Assignment 3 Solutions

COMP2804 Winter 2020

March 24, 2020

Name: Lenny Learning Combinatorics
Student ID: 100000000

1 Probabilities of Scrabble Words

A scrabble hand is a set of 7 tiles, each having one of the english uppercase letters on them, drawn uniformly at random from a bag of 100 tiles. The number of tiles of each letter are as follows:

E×12, A×9, I×9, O×8, N×6, R×6, T×6, L×4, S×4, U×4, D×4, G×3, B×2, C×2, M×2, P×2, F×2, H×2, V×2, W×2, Y×2, K×1, J×1, X×1, Q×1, Z×1

Imagine that the one hundred Scrabble tiles are all unique so that, for the example the tiles with a A on them are called $A_1, \ldots, A_9$, the E's are $E_1, \ldots, E_{12}$, the P's are P_1, P_2, and so on.

Once tiles are distinguished this way, there are exactly $\binom{100}{7}$ Scrabble hands, and each one is equally likely.

1. What is the probability that a Scrabble hand contains the word HEXAGON?

   There 2 ways to choose the H, 12 ways to choose the E, 1 way to choose the X, 9 ways to choose the A, 3 ways to choose the G, 8 ways to choose the O, and 6 ways to choose the N. Therefore, the probability of a Scrabble hand containing the word HEXAGON is

$$\frac{2 \cdot 12 \cdot 1 \cdot 9 \cdot 3 \cdot 8 \cdot 6}{\binom{100}{7}} = \frac{324}{166745425} \approx 1.943 \times 10^{-6} \ .$$

```python
from math import factorial
from fractions import Fraction
binom = lambda n,k: factorial(n)//(factorial(k)*factorial(n-k))
print(Fraction(2*12*1*9*3*8*6, binom(100,7)))
# 324/166745425
```

2. What is the probability that a scrabble hand contains the word GARBAGE?

   We use the same analysis as before, but we have some duplicate letters: There are $\binom{3}{2}$ ways to choose 2 G's, $\binom{9}{2}$ ways to choose 2 A's, 6 ways to choose an R, 2 ways to choose a B, and 12 ways to choose an E. Therefore, the probability that a Scrabble hand contains the word GARBAGE is

$$\frac{\binom{3}{2} \cdot 2 \cdot \binom{9}{2} \cdot 6 \cdot 2 \cdot 12}{\binom{100}{7}} = \frac{162}{166745425} \approx 9.715 \times 10^{-7} \ .$$

```
print(Fraction(binom(3,2)*binom(9,2)*6*2*12, binom(100,7)))
# 162/166745425
```

3. What is the probability that a scrabble hand contains the word APPLE?

   This one is trickier because APPLE has only five letters and our Scrabble hand has 7, so we have to take care to avoid double-counting. Still, there is only $\binom{2}{2} = 1$ ways to get 2 P's. After that we need to make a 5-letter hand that contains at least one each of A, L, and E using the remaining 98 tiles. Let's break it down:

   (a) We could have a hand that contains $\{A, L, E, x, y\}$ where $x, y \notin \{A, L, E, P\}$. There are

   $$9 \cdot 4 \cdot 12 \cdot \binom{100 - 2 - 9 - 4 - 12}{2} = 9 \cdot 4 \cdot 12 \cdot \binom{73}{2} \tag{1}$$

   such hands.

   (b) We could have a hand that contains $\{A, L, E, x, y\}$ where $x \in \{A, L, E\}$ and $y \notin \{A, L, E\}$. There are

   $$\left(\binom{9}{2} \cdot 4 \cdot 12 \cdot 73\right) + \left(9 \cdot \binom{4}{2} \cdot 12 \cdot 73\right) + \left(9 \cdot 4 \cdot \binom{12}{2} \cdot 73\right) \tag{2}$$

   such hands. (The three terms correspond to 2 A's, 2 L's, or 2 E's.)

   (c) We could have a hand that contains one of A, L, or E, three times, i.e, $\{A, L, E, x, x\}$ for some $x \in \{A, L, E\}$. There are

   $$\left(\binom{9}{3} \cdot 4 \cdot 12\right) + \left(9 \cdot \binom{4}{3} \cdot 12\right) + \left(9 \cdot 4 \cdot \binom{12}{3}\right) \tag{3}$$

   such hands. (The three terms correspond to 3 A's, 3 L's, or 3 E's.)

   (d) We could have a hand that contains $\{A, L, E, x, y\}$ where $x, y \in \{A, L, E\}$ and $x \neq y$. There are

   $$\left(\binom{9}{2} \cdot \binom{4}{2} \cdot 12\right) + \left(9 \cdot \binom{4}{2} \cdot \binom{12}{2}\right) + \left(\binom{9}{2} \cdot 4 \cdot \binom{12}{2}\right) \tag{4}$$

   such hands. (The three terms correspond to 2 A's and L's, 2 L's and E's, and 2 A's and E's.)

   Therefore, the total number of Scrabble hands that contains APPLE is

   $$(1) + (2) + (3) + (4) = 1510236$$

   So the probability that a Scrabble hand contains APPLE is

   $$\frac{1510236}{\binom{100}{7}} = \frac{17979}{190566200} \approx 9.4345 \times 10^{-5}$$

```
s1 = 9*4*12*binom(73, 2)
s2 = binom(9,2)*4*12*73 + 9*binom(4,2)*12*73 + 9*4*binom(12,2)*73
s3 = binom(9,3)*4*12 + 9*binom(4,3)*12 + 9*4*binom(12,3)
s4 = binom(9,2)*binom(4,2)*12 + 9*binom(4,2)*binom(12,2) + binom(9,2)*4*binom(12,2)
print(s1+s2+s3+s4)
print(Fraction(s1+s2+s3+s4, binom(100,7)))
# 1510236
# 17979/190566200
```

2 Feeding Your Rat

A rat feeder is essentially a straw whose diameter is just large enough for 1 (medicine) pill or 1 (food) pellet, but is long enough to hold many pills and pellets. The pill and pellets are put in at one of the feeder and come out the other end (when the rat presses a pedal) in the same order they were put in.

Suppose we place 25 identical pellets and 4 identical pills uniformly at random into a rat feeder. The rat then comes and consumes one item x_1 from the feeder and then consumes another item x_2 from the feeder.

Let A be the event "x_1 is a pellet" and let B be the event "x_2 is a pill".

1. What is $\Pr(A \cap B)$?

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B \mid A) = \frac{25}{29} \cdot \frac{4}{28} = \frac{25}{203}$$

2. What is $\Pr(A \cup B)$?

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = \frac{25}{29} + \frac{4}{29} - \frac{25}{203} = \frac{178}{203}$$

```
print(Fraction(25,29) + Fraction(4,29) - Fraction(25,29)*Fraction(4,28))
# 178/203
```

3. Are the events A and B independent? In other words, is $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$?

No they are not independent, because

$$\Pr(A) \cdot \Pr(B) = \frac{25}{29} \cdot \frac{4}{29} = \frac{100}{841} \neq \frac{25}{203} = \Pr(A \cap B)$$

3 A Coin-Flipping Game

Consider the following coin tossing games. For each one, compute the probability that you win the game. For each question, the answer is a rational number so you should give this number exactly and give a decimal approximation of it as well.

1. You toss a fair coin twice and win if it comes up heads at least once.

This one is easy to do by hand. This is a uniform probability space over $S = \{TT, TH, HT, HH\}$ and the event $W =$ "you win this game" $= \{HH, TH, HT\}$. So, $\Pr(W) = |W|/|S| = 3/4 = 0.75$.

2. You toss a fair coin 10 times and win if comes up heads at least five times.

Let $A =$ "you win this game". You could do this the long way, by counting how many of the 2^{10} 10-flip sequences contains n heads, for each $n \in \{5, 6, 7, 8, 9, 10\}$, which leads to the answer:

$$\Pr(A) = \frac{\sum_{k=5}^{10} \binom{10}{k}}{2^{10}} = \frac{319}{512} \approx 0.623046875 \ .$$

Python can calculate this easily:

```
print(Fraction(sum([binom(10,k) for k in range(5,11)]), 2**10))
# 319/512
```

Here's another way: For each $k \in \{0, \ldots, 10\}$, there's an obvious bijection between the flip sequences with exactly k heads and flip sequences with exactly k tails. So, for each $i \in \{0, \ldots, 4\}$, let $A_i =$ "You win by getting exactly $10 - i$ heads" and $B_i =$ "You lose by getting exactly $10 - i$ tails". Then $\Pr(A_i) = \Pr(B_i)$ for each $i \in \{0, \ldots, 4\}$. Therefore

$$1 = \sum_{i=0}^{4} (\Pr(A_i) + \Pr(B_i)) + \Pr(A_5) = 2 \cdot \left(\sum_{i=0}^{4} \Pr(A_i) \right) + \Pr(A_5)$$

so

$$\sum_{i=0}^{4} \Pr(A_i) = \frac{1 - \Pr(A_5)}{2} \ .$$

Therefore,

$$\Pr(A) = \sum_{i=0}^{4} \Pr(A_i) + \Pr(A_5) = \frac{1}{2} + \frac{\Pr(A_5)}{2} = \frac{1}{2} + \frac{\binom{10}{5}}{2^{10}} = \frac{319}{512} \approx 0.623046875 \ .$$

```
print(Fraction(1,2)+Fraction(1,2)*Fraction(binom(10,5),2**10))
# 319/512
```

3. You toss a fair coin twice and win if it comes up heads exactly once.

   This is another one we can enumerate: $S = \{TT, TH, HT, HH\}$, and $W = \{TH, HT\}$, so $\Pr(W) = |W|/|S| = 1/2 = 0.5$.

4. You toss a fair coin 10 times and win if comes up heads exactly five times. There are exactly $\binom{10}{5}$ flip sequences containing exactly 5 heads, so the probability of winning this game is only

$$\frac{\binom{10}{5}}{2^{10}} = \frac{63}{256} \approx 0.24609375$$

   Python code:

```
print(Fraction(binom(10,5),2**10))
# 63/256
```

4 Blindfolded Musical Chairs

We are playing a game of blindfolded musical chairs with 20 blindfolded people and 40 chairs. When the music stops each person picks a chair uniformly at random and sits on it.

1. What is the probability that some chair has at least two people sitting on it?

   Let A = "some chair has at least two people sitting on it". Then $\overline{A}$ = "each chair has at most one person on it".

   Let S be the set of functions that map people (a set of size 20) to chairs (a set of size 40), so $|S| = 40^{20}$. The event $\overline{A}$ is the set of one-to-one functions in S. In class, we've seen that there are $|\overline{A}| = 20!\binom{40}{20}$. Therefore,

$$\Pr(\overline{A}) = \frac{20!\binom{40}{20}}{40^{20}}$$

   so

$$\Pr(A) = 1 - \Pr(\overline{A}) = \frac{40^{20} - 20!\binom{40}{20}}{40^{20}} = \frac{33452086084232679113339}{335544320000000000000000}$$

```
print(Fraction(40**20-factorial(20)*binom(40,20), 40**20))
```

2. What is the probability that some chair has at least three people sitting on it?

   Again, we should use the complement rule. Let $\overline{A}$ = "every chair has at most two people on it". For each $i \in \{0, \ldots, 10\}$, let $\overline{A}_i$ = "i chairs have two people on them and $20 - 2i$ chairs have one person on them".

   Then we can count $|\overline{A}_i|$ using the Product Rule:

   (a) Choose $2i$ people $Q_1, \ldots, Q_{2i}$.

   (b) Pair $Q_1, \ldots, Q_{2i}$ in groups $G_1, \ldots, G_i$ with two people in each group. This leaves $20 - 2i$ lonely people $P_{i+1}, \ldots, P_{20-i}$.

(c) Choose a one-to-one function $f : \{G_1, \ldots, G_i, P_{i+1}, \ldots, P_{20-i}\} \to \{1,\ldots,40\}$ that assigns people to chairs.

The number of ways of doing the first step is $\binom{20}{2i}$. The number ways of doing the last step is $(20-i)!\binom{40}{20-i}$. The number of ways doing the second step is a problem we've seen already in Assignment 2, Question 3.2, and its $\binom{2i}{i}i!/2^i$. Therefore

$$\Pr(\overline{A}) = \sum_{i=0}^{10} \Pr(\overline{A}_i) = \sum_{i=0}^{10} \frac{\binom{20}{2i} \cdot \binom{2i}{i} i! \cdot (20-i)!\binom{40}{20-i}}{2^i 40^{20}} = \frac{2418616818652159169703558 9}{4294967296000000000000000 0}$$

and

$$\Pr(A) = 1 - \Pr(\overline{A}) = \frac{18763504773478408302964411}{42949672960000000000000000} \approx 0.4368718893611433 \ .$$

Python code:

```
f = lambda i: Fraction(binom(20,2*i)
                       * binom(2*i,i)*factorial(i)*factorial(20-i)
                       * binom(40,20-i),
                       2**i * 40**20)
abar = sum([f(i) for i in range(11)])
a = 1-abar
print(a)
print(float(a))
```

5 Random Number Weirdness

We independently pick two random numbers R_1 and R_2 from the set $\{1, \ldots, 1000\}$. (Note: Independence means we may pick $R_1 = R_2$.)

Here we are working in the uniform probability space over the sample space $S = \{1, \ldots, 1000\}^2$ which has size 1000^2.

1. What is the probability that R_1 and R_2 are both even?

   Here the event $A =$ "R_1 and R_2 are both even" is

   $$A = \{2k : i \in \{1, \ldots, 500\}^2$$

   and has size 500^2, so

   $$\Pr(A) = \frac{|A|}{|S|} = \frac{500^2}{1000^2} = \frac{1}{4} \ .$$

2. Suppose I tell you that at least one of R_1 or R_2 is even. What is the (conditional) probability that R_1 and R_2 are both even?

   Here the event $B =$ "at least one of R_1 or R_2 is even" is

   $$B = \{(2k, R_2) : k \in \{1, \ldots, 500\},\ R_2 \in \{1, \ldots, 1000\}\} \cup \{(R_1, k) : k \in \{1, \ldots, 500\},\ R_1 \in \{1, \ldots, 1000\}\}$$

   and using the principle of inclusion-exclusion, we determine the size of B:

   $$|B| = 500 \cdot 1000 + 500 \cdot 1000 - |A| = 500 \cdot 1000 + 500 \cdot 1000 - 500^2 \ .$$

   Finally,

   $$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{500^2}{500 \cdot 1000 + 500 \cdot 1000 - 500^2} = \frac{1}{3}$$

```python
print(Fraction(500**2, 500*1000 + 500*1000 - 500**2))
# 1/3
```

3. What is the probability that at least one of R_1 or R_2 is equal to 1000?

   Here the event $C =$ "at least one of R_1 or R_2 equals 1000" is

$$C = \{(1000, R_2) : R_2 \in \{1, \ldots, 1000\} \cup \{(R_1, 1000) : R_1 \in \{1, \ldots, 1000\}\}$$

   and using the principle of inclusion exclusion,

$$|C| = 1000 + 1000 - |\{1000, 1000\}| = 1999 \ ,$$

   so

$$\Pr(C) = \frac{|C|}{|S|}\frac{1999}{1000^2} = \frac{1999}{1000000} \ .$$

4. Suppose I tell you that at least one of R_1 or R_2 is even. What is the probability that at least one of R_1 or R_2 is equal to 1000?

   Notice that, $C \subseteq B$ since 1000 is even, so $B \cap C = C$. Now we're looking at

$$\Pr(C \mid B) = \frac{\Pr(B \cap C)}{\Pr(B)} = \frac{|B \cap C|}{|B|} = \frac{|C|}{|B|} = \frac{1999}{500 \cdot 1000 + 500 \cdot 1000 - 500^2} = \frac{1999}{750000}$$

```python
print(Fraction(1999, 500*1000 + 1000*500 - 500**2))
# 1999/750000
```

5. Suppose I tell you that at least one of R_1 or R_2 is even. What is the probability that at least one of R_1 or R_2 is equal to 999?

   Let $D =$ "at least one of R_1 or R_2 equals 999". Then

$$D = \{(999, R_2) : R_2 \in \{1, \ldots, 1000\} \cup \{(R_1, 999) : R_1 \in \{1, ldots, 1000\}$$

   and

$$|D| = 1000 + 1000 - |\{(999, 999)\}| = 1999 \ .$$

   Now, $D \cap B =$ "at least one R_i is even and the other is 999" and

$$D \cap B = \{(999, 2k) : k \in \{1, \ldots, 500\}\} \cup \{(2k, 999) : k \in \{1, \ldots, 500\}\}$$

   The two sets that define $D \cap B$ are disjoint, so we can use the Sum Rule to get

$$|D \cap B| = 500 + 500 = 1000$$

   Finally,

$$\Pr(D \mid B) = \frac{\Pr(D \cap B)}{\Pr(B)} = \frac{|D \cap B|}{|B|} = \frac{1000}{750000} \ .$$

6 Uniqueness of Maximum and Median

Let n be an odd integer. The *median* of a sequence $x_1, \ldots, x_n$, denoted of numbers, $\mathrm{median}(x_1, \ldots, x_n)$, is the unique value x such that there are at least $\lceil n/2 \rceil$ values less than or equal to x and at least $\lceil n/2 \rceil$ values greater than or equal to x.

For example,

$$\mathrm{median}(8, 4, 3, 4, 7, 5, 6) = 5$$

since $4, 3, 4, 5 \leq 5$ and $8, 7, 5, 6 \geq 5$; and

$$\text{median}(8, 5, 3, 4, 7, 5, 6) = 5$$

since $5, 3, 4, 5 \leq 5$ and $8, 5, 7, 5, 6 \geq 5$.

Consider a uniform random sequence $x_1, \dots, x_7$ of 7 numbers each chosen from the set $\{1, 2, 3, \dots, 10\}$

1. What is the probability that $\max\{x_1, \dots, x_7\}$ occurs exactly once in $x_1, \dots, x_7$?

   Here the underlying probability space is uniform over the set $S = \{1, \dots, 10\}^7$ and as size 10^7.

   Let A = "the maximum of $x_1, \dots, x_7$ is unique". We need to break this down a bit. For each $i \in \{1, \dots, 10\}$, let A_i = "the maximum of $x_1, \dots, x_7$ is unique and is equal to i". First we use the Product Rule to determine $|A_i|$ using the following procedure:

   (a) Pick an index $j \in \{1, \dots, 7\}$ and set $x_j \leftarrow i$.
   (b) Assign values from $\{1, \dots, i-1\}$ to each of $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_7$.

   There are 7 ways to perform the first step and $(i-1)^6$ ways to perform the second step, so

   $$|A_i| = 7 \cdot (i-1)^6 \ .$$

   Notice that $A_1, \dots, A_{10}$ are pairwise disjoint and $A = A_1 \cup A_2 \cup \dots \cup A_{10}$, so we can use the Sum Rule to finish with

   $$\Pr(A) = \sum_{i=1}^{10} \Pr(A_i) = \sum_{i=1}^{10} \frac{|A_i|}{|S|} = \sum_{i=1}^{10} \frac{7 \cdot (i-1)^6}{10^7} = \frac{1369767}{2000000} \approx 0.6848835$$

```
print(sum([Fraction(7 * (i-1)**6, 10**7) for i in range(1, 11)]))
# 1369767/2000000
```

2. What is the probability that $\text{median}(x_1, \dots, x_7)$ occurs exactly once in $x_1, \dots, x_7$?

   We follow the same basic approach as the previous question. Let A = "the median of $x_1, \dots, x_7$ is unique". For each $i \in \{1, \dots, 10\}$, let A_i = "the median of $x_1, \dots, x_7$ is unique and is equal to i".

   Let's determine $|A_i|$ using the Product Rule with the following procedure:

   (a) Choose an index $j \in \{1, \dots, 7\}$ and set $x_j \leftarrow i$.
   (b) Choose three indices $\{a, b, c\} \subset \{1, \dots, 7\} \setminus \{j\}$.
   (c) Choose three values from $\{1, \dots, i-1\}$ to assign to x_a, x_b, and x_c.
   (d) Choose three values from $\{i+1, \dots, 10\}$ to assign to x_r for each $r \in \{1, \dots, 7\} \setminus \{j, a, b, c\}$.

   The number of ways to execute this procedure and, hence, the size of A_i is

   $$|A_i| = 7 \cdot \binom{6}{3} \cdot (i-1)^3 \cdot (10-i)^3 \ .$$

   Again, the $A_1, \dots, A_{10}$ are pairwise disjoint and $A = A_1 \cup \dots \cup A_{10}$, so we use the Sum Rule to get

   $$\Pr(A) = \sum_{i=1}^{10} \Pr(A_i) = \sum_{i=1}^{10} \frac{|A_i|}{|S|} = \sum_{i=1}^{10} \frac{7 \cdot \binom{6}{3} \cdot (i-1)^3 \cdot (10-i)^3}{10^7} = \frac{7476}{15625} \approx 0.478464 \ .$$

```
print(sum([Fraction(7 * binom(6,3) * (i-1)**3 * (10-i)**3, 10**7)
           for i in range(1, 11)]))
# 7476/15625
```