Probability puzzles (Getting Serious)
Hong Wai Ng

This notes contains my attempts to answer all questions by the android app ‘Probability Puzzles’, the getting serious category. If you found any mistake in this note, do inform me at hongwai1920@gmail.com.

1. Let X be the number of tosses required to get a number greater than 4. Then X follows a discrete uniform distribution with $p = \frac{2}{6} = \frac{1}{3}$. Therefore,
$$\mathbb{E}(X) = \frac{1}{p} = 3.$$

2. By the complement rule, we have
$$\begin{aligned}\mathbb{P}(\text{at least two people share birthday}) &= 1 - \mathbb{P}(\text{all people have different birthday}) \\ &= 1 - \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \\ &= 1 - \frac{364 \times 363}{365^2}\end{aligned}$$

3. The smallest group size refers to when there is exactly a pair of people sharing the same birthday. Let n be the group size. So,
$$\begin{aligned}\mathbb{P}(\text{exactly 2 people share the same birthday}) &= 1 - \mathbb{P}(\text{nobody shares the same birthday}) \\ &= 1 - \frac{365}{365} \times \frac{365-1}{365} \times \cdot \times \frac{365-(n-1))}{365} \\ &= 1 - 1 \times \left(1 - \frac{1}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right)\end{aligned}$$
For small enough x, we have the approximation
$$e^x \approx 1 + x.$$
It follows that
$$\begin{aligned}\mathbb{P}(\text{exactly 2 people share the same birthday}) &= 1 - 1 \times \left(1 - \frac{1}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right) \\ &\approx 1 - e^{-\frac{1}{365}} \times \cdots \times e^{-\frac{n-1}{365}} \\ &= 1 - e^{-\frac{1}{365} \times \frac{(n-1)n}{2}}\end{aligned}$$
Therefore, we need to find n such that
$$\begin{aligned}&1 - e^{-\frac{1}{365} \times \frac{(n-1)n}{2}} > \frac{1}{2} \\ &\Leftrightarrow n^2 - n - 730 \ln 2 > 0 \\ &\Leftrightarrow n < -22 \quad \text{or} \quad n > 22.999\end{aligned}$$

So, the minimum n is 23,

4. Let X be the number of tosses required to get two consecutive heads. By first step analysis,

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X|H)\cdot\mathbb{P}(H) + \mathbb{E}(X|T)\cdot\mathbb{P}(T)\\ &= \frac{1}{2}\left[\mathbb{E}(X|H) + \mathbb{E}(X|T)\right]\end{aligned}$$

Since $\mathbb{E}(X|T) = 1 + \mathbb{E}(X)$, so

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X|H)\cdot\mathbb{P}(H) + \mathbb{E}(X|T)\cdot\mathbb{P}(T)\\ &= \frac{1}{2}\left[\mathbb{E}(X|H) + 1 + \mathbb{E}(X)\right]\end{aligned},$$

which implies that

$$\mathbb{E}(X) = \mathbb{E}(X|H) + 1$$

On the other hand, by first step analysis on $\mathbb{E}(X|H)$, we have

$$\begin{aligned}\mathbb{E}(X|H) &= \frac{1}{2}\left[\mathbb{E}(X|HH) + \mathbb{E}(X|HT)\right]\\ &= \frac{1}{2}\left[2 + \mathbb{E}(X) + 2\right]\end{aligned}$$

Therefore,

$$\mathbb{E}(X) = 6.$$

General problem: The expected number of tosses to get n consecutive heads is

$$2^{n+1} - 2.$$

5. By Bayes' theorem, we have

$$\begin{aligned}\mathbb{P}(\text{fair}|H) &= \frac{\mathbb{P}(H|\text{fair})\cdot\mathbb{P}(\text{fair})}{\mathbb{P}(H)}\\ &= \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+\frac{2}{2}\cdot\frac{1}{3}+\frac{0}{2}\cdot\frac{1}{3}}\\ &= \frac{1}{3}\end{aligned}$$

Alternatively, since there are 3 heads in total and the fair coin consists of 1 head among the 3, so the required probability is $\frac{1}{3}$.

6. Let X be the number of boys. Then X follows a binomial distribution with $n = 3$ and $p = 0.51$. So,

$$\begin{aligned}\mathbb{P}(X=2) &= \binom{3}{2}(0.51)^2(0.49)^1 \\ &= 3\cdot(0.51)^2\cdot 0.49\end{aligned}$$

7. By Bayes' theorem, we have

$$\begin{aligned}\mathbb{P}(\text{have disease} \mid \text{positive}) &= \frac{\mathbb{P}(\text{positive}|\text{have disease})\cdot\mathbb{P}(\text{have disease})}{\mathbb{P}(\text{positive})} \\ &= \frac{0.99\cdot\frac{150}{100000}}{0.99+\frac{100000-150}{100000}\cdot 0.01} \\ &= \frac{148.5}{148.5+998.5}\end{aligned}.$$

8. By Bayes' theorem, we have

$$\begin{aligned}\mathbb{P}(\text{double-headed coin}|\ 3\text{H}) &= \frac{\mathbb{P}(3\text{H}|\text{double-headed coin})\cdot\mathbb{P}(\text{double-headed coin})}{\mathbb{P}(3\text{H})} \\ &= \frac{1\cdot\frac{1}{10}}{1\cdot\frac{1}{10}+\frac{1}{2^3}\cdot\frac{9}{10}} \\ &= \frac{8}{17}.\end{aligned}$$

9. Let X be the number of boxes of cereal to collect a complete collection. Then

$$X = X_1+X_2+X_3+X_4$$

where each X_i is the number of boxes of cereal required to collect ith toy after (i-1) many toys are collected for $i=1,2,3,4$. Since collecting toys is independent, so each X_i follows a discrete uniform distribution with

$$p_i = \frac{4-(i-1)}{4}.$$

Therefore,

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X_1)+\mathbb{E}(X_2)+\mathbb{E}(X_3)+\mathbb{E}(X_4) \\ &= \frac{4}{4}+\frac{4}{3}+\frac{4}{2}+\frac{4}{1} \\ &= \frac{25}{3}.\end{aligned}.$$

This question is related to Coupon collector's problem.

10. Let X and Y be the number of successful tosses for 4 and 6 respectively. Then $X\sim B(4,p)$ and $Y\sim B(6,p)$ where p is his skill. We need to find p that satisfies

$$\mathbb{P}(X\geq 2) = \mathbb{P}(Y\geq 3),$$

that is,

$$\begin{aligned}\mathbb{P}(X \leq 1) &= \mathbb{P}(Y \leq 2) \\ \Leftrightarrow (1-p)^4 + 4p(1-p)^3 &= (1-p)^6 + 6p(1-p)^5 + 15p^2(1-p)^4 \\ \Leftrightarrow 1 - p + 4p &= (1-p)^3 + 6p(1-p)^2 + 15p^2(1-p) \\ \Leftrightarrow p &= \frac{3}{5}.\end{aligned}$$

11. We consider the following 3 cases.
    a. All A. B and C miss in their first try
    b. A misses B but B kills C
    c. A kills B but C misses A

In case a, the game restarts as if nothing happens.
In cases b and c, we just have to consider the game for two people.
For example, in case b,

$$\begin{aligned}\mathbb{P}(\text{A wins}|\text{A,B remain}) &= \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{2} + \cdots \\ &= \frac{1}{2} \sum_{k=0}^{\infty} \left(\frac{1}{8}\right)^k \\ &= \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{8}} \\ &= \frac{4}{7}.\end{aligned}$$

On the other hand, in case c,

$$\begin{aligned}\mathbb{P}(\text{A wins}|\text{A,C remain}) &= \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \cdots \\ &= \frac{1}{2} \sum_{k=0}^{\infty} \left(\frac{1}{4}\right)^k \\ &= \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{4}} \\ &= \frac{2}{3}.\end{aligned}$$

Therefore, applying the law of total probability to cases a,b,c, we have

$$\begin{aligned}\mathbb{P}(\text{A wins}) &= \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{2} \cdot \mathbb{P}(\text{A wins}) + \frac{1}{2} \cdot \frac{3}{4} \cdot \mathbb{P}(\text{A wins}|\text{A, B remain}) + \frac{1}{2} \cdot \frac{1}{2} \cdot \mathbb{P}(\text{A wins}|\text{A, C remain}) \\ &= \frac{1}{16}\mathbb{P}(\text{A wins}) + \frac{3}{8} \cdot \frac{4}{7} + \frac{1}{4} \cdot \frac{2}{3} \\ &= \frac{1}{16}\mathbb{P}(\text{A wins}) + \frac{8}{21}.\end{aligned}$$

It follows that

$$\mathbb{P}(\text{A wins}) = \frac{128}{315}.$$

12. Let $f(n)$ be the number of ways to reach nth stone. Since the grasshopper can hop either 1 or 2 steps, so

$$f(n) = f(n-1) + f(n-2), \quad \text{for all } n \geq 3$$
$$f(1) = 1, \quad f(2) = 2.$$

Clearly, by some manual calculations, we have

$$f(9) = 55.$$

One might note that the recurrence is the Fibonacci sequence and thus one might use the closed form formula to evaluate $f(9)$.

13. Observe that the number of ways is determined by the length of the board. A vertical 2-by-1 tile will take 1 slot whereas a horizontal 2-by-1 tile will take 2 slots. So, this question is similar to question 12.

Let $f(n)$ be the number of ways to assign 2-by-1 tiles into the board when the board has length n. Clearly

$$f(n) = f(n-1) + f(n-2), \quad \text{for all } n \geq 3$$
$$f(1) = 1, \quad f(2) = 2.$$

Therefore, $f(10) = 89$.

14. Let $f(n)$ be the number of binary strings of length n that contain at least 2 consecutive ones. Then there are 3 cases:
    a. The string ended with 0
    b. The string ended with 1
        i. The string ended with 01
        ii. The string ended with 11

For case a, there are $f(n-1)$ possible strings. For case b(i), there are $f(n-2)$ possible strings. For case b(ii), since it contains 2 consecutive ones already, there are 2^{n-2} possible strings. It follows that the recurrence is

$$f(n) = f(n-1) + f(n-2) + 2^{n-2} \quad \text{for all } n \geq 3$$
$$f(1) = 0, \quad f(2) = 1$$

Therefore, $f(8) = 201$. Hence, the required probability is

$$\frac{201}{2^8}$$

15. This question is similar to pricing American option with 6 states at each time step. We use dynamic programming to solve the question.

At the final stage, our expected value is 3.5. At the second last stage, we should continue the game if the score is less than 3.5. Otherwise, we take the score. This resulted in expected value

$$\frac{1}{6}\cdot 3.5+\frac{1}{6}\cdot 3.5+\frac{1}{6}\cdot 3.5+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 6=4.25.$$

Lastly, at the first stage, we perform similar analysis as above: continue the game if the score is less than 4.25. Otherwise, we take it. This gives expected value

$$\frac{1}{6}\cdot 4.25+\frac{1}{6}\cdot 4.25+\frac{1}{6}\cdot 4.25+\frac{1}{6}\cdot 4.25+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 6=\frac{14}{3}$$

16. This question can be solved by symmetry. Either the first seat is taken before the last seat or vice versa. Both events have the same probability, which is 0.5.

17. Let X be the money received and t be the threshold. Then $X\sim U(0,100)$. By the law of total expectation,

$$\begin{aligned}\mathbb{E}(X)&=\mathbb{E}(X|X>t)\cdot\mathbb{P}(X>t)+\mathbb{E}(X|X\leq t)\cdot\mathbb{P}(X\leq t)\\&=\left(\frac{100+t}{2}\right)\cdot\left(\frac{100-t}{100}\right)+0.9\mathbb{E}(X)\cdot\frac{t}{100}\\&=\frac{100^2-t^2}{200}+\frac{9t}{1000}\mathbb{E}(X)\end{aligned}$$

Simplifying equation above leads to

$$\mathbb{E}(X)=\frac{5(100^2-t^2)}{1000-9t}$$

Differentiate the expectation with respect to t gives

$$\frac{d\mathbb{E}(X)}{dt}=\frac{(1000-9t)(-10t)-5(100^2-t^2)(-9)}{(1000-9t)^2}=0.$$

Simplifying equation above, we have

$$\begin{aligned}9t^2-2000t+90000=0\\t\approx 159.5433\quad\text{or}\quad t\approx 62.6789\end{aligned}.$$

Clearly the former is a local minimum and the latter is a local maximum, so the answer is 62.6789.

18. Let P be the probability that the entire population dies out. By the law of total probability, we have

$$\begin{aligned}P&=\mathbb{P}(\text{both die})\cdot 1+\mathbb{P}(\text{one die another don't})\cdot P+\mathbb{P}(\text{both survive})\cdot P^2\\&=\frac{1}{4}\cdot\frac{1}{4}+2\cdot\frac{1}{4}\cdot\frac{3}{4}\cdot P+\frac{3}{4}\cdot\frac{3}{4}\cdot P^2\\&=\frac{1}{16}+\frac{3}{8}P+\frac{9}{16}P^2.\end{aligned}$$

Simplifying the equation above gives

$$9P^2 - 10P + 1 = 0.$$

Therefore, $P = \dfrac{1}{9}$.

19. Let X, YZ be the integer multiples of $7, 11$ and 13 respectively. Then

$$|X| = 142, \quad |Y| = 90 \quad \text{and} \quad |Z| = 76.$$

Now, by the exclusion-inclusion principle, we have

$$\begin{aligned} |X \cup Y \cup Z &= |X| + |Y| + |Z| - (|X \cap Y| + |X \cap Z| + |Y \cap Z|) + |X \cap Y \cap Z| \\ &= 142 + 90 + 76 - (12 + 10 + 6) + 0 \\ &= 280. \end{aligned}$$

Therefore, the required probability is $\dfrac{280}{1000}$.

20. On the first day we have at least 1 weather block with probability 1. On the next day, the probability of the weather changing and giving us another block is

$$\frac{4}{5} \times \frac{3}{5} + \frac{1}{5} \times \frac{4}{5} = \frac{16}{25}.$$

This occurs $n-1$ days and thus the expected number of weather blocks for n days is given by

$$E(X_n) = 1 + (n-1)\frac{16}{25}.$$

Plugging in $n = 10$ we expect $\dfrac{169}{25}$ weather groups.

21. Using similar reasoning as above, the required probability is

$$1 + 9 \times (0.5 \times 0.5 + 0.25 \times 0.75 + 0.25 \times 1.05)$$

22. Let n be the number of red marbles. So there are also n many black marbles.
Assume that 2 of the same colours are taken out, say $n-2$ black and n red marbles.
Then there are $\dbinom{n-2}{2}$ ways to choose 2 black marbles and $\dbinom{n}{2}$ ways red marbles.
Also, there are $\dbinom{2n-2}{2}$ ways to choose any two marbles. Therefore,

$$\begin{aligned} \frac{\binom{n-2}{2}+\binom{n}{2}}{\binom{2n-2}{2}} &= \frac{1}{2} \\ \Leftrightarrow (n-2)(n-3)+n(n-1) &= \frac{1}{2}(2n-2)(2n-3) \\ \Leftrightarrow n &= 3 \end{aligned}.$$

Hence, there are in total 6 marbles.
Note that we need not to consider the case for n black and $n-2$ red marbles.

23. Let M and W be maths and writing abilities respectively. So, $M, W \sim U(0,1)$. It follows that

$$\begin{aligned} \mathbb{P}(M > 0.9 | M + W > 1) &= \frac{0.5(1+0.9)0.1}{0.5} \\ &= 0.19 \end{aligned}.$$

24. The proportion will always be the same as the probability. So, the answer is 0.49.

25. Answer is $2 \times 0.51 \times 0.49^2$.

26. Assuming that the line passes through $(0, 20)$ and $(100, 80)$, the slope is

$$\frac{80-20}{100-0} = 0.6.$$

27. Clearly $\text{shorter piece} \sim U(0, 0.5)$ and $\text{longer piece} \sim U(0.5, 1)$. Therefore,

$$\mathbb{E}(\text{longer piece}) = 0.75$$

28. The answer is $\frac{1}{12}$ because this is a Markov Chain with invariant distribution

$$\left(\frac{1}{12}, \ldots, \frac{1}{12}\right).$$

Convergence to equilibrium is guaranteed from any starting state, because the Chain is aperiodic and positive-recurrent.

29. For nonnegative integer n and $k \in \{0,1,2,3,4,5\}$, let $p_{n,k}$ denote the probability that after n minutes there are k persons in room that does *not* contain more persons than other rooms.

Then we are interested in the sequence $(p_{n,5})_n$.
We find the following equalities:

$$2p_{0,0} = 1 = 2p_{0,1} \quad \text{and} \quad p_{0,2} = p_{0,3} = p_{0,4} = p_{0,5} = 0.$$

And for every n:

$$10p_{n+1,0} = p_{n,1}$$
$$10p_{n+1,1} = 10p_{n,0} + 2p_{n,2}$$
$$10p_{n+1,2} = 9p_{n,1} + 3p_{n,3}$$
$$10p_{n+1,3} = 8p_{n,2} + 4p_{n,4}$$
$$10p_{n+1,4} = 7p_{n,3} + 10p_{n,5}$$
$$10p_{n+1,5} = 6p_{n,4}$$

Pre-assuming that for every k limit $p_k := \lim_{n\to\infty} p_{n,k}$ exists we find:

$$p_0 + p_1 + p_2 + p_3 + p_4 + p_5 = 1$$
$$10p_0 = p_1$$
$$10p_1 = 10p_0 + 2p_2$$
$$10p_2 = 9p_1 + 3p_3$$
$$10p_3 = 8p_2 + 4p_4$$
$$10p_4 = 7p_3 + 10p_5$$
$$10p_5 = 6p_4$$

This can be solved and leads to:

$$p_5 = \frac{126}{512} = \frac{63}{256} = 0.24609375.$$

30. Note that

$$\begin{aligned}
\mathbb{P}(S\uparrow|(N+1)\uparrow, N\downarrow) &= \frac{\mathbb{P}((N+1)\uparrow, N\downarrow|S\uparrow)\cdot\mathbb{P}(S\uparrow)}{\mathbb{P}((N+1)\uparrow, N\downarrow)} \\
&= \frac{(0.6)^{N+1}\cdot(0.4)^N\cdot 0.5}{(0.6)^{N+1}\cdot(0.4)^N\cdot 0.5 + (0.6)^N\cdot(0.4)^{N+1}\cdot 0.5} \\
&= \frac{0.6}{0.6+0.4} \\
&= 0.6
\end{aligned}$$

31. Let X be the number of distinct objects in our sample. Then

$$X = \frac{X_1 + X_2 + \cdots + X_N}{N}$$

where $X_i \in \{0, 1\}$ with $X_i = 1$ means that the ith object is in our sample. It follows that

$$
\begin{aligned}
\mathbb{E}(X) &= \frac{1}{N}\sum_{k=1}^{N}\mathbb{E}(X_k) \\
&= \frac{1}{N}\sum_{k=1}^{N}\mathbb{P}(X_k = 1) \\
&= \frac{1}{N}\sum_{k=1}^{N}[1-\mathbb{P}(X_k = 0)] \\
&= 1-\frac{1}{N}\left[\sum_{k=1}^{N}\mathbb{P}(X_k = 0)\right] \\
&= 1-\frac{1}{N}\sum_{k=1}^{N}\left(\frac{(N-1)^N}{N^N}\right) \\
&= 1-\left(1+\frac{-1}{N}\right)^N
\end{aligned}
$$

Since

$$\lim_{N\to\infty}\left(1+\frac{-1}{N}\right)^N = e^{-1},$$

so

$$\lim_{N\to\infty}\mathbb{E}(X) = 1-e^{-1}.$$

32. This question is related to Bertrand's ballot theorem and Catalan number
Imagine a xy-plane where x-axis is the number of votes and y-axis is the number of C's votes that exceed T's votes. In other words, a C's vote is +1 whereas a T's vote is -1.
Note that the first vote chosen must be C's vote with probability $\frac{7}{10}$. It follows that we are left with 6 votes of C and 3 votes of T.
Now, we are going to calculate the number of paths that start at $(1,1)$ which will reach $y=0$ at some point before $x=10$, since the path must end at $(10,4)$. By the reflection principle, it is the same as the number of paths that start at $(1,1)$ which will end at $(10,-4)$. There are $\binom{9}{2}$ such paths. Hence, the required probability is

$$\frac{7}{10}\times\frac{\binom{9}{3}-\binom{9}{2}}{\binom{9}{3}} = 0.4.$$

33. Answer is the same as the probability that the true message is 1111, which is 0.6.