pulled out of the union since they are a part of A. Since the remaining elements of A are mutually exclusive, and by what was proven in part two, we can express the union of the A events in term of a summation:

$$\Big(P(A_1)+P(A_2)+...P(A_{m-1})+P(A_m))-P(B_1 \cup B_2)\Big)+P(B_1 \cup B_2) \leq \sum_{i=1}^{n} P(A_i)$$

$$\left(\sum_{i=1}^{m} P(A_i) - P(B_1 \cup B_2)\right) + P(B_1 \cup B_2) \leq \sum_{i=1}^{n} P(A_i)$$

Now we pull out $P(B_1) and P(B_2)$ out of the right summation

$$\left(\sum_{i=1}^{m} P(A_i) - P(B_1 \cup B_2)\right) + P(B_1 \cup B_2) \leq \sum_{i=1}^{m} P(A_i) + P(B_1) + P(B_2)$$

$$\left(\sum_{i=1}^{m} P(A_i) - P(B_1) - P(B_2) + P(B_1 B_2)\right) + P(B_1) + P(B_2) - P(B_1 B_2) \leq \sum_{i=1}^{m} P(A_i) + P(B_1) + P(B_2)$$

$$\sum_{i=1}^{m} P(A_i) < \sum_{i=1}^{m} P(A_i) + P(B_1) + P(B_2)$$

If there were more elements of A, B_j that weren't mutually exclusive we would continue to see the same pattern such that:

$$P\left(\bigcup_{i=1}^{n} A_i\right) < \sum_{i=1}^{m} P(A_i) + \sum_{j=1}^{k} P(B_j)$$

Where k is the number of non-mutually exclusive elements of A, and n = m+k

Problem 3

(Chain Rule) Show that

$$P\left(\bigcap_{i=n}^{1} A_i\right) = P\left(A_n \middle| \bigcap_{i=n}^{1} A_i\right) ... P(A_2|A_1)P(A_1)$$

The definition of a conditional probability is $P(B|A) = \frac{P(AB)}{P(A)}$. Using this definition we can "unravel" the equation on the right hand side:

$$P\left(A_n \middle| \bigcap_{i=n}^{1} A_i\right) ... P(A_2|A_1)P(A_1) =$$

$$\left(\frac{P(\bigcap_{i=n}^{1} A_i)}{P(\bigcap_{i=n-1}^{1} A_i)}\right) ... \left(\frac{P(A_3A_2A_1)}{P(A_2A_1)}\right)\left(\frac{P(A_2A_1)}{P(A_1)}\right)P(A_1)$$

$$= P\left(\bigcap_{i=n}^{1} A_i\right)$$

Problem 4

An urn contains r red balls and g green balls. One of the balls is drawn at random, but when it is put back in the urn b additional balls of the same color are put in with it. Now suppose that we draw another ball. Show that the probability that the first ball drawn was red given that the second ball drawn was green is $\frac{r}{r+g+b}$

$$P(R|G) = \frac{P(RG)}{P(G)}$$

Before going further, note that order does matter. Let A be the first draw event, and B be the second draw event. P(A) and P(B) will be the events that we draw red for the first and second rounds respectively and P($\overline{A}$), P($\overline{B}$) will be the probability of drawing a green on the first and second rounds likewise. We need to prove:

$$P(A|\overline{B}) = \frac{r}{r+g+b}$$

By definition of conditional probability:

$$P(A|\overline{B}) = \frac{P(A\overline{B})}{P(\overline{B})}$$

For this problem it is usually helpful to draw out a tree and list all the probable outcomes. For timesake, they are $P(A \cap B), P(A \cap \overline{B}), P(\overline{A} \cap B), P(\overline{A} \cap \overline{B})$. Note that $P(B) = P(A \cap B) + P(\overline{A} \cap B)$, and $P(A\overline{B}) = P(A \cap \overline{B})$ Let's substitute this in:

$$P(A|\overline{B}) = \frac{P(A \cap \overline{B})}{P(A \cap \overline{B}) + P(\overline{A} \cap \overline{B})}$$

Now substitute in $P(A), P(B), etc.$ for probabilistic weights:

$$P(A|\overline{B}) = \frac{\frac{r}{r+b}\frac{g}{r+g+b}}{\frac{g}{r+b}\frac{r}{r+g+b} + \frac{g}{r+b}\frac{g+b}{r+g+b}}$$

Note that $P(\overline{A} \cap \overline{B}) = \frac{g+b}{r+g+b}$, since there are b extra red balls due to the success in the first round.

$$P(A|\overline{B}) = \frac{\frac{r}{r+b}\frac{g}{r+g+b}}{\frac{g}{r+b}\frac{r}{r+g+b} + \frac{g}{r+b}\frac{g+b}{r+g+b}}$$

$$P(A|\overline{B}) = \frac{\frac{1}{r+b}\frac{1}{r+g+b}rg}{\frac{1}{r+b}\frac{1}{r+g+b}\big(gr + g(g+b)\big)}$$

$$P(A|\overline{B}) = \frac{rg}{g\big(r+g+b\big)}$$

And we therefore get:

$$P(A|\overline{B}) = \frac{r}{\big(r+g+b\big)}$$

Problem 5

A pair of fair dice is rolled 7 times. Find the probability that "seven" will show at least three times. $\boxed{0.096}$

In this problem, order does not matter since each dice roll is independent of each other. Let event A be the event in which we roll a seven, such that A = (1,6),(2,5),(3,4),(4,3),(5,2),(6,1). We use ordered pairs to represent each die.

First, let's find P(A). The total combinations of dice rolls is m^j where m is the number of sides, and j is the number of dice, or $6^2 = 36$. Since A has 6 pairs, or outcomes, $P(A) = 6/36 = 1/6$.

Now let's set up the experiment. We want to find P(A$\geq$ 3) for trials $n = 7$. Note that order does not matter, and that we need at least 3 successful outcomes and not exactly 3 outcomes. We can use the Fundamental Theorem (eq 3-13 in the book pg 53) :

$$p_n(k) = P(A \text{ occurs k times in any order}) = \binom{n}{k} p^k (p-1)^{n-k}$$

The tricky part is that this gives $P(A = k)$ and not $P(A \leq k)$.We would need to find the sum of all $P(A = k)$ for $3 \leq k \leq 7$. I will do this using python, since writing and calculating this summation would be lengthy.

I wrote a python function called fundCalc() that is included in the code below. This will implement the Fundamental Theorem equation stated below. For our input parameters we have n (number of favorable outcomes), k (total number of trials), and p (probability of the favorable event). P($\overline{A}$) = 1-P(A) = 5/6, so we will use this for parameter p.

```
In [24]: fundCalc(7,3,1/6)+fundCalc(7,4,1/6)+fundCalc(7,5,1/6)+fundCalc(7,6,1/6)+fundCalc(7,7,1/6)
Out[24]: 0.09577546296296295

In [25]: fundCalc(7,0,1/6)+fundCalc(7,1,1/6)+fundCalc(7,2,1/6)
Out[25]: 0.9042245370370372

In [26]: fundCalc(7,0,1/6)+fundCalc(7,1,1/6)+fundCalc(7,2,1/6)+fundCalc(7,3,1/6)+fundCalc(7,4,1/6)+fundCalc(7,5,1/6)
+fundCalc(7,6,1/6)+fundCalc(7,7,1/6)
Out[26]: 1.0000000000000002

In [27]:
```

Figure 1: Calculation of $p_7(3) = P(A \geq 3)$. Note that the sum of $p_7(k) = 1$ for all k is 1 on line 26.

Line 24 in figure 1 shows that $p_7(3) + ... + p_7(7) \approx \boxed{0.096}$. Since P(A)=1/6, we would expect that in a series of 7 dice rolls we are most likely to see one of those rolls be favorable. See Figure 2 below for evidence.

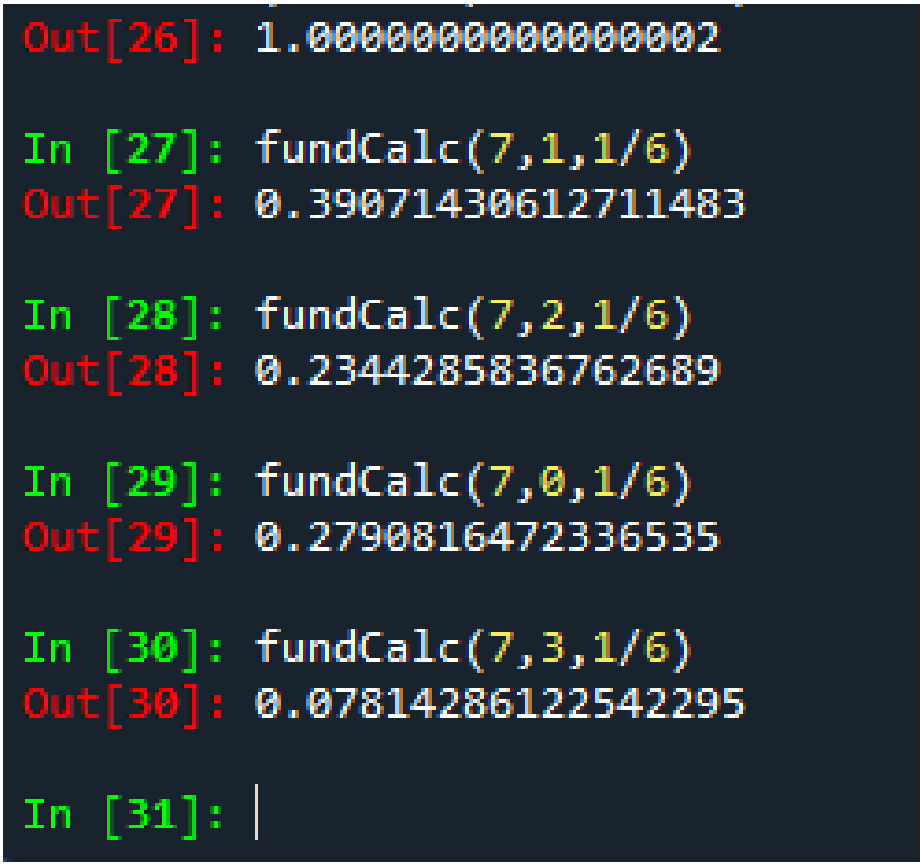

Figure 2: Visual confirmation that $p_7(1)$ has the highest probability, since $P(A) = 1/6$. We are most likely to roll a 7 once out of seven trials.

Problem 6

An urn contains n balls numbered from 1 to n. The balls are drawn one at a time without being replaced in the urn. What is the probability that in the first k draws the numbers on the balls will coincide with the numbers of the draws?

Let A_k be the event where we draw a ball numbered k on the kth draw. This problem is asking to find:

$$P\left(\bigcap_{i=1}^{k} A_i\right)$$

Note that we are looking for $P(A_1 \cap A_2 \cap ... \cap A_k)$ and not for probabilities of something like $P(\overline{A_1} \cap A_2 \cap ... \cap A_k)$ where we missed a k coinciding with the ball drawn. Let's see if we can find a pattern:

$$P(A_1) = 1/n, k = 1$$
$$P(A_1 \cap A_2) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right), k = 2$$
$$P(A_1 \cap A_2 \cap ... \cap A_k) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)...\left(\frac{1}{n-k}\right)$$

We have a pattern established and can represent it as:

$$P(A_1 \cap A_2 \cap ... \cap A_k) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)...\left(\frac{1}{n-k}\right)$$
$$P(A_1 \cap A_2 \cap ... \cap A_k) = \boxed{\frac{(k-1)!}{n!}}$$

Problem 7

Simulate a coin tossing experiment, where p represents the probability of obtaining heads, by modifying the code:

Listing 1: coinToss.m

```
% Set up a simulation for n coin tosses
x = rand(n,1);
number = sum(x < p);
```

Using n repetitions of the experiment, count the number of times m heads occur. Note that

$$P(m \text{ heads}) = \frac{\text{Number of times m heads observed}}{n}$$

1. For a fair coin, what is the simulated probability of obtaining m= 3 heads in n= 10 coin tosses? Compare your results to that obtained with the theoretical formula

$$P(m) = \binom{n}{m} p^m (1-p)^{n-m}$$

Let's calculate the theoretical value. Notice that $p = 0.5$ is the probability of landing heads:

$$P(3) = \binom{10}{3} (0.5)^3 (1-(0.5))^{10-3}$$

$$P(3) = \frac{10!}{3!(10-3)!} \frac{1}{1024} = \boxed{0.117188}$$

This is roughly 1 times out of 10 experiments (where we toss 10 times per experiment), so let's see how many experiments it takes to get exactly 3 heads in 10 coin tosses in figure 3 below:

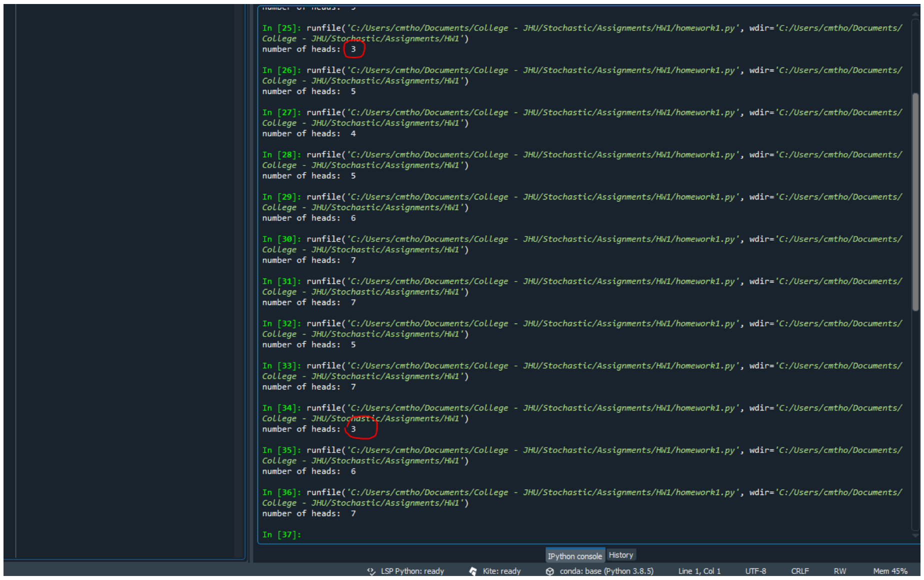

Figure 3: Sequential execution of experiment, or simulated coin tosses. The experiments where heads was tossed 3 times was circled in red.

2. Repeat part (a) but instead consider a biased coin with p= 0.3. Make a plot of P(m heads) versus m heads using the theoretical formula.

   For comparison, let's calculate $P(m), m = 3$:

$$P(3) = \binom{10}{3}(0.3)^3(1-(0.3))^{10-3}$$

$$P(3) = (120)(2.22357 * 10^{-3}) = \boxed{0.266}$$

   Let's plot $P(m)$ in Python in figure 4. Notice that the plot aligns with our calculation above, and we are most likely to have 3 heads in each experiment.

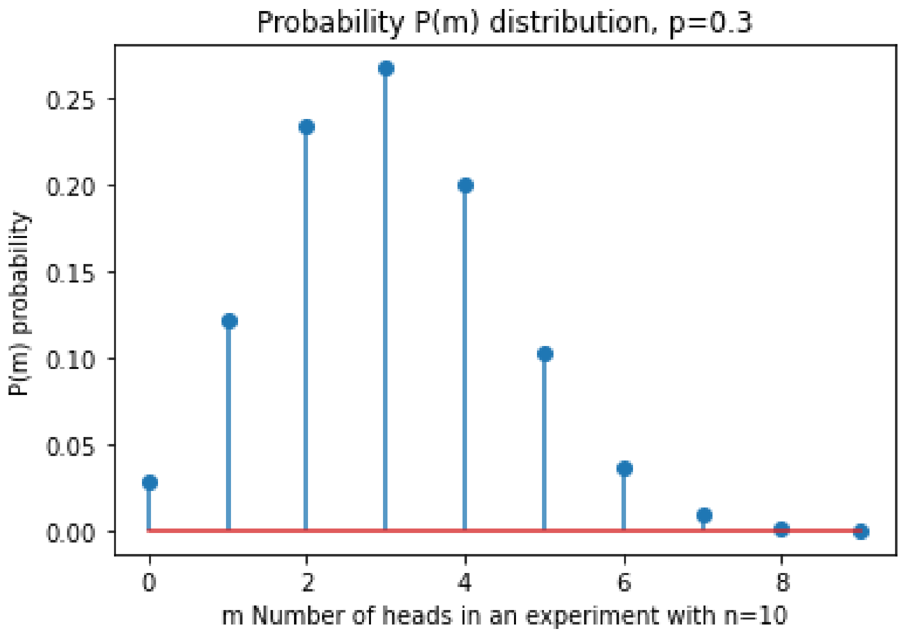

Figure 4: A plot of our biased coin toss experiment, with n=10 coin tosses.

Python Code

```
import numpy as np
import math
import matplotlib.pyplot as plt
##############################
#Problem 5
##############################

# calculates the fundamental theorem, or the probability of p occuring k out of n times.
def fundCalc(n,k,p):
    return math.comb(n,k)*(p**k)*(1-p)**(n-k)

##############################
#Problem 7
##############################

# Set up a simulation for n coin tosses
n = 10    # number of experiments
p = 0.5    # probability of obtaining heads

# Create a repetition of experiments by using the numpy rand()
# function. This produces an n by 1 matrix with values between
# 0 and 1 (uniform distribution, like the coin toss)
x = np.random.rand(n,1)

# Calculate the number of m heads that occur
m = sum(x<p)
print("number of heads: ",m[0])


# plot of the probability P(m) with a biased coin of p=0.3
p = 0.3
m = np.arange(0,10)
Pm = np.zeros(10)
for idx in range(0,10):
   # Pm[idx] = math.comb(n,m[idx])*p**(m[idx])*(1-p)**(n-m[idx])
   Pm[idx] = math.comb(n,m[idx])*p**(m[idx])*(1-p)**(n-m[idx])
   print(math.comb(n,m[idx])*p**(m[idx])*(1-p)**(n-m[idx
```

```
        ]))
# Pm = math.comb(n,m)*p**(m)*(1-p)**(n-m)
print(Pm)
# create the plot
plt.stem(m,Pm)
plt.title('Probability P(m) distribution, p=0.3')
plt.xlabel('m Number of heads in an experiment with n=10'
    )
plt.ylabel('P(m) probability')
plt.show()
```