Probability puzzles (Outrageous)
Hong Wai Ng

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

1. Let X be the number of foxes immediately followed by a hound and X_i the fox is immediately followed by a hound. By linearity of expectation, we have

$$\begin{aligned}\mathbb{E}(X) &= \sum_{i=1}^{5} \mathbb{E}(X_i) \\ &= \sum_{i=1}^{5} \mathbb{P}(X_i = 1) \\ &= 5\mathbb{P}(X_1 = 1) \\ &= 5 \cdot \frac{7}{12} \\ &= \frac{35}{12}\end{aligned}.$$

2. Each path consists of 7 norths and 4 east. Therefore, there are $\binom{11}{7}$ paths.

3. We can calculate the total number of paths minus the number of paths that travel (7,4).

   Clearly the total number of paths is $\binom{11}{7}$.
   Therefore, the required probability is

$$\frac{\binom{11}{7} - \binom{7}{2} \cdot \binom{4}{2}}{\binom{11}{7}} \approx 0.6182$$

4. Let X_A and X_B be the next arrival time of train A and B respectively. Then $X_A \sim U(0,5)$ and $X_B \sim U(0,3)$. Since X_A and X_B are independent, so X_A and X_B are independent, so

$$
\begin{aligned}
\mathbb{E}[(\min(X_A, X_B)] &= \int\int \min(x_A, x_B)\frac{1}{3}\cdot\frac{1}{5}dx_A dx_B \\
&= \frac{1}{15}\left(\int_0^3\int_0^{x_B} x_A dx_A dx_B + \int_0^3\int_{x_B}^5 x_B dx_A dx_B\right) \\
&= \frac{1}{15}\int_0^3 \frac{x_B^2}{2}dx_B + \frac{1}{15}\int_0^3 x_B(5-x_B)dx_B \\
&= \frac{1\cdot 3^3}{15\cdot 6} + \frac{1}{15}\left(5\cdot\frac{3^2}{2} - \frac{3^3}{3}\right) \\
&= \frac{3}{10} + \frac{9}{10} \\
&= 1.2
\end{aligned}.
$$

An alternative is to compute the expectation of minimum of two IID uniform distributions is to find its probability density function. Let

$$
Y = \min(X_A, X_B).
$$

For any y, we have

$$
\begin{aligned}
\mathbb{P}(Y \le y) &= \mathbb{P}[\min(X_A, X_B) \le y] \\
&= 1 - \mathbb{P}[\min(X_A, X_B) > y] \\
&= 1 - \mathbb{P}(X_A > y)\cdot\mathbb{P}(X_B > y) \\
&= 1 - \left(\frac{5-y}{5}\right)\left(\frac{3-y}{3}\right) \\
&= 1 - \frac{1}{15}(5-y)(3-y)
\end{aligned}.
$$

Differentiating the CDF with respect to y gives

$$
f_Y(y) = \frac{8}{15} - \frac{2}{15}y
$$

Therefore,

$$
\begin{aligned}
\mathbb{E}(Y) &= \int_0^3 y\left(\frac{8}{15} - \frac{2}{15}y\right)dy \\
&= \frac{8\cdot 3^2}{15\cdot 2} - \frac{2\cdot 3^3}{15\cdot 3} \\
&= \frac{12}{5} - \frac{6}{5} \\
&= 1.2
\end{aligned}.
$$

5. Let X be the number of potential couples and X_i be the number of man sitting adjacent to the ith woman for all $i = 1, 2, ..., 8.$ Then
$$X = X_1 + X_2 + \cdots + X_8.$$
By linearity of expectation, we have
$$\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{8} \mathbb{E}(X_i) \\
&= 8\mathbb{E}(X_1) \\
&= 8 \sum_{x=0}^{2} x\mathbb{P}(X_1 = x) \\
&= 8 \left[\mathbb{P}(X_1 = 1) + 2\mathbb{P}(X_1 = 2)\right] \\
&= 8 \left[\frac{\binom{5}{1} \cdot \binom{7}{1}}{\binom{12}{2}} + 2 \cdot \frac{\binom{5}{2}}{\binom{12}{2}} \right] \\
&= 8 \left(\frac{35}{66} + \frac{10}{33} \right) \\
&= \frac{20}{3}.
\end{aligned}$$
Alternatively, we can define Y_i to be the number of woman sitting adjacent to the ith man for all $i = 1, 2, ..., 5.$ Then
$$\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{5} \mathbb{E}(Y_i) \\
&= 5\mathbb{E}(Y_1) \\
&= 5 \sum_{y=0}^{2} y\mathbb{P}(y_1 = y) \\
&= 5 \left[\mathbb{P}(Y_1 = 1) + 2\mathbb{P}(Y_1 = 2)\right]. \\
&= 5 \left[\frac{\binom{8}{1} \cdot \binom{4}{1}}{\binom{12}{2}} + 2 \cdot \frac{\binom{8}{2}}{\binom{12}{2}} \right] \\
&= 5 \left(\frac{16}{33} + \frac{28}{33} \right) \\
&= \frac{20}{3}.
\end{aligned}$$
Another approach is to consider X_i be that $X_i = 1$ mean ith and (i+1)th people are of different gender and $X_i = 0$ mean they are of the same gender. Note that $i = 1, 2, ..., 13.$ Then the expected number of potential couples in a circular table is

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{13} \mathbb{E}(X_i) \\
&= 13\mathbb{E}(X_1) \\
&= 13\left[\frac{5}{13}\cdot\frac{8}{12} + \frac{8}{13}\cdot\frac{5}{12}\right]. \\
&= \frac{2\cdot 5\cdot 8}{12} \\
&= \frac{20}{3}
\end{aligned}
$$

6. Let X be the number of potential couples and X_i be that $X_i = 1$ mean ith and (i+1)th people are of different gender and $X_i = 0$ mean they are of the same gender. Note that $i = 1, 2, ..., 12$. Then
$$X = X_1 + X_2 + \cdots + X_{12}.$$
By linearity of expectation, we have
$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{12} \mathbb{E}(X_i) \\
&= 12\mathbb{E}(X_1) \\
&= 12\left[\frac{5}{13}\cdot\frac{8}{12} + \frac{8}{13}\cdot\frac{5}{12}\right]. \\
&= \frac{2\cdot 5\cdot 8}{13} \\
&= \frac{80}{13}
\end{aligned}
$$
General formula: If there are b boys and g girls, then there are expected number of potential couples in a line is
$$\mathbb{E}(X) = \frac{2mg}{m+g}.$$

7. Recall that if H is an event with positive probability, then
$$\mathbb{E}(X|H) = \frac{\mathbb{E}(X1_H)}{P(H)}.$$
Let $M, W \sim U(0,1)$. We need to calculate
$$Corr(M,W) = \frac{Cov(M,W|M+W>1)}{Var(M|M+W>1)}.$$
Since

$$\begin{aligned}\mathbb{E}(M1_{\{W+M>1\}}) &= \int_0^1 \int_{1-m}^1 m\,dm\,dw \\ &= \frac{1}{3}\end{aligned}.$$

and

$$P(M+W>1) = \frac{1}{2}.$$

it follows that

$$\begin{aligned}\mathbb{E}(M|W+M>1) &= \frac{\mathbb{E}(M1_{\{W+M>1\}})}{P(M+W>1)} \\ &= \frac{2}{3}\end{aligned}.$$

On the other hand,

$$\begin{aligned}\mathbb{E}(M^2 1_{\{W+M>1\}}) &= \int_0^1 \int_{1-m}^1 m^2\,dw\,dm \\ &= \frac{1}{4}\end{aligned}.$$

It follows that

$$\begin{aligned}\mathbb{E}(M^2|M+W>1) &= \frac{\mathbb{E}(M^2 1_{\{W+M>1\}})}{P(M+W>1)} \\ &= \frac{1}{2}\end{aligned}$$

Therefore,

$$\begin{aligned}Var(M|M+W>1) &= \mathbb{E}(M^2|M+W>1) - [\mathbb{E}(M|W+M>1)]^2 \\ &= \frac{1}{18}\end{aligned}$$

On the other hand,

$$\mathbb{E}(MW1_{\{M+W>1\}}) = \int_0^1 \int_{1-m}^1 mw\,dw\,dm = \frac{5}{24}.$$

So,

$$\begin{aligned}Cov(M,W|M+W>1) &= \mathbb{E}(MW|M+W>1) - [\mathbb{E}(M|M+W>1)]\cdot[\mathbb{M}(M|M+W>1)] \\ &= \frac{5}{12} - \frac{4}{9} \\ &= -\frac{1}{36}\end{aligned}$$

Hence,

$$
\begin{aligned}
Corr(M,W) &= \frac{Cov(M,W|M+W>1)}{Var(M|M+W>1)} \\
&= -\frac{1}{2}
\end{aligned}.
$$

8. This question is related to Derangement (Wiki). Let S_k be the set of permutations that its kth position is fixed, for $k=1,2,...,n$. By the inclusion-exclusion principle,

$$
\begin{aligned}
|S_1 \cup S_2 \cdots \cup S_n| &= \sum_{i=1}^{n} |S_i| - \sum_{1\le i<j\le n} |S_i \cap S_j| + \sum_{1\le i<j<k\le n} |S_i \cap S_j \cap S_k| \cdots \\
&= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)! + \cdots \\
&= \sum_{k=1}^{n} \binom{n}{k}(-1)^{k+1}(n-k)! \\
&= n!\sum_{k=1}^{n} \frac{(-1)^{k+1}}{k!}
\end{aligned}.
$$

Therefore,

$$
\mathbb{P}(\text{at least one fixed point}]) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k!}.
$$

It follows that

$$
\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(\text{at least one fixed point}]) &= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k!} \\
&= -\sum_{k=1}^{\infty} \frac{(-1)^{k}}{k!} \\
&= -\left(\sum_{k=0}^{\infty} \frac{(-1)^{k}}{k!} - 1\right) \\
&= 1 - e^{-1}
\end{aligned}.
$$

9. This question is similar to the Gambler's ruin Problem(Wiki). Let $S_0 = 2$,

$$
S_n = S_0 + X_1 + \cdots + X_n
$$

where X_i is a Bernoulli's random variable with

$$
\mathbb{P}(X_i = 1) = \frac{3}{4} \quad \text{and} \quad \mathbb{P}(X_i = -1) = \frac{1}{4}
$$

and

$$
\tau_N = \min\{n \ge 0 : S_n = N\},
$$

Note that τ_N is the first hitting of the random walk to level N. We also let

$$f(x) = \mathbb{P}(\tau_0 < \infty | S_0 = x).$$

By the law of total probability and the markov property of random walk, we have the following recurrence relation.

$$\begin{aligned}
f(x) &= \mathbb{P}(\tau_0 < \infty | S_1 = x+1, S_0 = x) \cdot \mathbb{P}(S_1 = x+1 | S_0 = x) + \mathbb{P}(\tau_0 < \infty | S_1 = x-1, S_0 = x) \cdot \mathbb{P}(S_1 = x-1 | S_0 = x) \\
&= \mathbb{P}(\tau_0 < \infty | S_0 = x+1) \cdot \frac{3}{4} + \mathbb{P}(\tau_0 < \infty | S_0 = x-1) \cdot \frac{1}{4} \\
&= \frac{3}{4} f(x+1) + \frac{1}{4} f(x-1).
\end{aligned}.$$

Simplifying it gives

$$3f(x+1) - 4f(x) + f(x-1) = 0.$$

It follows that its characteristic equation is

$$3t^3 - 4t + 1 = 0.$$

It has roots $\frac{1}{3}$ and 1. Therefore,

$$f(x) = A\left(\frac{1}{3}\right)^x + B$$

for some constants A, B. Since

$$f(0) = 1 \quad \text{and} \quad \lim_{x\to\infty} f(x) = 0$$

we deduce that

$$A = 1 \quad \text{and} \quad B = 0.$$

Hence,

$$f(x) = \left(\frac{1}{3}\right)^x.$$

Thus,

$$f(2) = \frac{1}{9}.$$

NOTE: Since the random walk is not symmetric, it is not a martingale. So, we cannot apply the martingale method to solve this question.

In general, let p and q are the up and down probabilities of the random walk respectively. If $p \neq \frac{1}{2}$, $N > 0$ and $m \geq 0$, then

$$\mathbb{P}(\tau_0 < \tau_N | S_0 = m) = \frac{1 - \left(\frac{p}{q}\right)^{N-m}}{1 - \left(\frac{p}{q}\right)^N}$$

By taking

$$p = \frac{3}{4}, \quad q = \frac{1}{4} \quad \text{and} \quad N \to \infty,$$

we will have the same answer.

10. Applying the stars and bars method, the answer is (10+9) choose 9, which is 92378. https://en.wikipedia.org/wiki/Stars_and_bars_%28combinatorics%29?wprov=sfla1

11. First of all, it is easy to see that the probability that you are back at the origin after the 2n-th step is

$$\mathbb{P}(S_{2n} = (0,0)|S_0 = (0,0)) = \left(\frac{1}{4}\right)^{2n} \sum_{m=0}^{n} \frac{(2n)!}{(m!)^2[(n-m)!]^2} = \left(\frac{1}{2}\right)^{4n} \binom{2n}{n}^2$$

Recall the Stirling approximation that for large n,

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

It follows that

$$\begin{aligned}\binom{2n}{n} &= \frac{(2n)!}{(n!)^2} \\ &= \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{2\pi n\left(\frac{n}{e}\right)^{2n}} \\ &= \frac{2^{2n}}{\sqrt{\pi n}}\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{n=1}^{\infty} \mathbb{P}(S_{2n} = (0,0)|S_0 = (0,0)) &\approx \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{4n} \frac{2^{4n}}{\pi n} \\ &= \sum_{n=1}^{\infty} \frac{1}{\pi n} \\ &= \infty\end{aligned}$$

Hence,

$$\lim_{n\to\infty} \mathbb{P}(S_{2n} = (0,0)|S_0 = (0,0)) = 1.$$

12. We can interpret this question as a variant of Gambler's ruin Problem(Wiki). Let

$$S_0 = 1, \quad S_n = S_0 + X_1 + \cdots + X_n$$

where X_i is a Bernoulli's random variable with

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$$

and

$$\tau_N = \min\{n \geq 0 : S_n = N\},$$

Therefore,

$$\mathbb{P}(\tau_9 < \tau_0 | S_0 = 1) = \frac{1}{9}.$$

In general, if $p = q = \frac{1}{2}$ and there are N people, then the required probability

$$\mathbb{P}(\tau_{N-1} < \tau_0 | S_0 = m) = \frac{1}{N-1}.$$

In fact, the answer is independent of our location and the host. We just need to consider the case when one of our neighbors receives the coin.
After that, we just need to make sure our other neighbor receives the coin first. Then we are the winner. Then the problem is identical to this question.

13. By labelling all vertices with distance to the destination, they can be partitioned into 0,1,2,3. A sketch of a markov chain is shown below.

    It follows that

$$\mu_3 = 1 + \mu_2$$
$$\mu_2 = 1 + \frac{1}{3}\mu_3 + \frac{2}{3}\mu_1.$$
$$\mu_1 = 1 + \frac{2}{3}\mu_2$$

    Solving the equations above gives $\mu_3 = 1$

14. Traditional markov chain method: Observe that

$$\mu_s = 1 + \frac{1}{2}\mu_H + \frac{1}{2}\mu_s$$
$$\mu_H = 1 + \frac{1}{2}\mu_{HT} + \frac{1}{2}\mu_H$$
$$\mu_{HT} = 1 + \frac{1}{2}\mu_{HTT} + \frac{1}{2}\mu_H$$
$$\mu_{HTT} = 1 + \frac{1}{2}\mu_s$$

    Solving equations above leads to

$$\mu_s = 18.$$

    Martingale method:
    Suppose that at each time N a person arrives and bets 1 dollar on the Nth roll being H. If they win, they then bet 2 on T; if they win again, 4 on T; and if they win again, 8 on H. They stop betting as soon as they either lose once or win four bets in a row. Then the net winnings of all the bettors up to a given time is a martingale, because its change at any future time is a sum of mean-0 random variables and therefore itself has mean 0.

A single given bettor's net winnings will be 0 (before they start), -1 (if they ever lose, and forever thereafter), +1 (if they have just seen H), +3 (if they have just seen HT), +7 (if they have just seen HTT), or +15 (if they have just seen HTTH, and forever thereafter). Suppose the first HTTH happens when its final H is on turn n. Total winnings at this point are -1 from the first $n-4$ bettors, +15 from the one who started with the first H, -1 from the next, -1 from the next, and +1 from the last: so a total of

$$18 - n.$$

Hence the expected value of the total winnings when HTTH first occurs is $18 - t$ where t is the expected turn number when that happens.

By the Optional Stopping Theorem, which says that the expected value of a martingale at a stopping time equals its initial expected value. (The first time when we get HTTH is certainly a stopping time.) The initial expected value is 0 (again, all the individual bets have expectation 0) and therefore

$$18 - t = 0.$$

Thus,

$$t = 18.$$

Fun further exercise: consider applying the same reasoning to other sequences besides HTTH and figure out the general formula for the expected time to see them.

For a more theoretical approach, please refer to the paper (1980) A Martingale Approach to the Study of Occurrence of Sequence Pattern in Repeated Experiments.pdf

15. Let n_a and n_i be the number of tosses such that HHHH and HTTH first appear respectively. By using martingale approach, we have

$$E[(n_a - 4)(-1) + 15 + 7 + 3 + 1] = 0$$
$$E(n_a) = 30.$$

Similarly,

$$\mathbb{E}[(n_i - 4)(-1) + 15 + (-1) + (-1) + 1] = 0$$
$$\mathbb{E}(n_i) = 18.$$

Therefore,

$$\mathbb{E}(T_a - T_i) = 30 - 18 = 12$$

16. For general λ, we have

$$\begin{aligned}
\mathbb{P}(X \text{ is even}) &= \sum_{k=0}^{\infty} \mathbb{P}(X = 2k) \\
&= \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^{2k}}{(2k)!} \\
&= \frac{e^{-\lambda}}{2} \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \right) \\
&= \frac{e^{-\lambda}}{2} (e^{\lambda} + e^{-\lambda}) \\
&= \frac{1 + e^{-2\lambda}}{2}
\end{aligned}$$

Since $\lambda = 1,$ so

$$\mathbb{P}(X \text{ is even}) = \frac{1 + e^{-2}}{2}.$$

17. Note that if we start at (0,1), then there is only one path that goes to (5,6) and (11,0). The hard part of this question is to calculate the number of paths that do not go below x-axis.

    By the reflection principle, the total number of paths that do not go below x-axis is

$$\binom{10}{5} - \binom{10}{4} = \frac{1}{6}\binom{10}{5} = 42.$$

    Therefore, the required probability is

$$\frac{1}{42}.$$

18. This question is the duplicate from question 32 in the getting series category.