Probability & Statistics: Assignment 2

Due on September 28, 2018 at 8:30am

Prof. D. Eynard

A. Romanelli / A. Vicini

Problem 1

a) The proof that Ω is a discrete probability space lies within the definition of the set Ω itself which is given to us being the set of all integer numbers between 1 and 12, thus a bounded interval and surely a finite one.

b) In order to prove that p is a probability measure we need to show the three following properties:

    i. Normalisation

    ii. Positivity

    iii. Additivity for disjoint sets

i. Normalisation   We want to show that $P(\Omega) = 1$, knowing that $P(\{\omega\}) = \frac{1}{12}, \forall \omega \in \Omega$, as given to us by the problem's premises. This is indeed true if we compute the summation of the probabilities of all the faces of the given die:

$$\sum_{i=1}^{12} \left(\frac{1}{12} \right) = \frac{1}{12} + \frac{1}{12} + \cdots + \frac{1}{12} = 1$$

ii. Positivity   To prove this property, we will need to make sure that it doesn't exist a $P(\{\omega\}) < 0, \forall \omega \in \Omega$. But as we already know that all the faces of the die share the same probability $P(\{\omega\}) = \frac{1}{12} > 0$, thus $P(\{\omega\}) > 0, \forall \omega \in \Omega$.

iii. Additivity   Let's consider to possible events of Ω, namely A and B. When throwing the die, the chance of an event $\omega \in (A \cup B)$ happening is twice the chance of an event $\omega \in A$ or $\omega \in B$, which we know being $\frac{1}{12}$.

$$P(A \cup B) = P(A) + P(B)$$
$$\sum_{\omega \in A \cup B} p(\omega) = \sum_{\omega \in A} p(\omega) + \sum_{\omega \in B} p(\omega)$$
$$\frac{2}{12} = \frac{1}{12} + \frac{1}{12}$$
$$\frac{2}{12} = \frac{2}{12} \quad \square$$

Moreover we can say that over a single throw there can be no overlapping of events, which means that all events are disjoint sets.

c) Given the following set of events:

- A_1 : the outcome is a even number : $\{2, 4, 6, 8, 10, 12\}$
- A_2 : the outcome is greater than six : $\{7, 8, 9, 10, 11, 12\}$
- A_3 : the outcome is a number which is a multiple of three : $\{3, 6, 9, 12\}$

We can see already that we cannot compute $P(\bigcup_{i=1}^{3} A_i)$ because there are clearly elements in common between the three events, like the outcome involving a 12, which is contained within each of the three events. Thus, we cannot compute the probability because $\exists i, j \in \{1, 2, 3\} : A_i \cap A_j \neq \emptyset$.

2

d) For the inclusion/exclusion principle, we have no constraints and we can thus compute the probability of $P(A_1 \cup A_2 \cup A_3)$:

$$P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_3 \cap A_1) + P(A_1 \cap A_2 \cap A_3)$$

$$\frac{|A_1|}{12} + \frac{|A_2|}{12} + \frac{|A_3|}{12} - \frac{|A_1 \cap A_2|}{12} - \frac{|A_2 \cap A_3|}{12} - \frac{|A_3 \cap A_1|}{12} + \frac{|A_1 \cap A_2 \cap A_3|}{12}$$

$$\frac{1}{2} + \frac{1}{2} + \frac{1}{3} - \frac{1}{4} - \frac{1}{6} - \frac{1}{6} + \frac{1}{12} = 0.8\overline{3}$$

Problem 2

a) Based on my catch our estimate would be to have a 30% of red fishes and 70% of white fishes, although according to the *Law of Large Numbers* we would need several attempts to verify whether this estimate is accurate or not.

b)
$$p(\text{catch}) = \frac{90 \times 89 \times 88 \times 87 \times 86 \times 85 \times 84 \times 10 \times 9 \times 8}{100 \times 99 \times 98 \times 97 \times 96 \times 95 \times 94 \times 93 \times 92 \times 91} \approx 0.000431...$$

This gives us the probability of one of the multiple permutations of our catch. In order to compute all these possible permutation we need to compute them with the following formula:

$$\frac{10!}{7! \cdot 3!} = 120$$

$$p(\text{catch}) \times 120 \approx 0.05179$$

c)
$$p(\text{catch}) = \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 90 \times 89 \times 88}{100 \times 99 \times 98 \times 97 \times 96 \times 95 \times 94 \times 93 \times 92 \times 91} \approx 6.786 \times 10^{-9}$$

Again, we need to multiply this by the number of possible permutations, which are still 120:

$$6.786 \times 10^{-9} \times 120 = 8.144 \times 10^{-7}$$

d) 0%, since we couldn't possibly catch 7 white fishes if there are only five in the pond.
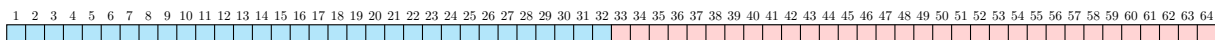
e)
$$\frac{R! \cdot (N-R)! \cdot (N-n)!}{N! \cdot (R-r)! \cdot (W-w)!} \cdot \frac{n!}{r! \cdot w!}$$
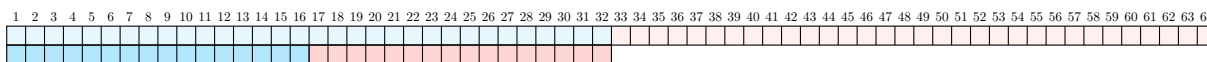
Problem 3

This strategy is not really a viable one if what we aim for is to have a higher percentage of male against females and here's why:

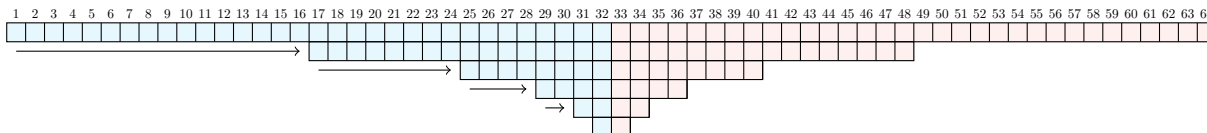Let's take 2^n of the female population as a sample:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64

Under the assumption that we have equal birth rate between males and females, half of those women are going to give birth to a girl, whereas the other half is going to give birth to a boy.

Now, only the women who did give birth to a male are going to give birth again:

3

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64

To better visualise what's going on, we're going to shift the sample so that we always have the midpoint in the middle of the page, since at each iteration we lose half of the population which gave birth to a girl. We repeat this until there's no one else left who can give birth to anyone.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64

It's clearly a zero-sum process. Whenever a girl is born, a male must also be born and the two cancel out, not influencing the ratio between females and males.

For any n of women, if we let the boys being born as B and the girls as G we know that $B = G = \frac{n}{2}$ we can say that there's going to be $\log_2(n)$ generations before no one can give birth.

If every woman is going to stop as soon as she gives birth to a woman, then we can certainly say that the number of girls is going to be surely n. We can compute the number of boys B as the following summation:

$$\sum_{i=1}^{\log_2 n} \left(\frac{1}{2^i} n \right)$$

But this can really be simplified to:

$$\sum_{i=1}^{\log_2 n} \left(\frac{1}{2^i} n \right) = n - 1$$

Which clearly shows how $B < G, \forall n \in \mathbb{N}$ and thus this is not a viable to way to increase the percentage of males in the realm.

Problem 4

To tackle this problem we can try to think about it in smaller terms and see how this does scale and if the scale of the problem (aka the seats available) count towards the solution. Let's imagine we have only two seats:

1	2

The chances of the first person getting his seat is exactly $\frac{1}{2}$, to randomly get his rightful place or get the wrong one. If we scale this problem up to 3 seats we get the following:

1	2	3

Where in $\frac{1}{3}$ of the cases, the person who lost the ticker finds again the right place, whereas in the other cases it's taking someone else's spot.
If it is taking someone else's spot $\frac{2}{3}$ of the cases, we need to count what's the chance that out of the two seats left the second guy picks the seat of the first guy, bringing back balance into the situation and granting the third person the rightful seat.
But that was the exact same situation that we explored previously, where with two seats we would have a 50% chance! We can then start to explore a recursive way to model this problem.

Let's reconsider again the problem but with the 100 seats:

$$\boxed{1}\,\boxed{2}\quad\cdots\quad\boxed{99}\,\boxed{100}$$

In this instance we will proceed exactly as we did for the previous example. The first person boarding the plane has a $\frac{1}{100}$ of getting the right spot but then we need to add to this chance the possibility of him sitting somewhere else and stealing a seat.

This happens in the rest of the cases and we need to consider what's the chance of <u>NOT</u> getting the seat of the last person, hence $\frac{99}{100} - \frac{1}{100} = \frac{98}{100}$ and we need to multiply this by the probability of the subproblem for 99 seats. We can then come up with the following formula:

$$p(n) = \frac{1}{n} + \frac{n-2}{n} \cdot p(n-1)$$

With $n \geq 2$ and $p(2) = \frac{1}{2}$

$$p(3) = \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$
$$p(4) = \frac{1}{4} + \frac{2}{4} \cdot \frac{1}{2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

We can see that there is a steady trend, can we prove though that $p(n) = \frac{1}{2}, \forall n \in \mathbb{N}$?

We'll give it a shot by induction:
We set our base case and verify whether $p(2) = \frac{1}{2}$

$$\frac{1}{2} + \frac{2-2}{2} \cdot p(1) = \frac{1}{2}$$

$$\frac{1}{2} + 0 = \frac{1}{2}$$

Our formula holds for $n = 2$, let's now assume that $p(n) = \frac{1}{2}$ is true for some n and attempt to prove this for $n+1$.

$$\frac{1}{n+1} + \frac{n+1-2}{n+1} \cdot p(n) = \frac{1}{2}$$
$$\frac{1}{n+1} + \frac{n+1-2}{n+1} \cdot \frac{1}{2} = \frac{1}{2}$$
$$\frac{2}{n+1} + \frac{n+1-2}{n+1} = 1$$
$$\frac{2}{n+1} + \frac{n-1}{n+1} = 1$$
$$\frac{n+1}{n+1} = 1 \Rightarrow 1 = 1, \forall n \in \mathbb{N}$$