Some Miscellaneous Notes on the Book Frequently Asked Questions in Quantitative Finance

John L. Weatherwax*

Nov 16, 2004

Notes on Common Interview Questions

While the answers to these problems are given in Wilmott's FAQ book and thus these answers are strictly not needed, several of the steps stated in the solutions are not fully worked out and I wanted to understand these problems as well as possible. Below you will find the solutions that I came up with. These solutions may provide an alternative way to view these problems.

Russian Roulette

When we spin the chamber we have a $\frac{2}{6} = \frac{1}{3}$ chance of having a bullet in the position where it will be next fired (behind the hammer of the gun). If we have fired the gun once and no bullet went out then we must have landed in one of the four empty spots. Once of these spots is next to the remaining two bullets and if we have just rotated to that spot then the second shot will result in injury. Thus we have a $\frac{1}{4}$ chance of getting shot if we elect to fire the gun a second time. Since $\frac{1}{4} < \frac{1}{3}$ we should *not* spin the chamber. If one fires the second time and does not get hit one has a $\frac{1}{3}$ of a chance of getting hit on the third shot. At this point one has an equal probability of death if one spins the chamber or takes this third shot. If one fires a third time and does not get shot then the fourth shot has a $\frac{1}{2}$ of a probability of being fatal. If not bullet is fired after the fourth shot, the fifth shot certainly will.

Matching Birthdays

In order for n people to *not* have a birthday in common requires that each person have a distinct birthday. This means that once the first persons birthday is picked the second person has 364 choices for their birthday that won't overlap with the first persons birthday. Once the first two people have their birthdays specified the third person will have 363 choices for their non-intersecting birthday. This pattern continues until we have considered the nth person. Thus the probability that all n people have distinct birthdays (denoted by P_d) is

$$\begin{aligned} P_d &= 1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{363}{365} \cdots \frac{365-(n-1)}{365} \quad \text{for} \quad n \geq 1 \\ &= \frac{\prod_{k=1}^{n-1}(365-k)}{365^{n-1}} . \end{aligned} \tag{1}$$

Then the probability that at least two people have the same birthday is given by

$$1 - \frac{\prod_{k=1}^{n-1}(365-k)}{365^{n-1}}.$$

If we want this probability to be greater than 1/2 we require

$$\frac{\prod_{k=1}^{n-1}(365-k)}{365^{n-1}} < \frac{1}{2}.$$

In the MATLAB script `matching_bdays.m` we compute the expression on the left-hand-side and find that the smallest value of n that makes this expression true is $n = 23$. Note that one needs to numerically compute the left-hand-side in a reasonable manner otherwise numerical overflow can make the results incorrect.

Another Birthday Problem

To begin this problem lets consider you are the second person in line. Then the probability you win a free ticket is given by $\frac{1}{365}$ since in only one way can you have the same birthday as the person ahead of you in line. Now assume you are the third person in line. The probability we win in this case then is given by

$$\frac{364}{365}\left(\frac{2}{365}\right).$$

Where the first fraction is the probability the first two people don't have the same birthday and the second fraction represents the probability you have the same birthday as one of the two people in line ahead of you. In the general case, when we are the $n+1$st person in line the above two conditions must hold. Then in order to win a free ticket in this case you must have that

- None of the n people ahead of you in line have matching birthdays
- You have the same birthday as one of the n people ahead of you in line.

Note that the second condition is conditioned on the first condition. We know that the probability of the first condition is given by Equation (1), and the second condition has the probability given by

$$\frac{n}{365}.$$

Thus in total if you have n people a head of you (i.e. stand at position $n+1$) the probability you win is given by

$$\frac{n}{365^n}\prod_{k=1}^{n-1}(365-k) \quad \text{for} \quad n \geq 1.$$

This expression agrees with the above ones when we are in the second person in line ($n = 1$) and the third ($n = 2$). If we plot the above probability as a function of n we get the plot shown in Figure 1. In the MATLAB script `position_in_line.m` we compute the expression we win a free ticket given there are n people in line ahead of use and find that the value of n that makes this expression a maximum. We find that the best chances to win are when

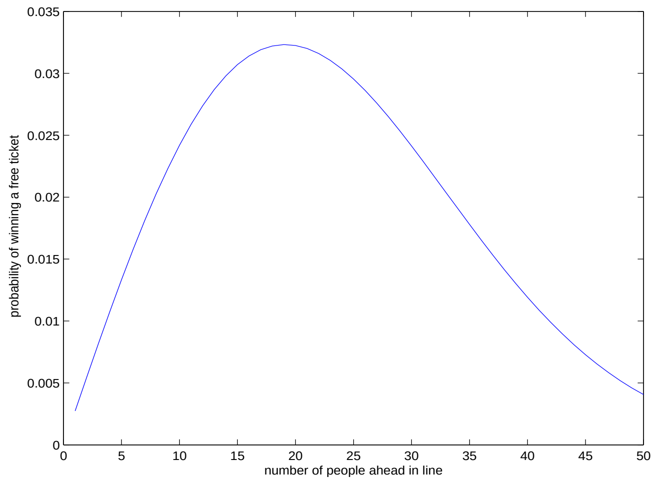

Figure 1: The plot that you win a free ticket given there are n people ahead of you in line.

Biased Coins

Let E_k be the event that the number of heads in the first k coins is odd. Let H_k be the even that the kth coin displays a head. Then we can calculate $P(E_k)$ by conditioning on what has happened in the past as follows

$$
\begin{aligned}
P(E_k) &= P(H_k|E_{k-1}^c)P(E_{k-1}^c) + P(H_k^c|E_{k-1}P(E_{k-1}) \\
&= \left(\frac{1}{2k+1}\right)(1 - P(E_{k-1})) + \frac{2k}{2k+1}P(E_{k-1})\,,
\end{aligned}
$$

when we use the information given in the problem statement. This can be simplified to

$$
P(E_k) = \frac{1}{2k+1} + \left(\frac{2k-1}{2k+1}\right)P(E_{k-1}) \quad \text{for} \quad k \geq 2\,.
$$

We see that this is a recursive relationship for $P(E_k)$ as a function of k. Since we know $P(E_1) = \frac{1}{3}$ by iterating the above we compute

$$
\begin{aligned}
P(E_2) &= \frac{1}{5} + \frac{3}{5}\left(\frac{1}{3}\right) = \frac{2}{5} \\
P(E_3) &= \frac{1}{7} + \frac{5}{7}\left(\frac{2}{5}\right) = \frac{3}{7} \\
&\vdots \\
P(E_k) &= \frac{k}{2k+1}\,.
\end{aligned}
$$

Two Heads

We interpret the statement "wait on average" to mean the *expected* number of flips before we get n heads. Let H_n be the expected number of flips to get n heads. Then using the conditional expectation formula by conditioning on the outcome of the next flip, we have

The first expression $\frac{1}{2}(1+H_{n-1})$ is the conditional expectation given that we get a head on the first flip, while the second expression is the conditional expectation given that we get a tail on this flip and need to start over. Thus we can solve for H_n in terms of H_{n-1} to get

$$H_n = 2 + 2H_{n-1}\,.$$

To evaluate this recurrence we can compute H_1 or the expected number of flips needed to get 1 head. With probability of 1/2 we get a head in the first flip. If we flip a tail then we start over. Thus we get

$$H_1 = \frac{1}{2} + \frac{1}{2}(1 + H_1)\,.$$

Solving for H_1 we get $H_1 = 2$. Then we find

$$\begin{aligned}
H_2 &= 2 + 2H_1 = 2 + 2^2 \\
H_3 &= 2 + 2^2 + 2^3 \\
H_4 &= 2 + 2^2 + 2^3 + 2^4 \\
&\vdots \\
H_n &= 2 + 2^2 + 2^3 + \cdots + 2^n = \sum_{k=1}^{n} 2^k = \frac{1-2^{n+1}}{1-2} - 1 = 2^{n+1} - 2\,.
\end{aligned}$$

For the case asked about $n = 2$ and we find $H_2 = 2^3 - 2 = 6$.

Balls in a Bag

First let W_k be the event that there are k white balls in the bag after the initial filling of the bag with balls using the method specified. Next let D (for drawn) be the event that 10 draws (with replacement) give all white balls. Then we want to evaluate $P(W_{10}|D)$. Using Bayes' rule we can write this probability as

$$P(W_{10}|D) = \frac{P(D|W_{10})P(W_{10})}{P(D)} = \frac{P(D|W_{10})P(W_{10})}{\sum_{k=0}^{10} P(D|W_k)P(W_k)}\,.$$

Lets now compute the expressions we need to evaluate the above. We have that W_k is a binomial random variable with probability of success $\frac{1}{2}$ and 10 trials or

$$P(W_k) = \binom{10}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{10-k} = \binom{10}{k}\left(\frac{1}{2}\right)^{10} \quad \text{for} \quad 0 \le k \le 10\,.$$

Next we have for $P(D|W_k)$ that

$$
\begin{aligned}
P(D|W_0) &= 0 \\
P(D|W_1) &= \left(\frac{1}{10}\right)^{10} \\
P(D|W_2) &= \left(\frac{2}{10}\right)^{10} \\
&\vdots \\
P(D|W_k) &= \left(\frac{k}{10}\right)^{10} \\
&\vdots \\
P(D|W_{10}) &= 1\,.
\end{aligned}
$$

Using these we can evaluate $P(W_{10}|D)$ and find

$$
P(W_{10}|D) = \frac{\left(\frac{1}{2}\right)^{10}}{\sum_{k=0}^{10}\left(\frac{k}{10}\right)^{10}\binom{10}{k}\left(\frac{1}{2}\right)^{10}} = \frac{1}{\sum_{k=0}^{10}\left(\frac{k}{10}\right)^{10}\binom{10}{k}}
$$

In the MATLAB/Octave script `balls_in_a_bag.m` we evaluate the above sum to find a probability of 0.070190.

Sums of Uniform Random Variables

Lets first consider the probability that the sum of n uniform random variables x_i is larger than 1, or $P\left(\sum_{k=1}^{n} x_k > 1\right)$. We can write this probability as

$$
1 - P\left(\sum_{k=1}^{n} x_k \leq 1\right),
$$

and thus if we can evaluate $P\left(\sum_{k=1}^n x_k \le 1\right)$ we can evaluate the probability of interest. To evaluate this later probability it helps to explicitly consider some cases for n. Note that

$$
\begin{aligned}
P(x_1 \le 1) &= 1 \\
P(x_1 + x_2 \le 1) &= \int_{x_1=0}^{1} \int_{x_2=0}^{1-x_1} 1 dx_2 dx_1 = \int_{x_1=0}^{1} (1 - x_1) dx_1 = x_1 - \left.\frac{x_1^2}{2}\right|_0^1 = \frac{1}{2} \\
P\left(\sum_{k=1}^{3} x_k \le 1\right) &= \int_{x_1=0}^{1} \int_{x_2=0}^{1-x_1} \int_{x_3=0}^{1-x_1-x_2} dx_3 dx_2 dx_1 = \int_{x_1=0}^{1} \int_{x_2=0}^{1-x_1} (1 - x_1 - x_2) dx_2 dx_1 \\
&= \int_{x_1=0}^{1} (1 - x_1) x_2 - \left.\frac{1}{2} {x_2}^2\right|_0^{1-x_1} dx_2 \\
&= \int_{x_1=0}^{1} \left((1 - x_1)^2 - \frac{1}{2}(1 - x_1)^2\right) dx_1 = \int_{x_1=0}^{1} \frac{1}{2}(1 - x_1)^2 dx_1 \\
&= -\frac{1}{6} \left.(1 - x_1)^3\right|_{x_1=0}^{1} = -\frac{1}{6}(0 - 1) = \frac{1}{6} \\
P\left(\sum_{k=1}^{4} x_i \le 1\right) &= \int_{x_1=0}^{1} \int_{x_2=0}^{1-x_1} \int_{x_3=0}^{1-x_1-x_2} \int_{x_4=0}^{1-x_1-x_2-x_3} dx_4 dx_3 dx_2 dx_1 \\
&= \int_{x_1=0}^{1} \int_{x_2=0}^{1-x_1} \int_{x_3=0}^{1-x_1-x_2} (1 - x_1 - x_2 - x_3) dx_3 dx_2 dx_1 \\
&= \int_{x_1=0}^{1} \int_{x_2=0}^{1-x_1} \left[(1 - x_1 - x_2) x_3 - \left.\frac{1}{2} {x_3}^2\right|_{x_3=0}^{1-x_1-x_2} dx_2 dx_1 \right. \\
&= -\int_{x_1=0}^{1} \int_{x_2=0}^{1-x_1} \frac{(1 - x_1 - x_2)^2}{2} dx_2 dx_1 \\
&= -\int_{x_1=0}^{1} \left.\frac{(1 - x_1 - x_2)^3}{6}\right|_{x_2=0}^{1-x_1} dx_1 = \int_{x_1=0}^{1} \frac{(1 - x_1)^3}{6} dx_1 = \frac{1}{24}\,.
\end{aligned}
$$

In general it looks like the pattern is given by

$$
P\left(\sum_{k=1}^{n} x_i \le 1\right) = \frac{1}{n!}\,,
$$

so

$$
P\left(\sum_{k=1}^{n} x_i > 1\right) = 1 - \frac{1}{n!}\,.
$$

We can use the above result to answer the question posed by this problem as follows. Note that the event $\sum_{k=1}^n x_i > 1$ can be decomposed into the following two mutually exclusive events depending on whether the sum of the first $n-1$ terms is greater than 1 or symbolically as

$$
\left(\sum_{k=1}^{n} x_i > 1\right) = \left(\sum_{k=1}^{n} x_i > 1\,, \sum_{k=1}^{n-1} x_i > 1\right) \cup \left(\sum_{k=1}^{n} x_i > 1\,, \sum_{k=1}^{n-1} x_i \le 1\right).
$$

The *second* event above is the one that happens when we have to sum n terms. Thus since these two event are disjoint we have

$$
\begin{aligned}
P\left(\sum_{k=1}^{n} x_i > 1\,, \sum_{k=1}^{n-1} x_i \le 1\right) &= P\left(\sum_{k=1}^{n} x_i > 1\right) - P\left(\sum_{k=1}^{n} x_i > 1\,, \sum_{k=1}^{n-1} x_i > 1\right) \\
&= P\left(\sum_{k=1}^{n} x_i > 1\right) - P\left(\sum_{k=1}^{n-1} x_i > 1\right) \\
&= \frac{1}{(n-1)!} - \frac{1}{n!} = \frac{n! - (n-1)!}{n!(n-1)!} = \frac{(n-1)(n-1)!}{n!(n-1)!} \\
&= \frac{n-1}{n!}\,.
\end{aligned}
$$

Thus the expectation of n is then given by

$$
E[n] = \sum_{n=2}^{\infty} n\left(\frac{n-1}{n!}\right) = \sum_{n=2}^{\infty} \frac{1}{(n-2)!} = \sum_{n=0}^{\infty} \frac{1}{n!} = e^1\,.
$$

Minimum and Maximum Correlation

From the given numbers we can construct the correlation matrix

$$
\left[\begin{array}{ccc} 1 & \rho_{xy} & \rho_{xz} \\ \rho_{xy} & 1 & \rho_{yz} \\ \rho_{xz} & \rho_{yz} & 1 \end{array}\right].
$$

Since this matrix must be positive definite the determinant of it must be positive. When we expand the determinant of the above using minors about the first column of the above we get

$$
1(1-\rho_{yz}^2) - \rho_{xy}(\rho_{xy} - \rho_{xz}\rho_{yz}) + \rho_{xz}(\rho_{xy}\rho_{yz} - \rho_{xz}) > 0\,.
$$

We will use this expression and solve for the unknown ρ_{xz}. Expanding all terms in the above gives for ρ_{xz} the quadratic expression for

$$
\rho_{xz}^2 - 2\rho_{xy}\rho_{yz}\rho_{xz} - (1 - \rho_{yz}^2 - \rho_{xy}^2) < 0\,.
$$

When we complete the square by adding and subtracting $\rho_{xy}^2\rho_{yz}^2$ to the left-hand-side we find

$$
(\rho_{xz} - \rho_{xy}\rho_{yz})^2 < \rho_{xy}^2\rho_{yz}^2 + (1 - \rho_{yz}^2 - \rho_{xy}^2) = (1-\rho_{xy}^2)(1-\rho_{yz}^2)\,.
$$

Thus

$$
|\rho_{xz} - \rho_{xy}\rho_{yz}| < \sqrt{(1-\rho_{xy}^2)(1-\rho_{yz}^2)}
$$

or

$$
\rho_{xy}\rho_{yz} - \sqrt{(1-\rho_{xy}^2)(1-\rho_{yz}^2)} < \rho_{xz} < \rho_{xy}\rho_{yz} + \sqrt{(1-\rho_{xy}^2)(1-\rho_{yz}^2)}\,.
$$

When we use the given information that $\rho_{xy} = 0.9$ and $\rho_{yz} = 0.8$ and evaluate the above bounds we find $0.458 < \rho_{xz} < 0.981$.

Airforce One

In this problem, if at the beginning, GW happens to sit (randomly) in the first seat everyone will fall into their correct seat and the last person to sit down will sit in their correct seat. If GW sits in the wrong seat say n then it is only when person n tries to sit down and finds their seat taken that another seat must be selected. For example, if GW sits in seat 90 then presidents $2, 3, \cdots, 89$ will all sit in their correct seat (since they are empty) and the 90th person will need to find a new seat. In this case however the 90th person can only select from the seats 1 or $91, 92, \cdots, 100$. The number of seats that the 90th person can thus select from is much smaller than the original 100 that GW could choose from. This second part, where the nth person has to randomly find a new place to sit is just like the first part of the problem, where GW enters the plane, but over a much smaller set of seats. For example in one way, if the nth person selects the first seat # 1 then everyone that follows person n will have their seat empty and the last person will sit in their assigned seat. This discussion motivates a recursive solution to this problem. Normally, with a recursive solution one has to start the recursion with a few Point cases. Let P_n be the probability the last person entering the plane sits in their correct seat. Then we have $P_1 = 1$,

$$P_2 = \frac{1}{2},$$

since only if GW sits in his own seat will the last person have his seat empty. Now for P_3 if GW sits in his seat (with probability $1/3$) then everyone else will find themselves in their correct seat. If GW sits in seat 2 the probability the last person ends in their seat is P_2. If GW sits in seat 3 the probability the last person ends in their seat is 0. Thus we get

$$P_3 = \frac{1}{3}1 + \frac{2}{3}\left(\frac{1}{2}P_2\right) = \frac{1}{2}.$$

Using the same logic we have

$$P_4 = \frac{1}{4} + \frac{3}{4}\left(\frac{1}{3}P_3 + \frac{1}{3}P_2 + \frac{1}{3}0\right) = \frac{1}{2}.$$

Computing P_5 for completeness we have

$$P_5 = \frac{1}{5} + \frac{4}{5}\left(\frac{1}{4}P_4 + \frac{1}{4}P_3 + \frac{1}{4}P_2 + \frac{1}{4}0\right) = \frac{1}{2},$$

when we put the previous expressions for P_k in. In general it looks like the recursion relationship is

$$P_n = \frac{1}{n} + \left(1 - \frac{1}{n}\right)\sum_{k=2}^{n}\frac{1}{n-1}P_k = \frac{1}{n} + \frac{1}{n}\sum_{k=2}^{n-1}P_k.$$

From the few cases we computed by hand we can hypothesize that $P_k = \frac{1}{2}$ and verify this using the above. We find that

$$P_n = \frac{1}{n} + \frac{1}{n}\frac{n-2}{2} = \frac{1}{2}.$$

Hit and Run Taxi

Let G be the event that a taxi is green and B the event that a taxi is blue. From the problem statement we have $P(G) = 0.85$ and $P(B) = 0.15$. Let $\hat{G}$ be the event the observer says the taxi is green while $\hat{B}$ is the event that observer says that the taxi is blue. Then we have

$$
\begin{aligned}
P(\hat{G}|G) = 0.8 \quad &\text{and} \quad P(\hat{G}|B) = 0.2 \\
P(\hat{B}|G) = 0.2 \quad &\text{and} \quad P(\hat{B}|B) = 0.8 \,.
\end{aligned}
$$

The question to answer is then what is the value of $P(B|\hat{B})$. Using Bayes' rule we have that

$$
\begin{aligned}
P(B|\hat{B}) &= \frac{P(\hat{B}|B)P(B)}{P(\hat{B})} = \frac{P(\hat{B}|B)P(B)}{P(\hat{B}|B)P(B) + P(\hat{B}|G)P(G)} \\
&= \frac{0.8(0.15)}{0.8(0.15) + 0.2(0.85)} = 0.4137 \,.
\end{aligned}
$$

Annual Returns

Let u be the percent we win when we are correct (here it is given by 0.5) and let d be the percent we loose when we are incorrect (in this problem it is 0.5). Assume that we play this game for L days. If during play we have n days where we win and $L - n$ days where we loose, we will have a return on our portfolio given by

$$
(1+u)^n (1-d)^{L-n} \,.
$$

If p is the probability we win on a given day then to have n winning days out of L is given by a binomial random variable and has a value given by

$$
\binom{L}{n} p^n (1-p)^{L-n} \,.
$$

For the trader to be ahead after the L days requires

$$
(1+u)^n (1-d)^{L-n} > 1 \,,
$$

or taking logarithms

$$
n \ln(1+u) + (L-n)\ln(1-d) > 0 \,.
$$

Solving for n we get

$$
n > -\frac{L \ln(1-d)}{\ln\left(\frac{1+u}{1-d}\right)} \,.
$$

Since n is a positive integer we must take the largest integer greater than or equal to the right-hand-side of the above expression, lets call that number n^* and note that it depends on the values of u, d, and L. If the trader wins on n^* or more days the trader will end positive. The probability that he ends ahead is then given by

$$
P_{\text{win}}(u, d, p, L) = \sum^{L} \binom{L}{n} p^n (1-p)^{L-n} \,.
$$

Using the numbers given for this problem of $u = 0.5$, $d = 0.5$, $p = 0.6$, and $L = 260$ evaluating the above gives $n^* = 165$ and then $P_{\text{win}} = 0.14$.

If we want to find the value for L such that P_{win} is the largest we can consider the above expression a function of L. When we evaluate P_{win} as a function of L for $1 \leq L \leq 500$, we find that the largest value of P_{win} happens when $L = 3$ with a value of $P_{\text{win}} = 0.648$. These calculations are done in the R script `ann_return.R`.

Dice Game

We can easily decide when to stop playing this game. Assume that we have x and play the game. Then the expected winning in this case are given by

$$1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) - x\left(\frac{1}{6}\right) = \frac{1}{6}(15 - x)\,.$$

Thus the expected winnings will only be positive if $x < 15$. The optimal stopping rule is to play until you have 15 or more dollars and then stop. When you play optimally the largest number one could have (and decide to keep playing) is 14. Then with one more play you could end up with $14 + \{1, 2, 3, 4, 5\} = \{15, 16, 17, 18, 19\}$ dollars. To evaluate the expected winnings under optimal play we can perform a Monte Carlo simulation. To do this we run this game under the stated optimal play and count the number of times we end with each of the above outcomes. Let N be the number of times we play this game and $n_{15}, n_{16}, n_{17}, n_{18}, n_{19}$ be the number of times (out of N) we get each of the above outcomes. Then the expectation of our win is given by

$$15\left(\frac{n_{15}}{N}\right) + 16\left(\frac{n_{16}}{N}\right) + 17\left(\frac{n_{17}}{N}\right) + 18\left(\frac{n_{18}}{N}\right) + 19\left(\frac{n_{19}}{N}\right)\,.$$

In the MATLAB script `dice_game.m` we implement this Monte Carlo simulation. There we find the probabilities of getting values n_i for $15 \leq i \leq 19$ given by

```
0.3450    0.2658    0.1999    0.1259    0.0634
```

and the expected winnings given by 16.2969.

100 kg of Berries

We start with 0.99 percent of 100 kilograms as water or 99 kilograms of water. The remaining 1 kilogram is pure berries. From this 99 kilograms of water some evaporates such that now there is x kilograms of water remaining. We are also told that after this evaporation the percent of water is 0.98 thus

$$\frac{x}{1 + x} = 0.98\,.$$

Solving for x in the above gives $x = 49$. Thus the berries plus water weights $49 + 1 = 50$ kilograms.

Urban Planning

Consider one "urban plan" where the cities are located at the four corners $(\pm l, \pm l)$ and

$\{(-l,+l),(+l,-l)\}$. The total length of the roads for this plan (since the side of the square is of length $2l$)

$$2\sqrt{4l^2+4l^2} = 2\sqrt{8l^2} = 4l\sqrt{2} = 5.65l\,.$$

This is to be compared with the "urban plan" where we place two intersections located at $(\pm a, 0)$ in such a way that we connect the intersection at $(-a,0)$ to the corners $\{(-l,+l),(-l,-l)\}$ and the intersection at $(0,+a)$ to the corners $\{(+l,-l),(+l,+l)\}$. The the total length of the roads in this construction is given by

$$4\sqrt{(l-a)^2+l^2} + 2a \quad \text{for} \quad 0 \le a \le l\,.$$

We desire to pick a such that the above expression is as small as possible. Denoting this expression as $J(a)$, to find the minimum of $J(a)$ we set its derivative equal to zero and solve for a. The derivative set equal to zero gives

$$J'(a) = 2 + \frac{2(2(l-a)(-1))}{\sqrt{(l-a)^2+l^2}} = 0\,.$$

When we solve for a in the above expression we get

$$|l-a| = \frac{l}{\sqrt{3}} \quad \text{or} \quad a = l \pm \frac{l}{\sqrt{3}} = l\left(1 \pm \frac{1}{\sqrt{3}}\right)\,.$$

We need to take the minus sign in the above expression to ensure that $0 \le a \le l$. Taking the second derivative of $J(a)$ to determine the sufficient condition for a minimum we find

$$\begin{aligned} J''(a) &= \frac{4}{\sqrt{(l-a)^2+l^2}} + \frac{4(l-a)2(l-a)}{2((l-a)^2+l^2)^{3/2}} \\ &= \frac{4}{\sqrt{(l-a)^2+l^2}} + \frac{4(l-a)^2}{((l-a)^2+l^2)^{3/2}} \ge 0\,, \end{aligned}$$

since each term in the sum is. Thus the value of a computed above is a minimum. The total distance then when $a = l\left(1-\frac{1}{\sqrt{3}}\right)$ is (note that $l - a = \frac{1}{\sqrt{3}}$) given by

$$4\sqrt{\frac{l^2}{3}+l^2} + 2l\left(1-\frac{1}{\sqrt{3}}\right) = 2l(1+\sqrt{3})\,,$$

when we simplify. If we compute the above expression we see that it is equal to $5.46l$, which is *less* than the value of $5.65l$ we computed under the first design.

Closer to the edge or the center

Consider a square centered at $(0,0)$ with sides of length $2a$, so that the corners of the square are located at $(\pm a, \pm a)$. Let $\mathcal{R}$ be the points in the square that are closer to the center than the edge. Then the probability that if we randomly sample from the points in this square we select a point that is closer to the center and therefore in $\mathcal{R}$ is equal to the area of the region $\mathcal{R}$ over the area of the square or $4a^2$. To find the area of $\mathcal{R}$, consider just the first quadrant (the problem is symmetric in the other quadrants), then the boundary between points that are closer to the center and points that are closer to the edges is given by

Let us further now consider the points (x, y) *below* the line $y = x$, where we know that $a - x < a - y$. Then the boundary described above becomes

$$\sqrt{x^2 + y^2} = a - x \, .$$

When we square this and solve for x in terms of y we have

$$x = \frac{1}{2a}(a^2 - y^2) \, .$$

We can also solve this expression for y as a function of x to get

$$y = \sqrt{a^2 - 2ax} \quad \text{in the region where} \quad y < x \, . \tag{2}$$

If we consider the case where the points (x, y) are *above* the line $y = x$ then $a - y < a - x$ and the boundary becomes

$$\sqrt{x^2 + y^2} = a - y \, .$$

When we square this and solve for y in terms of x to get

$$y = \frac{1}{2a}(a^2 - x^2) \quad \text{in the region where} \quad y > x \, . \tag{3}$$

Now these two curves meet in the first quadrant at the point where the parabolas given by Equation 2 and 3 meet the line $x = y$. Taking Equation 2 we have that the x value where these curves meet is the point where

$$x = \sqrt{a^2 - 2ax} \, . \tag{4}$$

On squaring both sides and rearranging we get

$$x^2 + 2ax - a^2 = 0 \, .$$

The quadratic formula gives then for the solution the values of

$$x = \frac{-2a \pm \sqrt{4a^2 + 4a^2}}{2} = a(-1 \pm \sqrt{2}) \, .$$

We must take the positive sign in the above expression so that $x < a$, and we call that solution x^*. Finally the radicand in Equation 2 will vanish when

$$a^2 - 2ax = 0 \quad \text{or} \quad x = \frac{a}{2} \, .$$

Thus in summary as we move in x from 0 to $\frac{a}{2}$ the boundary of the region $\mathcal{R}$ is given by the following curve

$$y(x) = \begin{cases} \frac{1}{2a}(a^2 - x^2) & 0 < x < x^* \\ \sqrt{a^2 - 2ax} & x^* < x < \frac{a}{2} \end{cases}$$

When we plot this curve using the script `closer_to_the_center.m` we get the result shown in Figure 2. To determine the area of the region $\mathcal{R}$ we need to integrate the above function

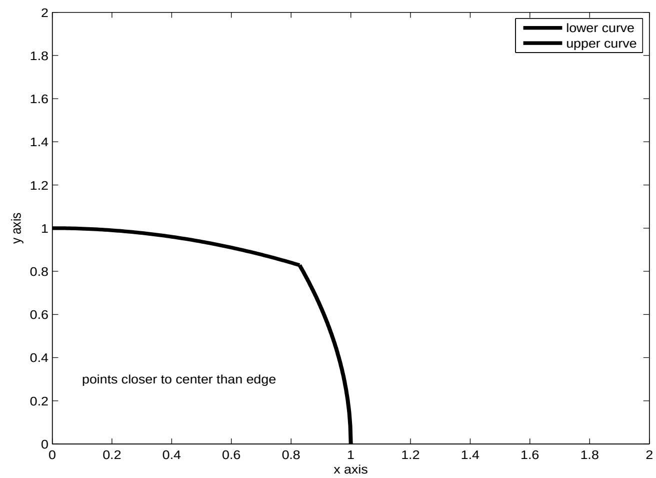

Figure 2: The region in the first quadrant where the the point (x, y) is closer to the center $(0, 0)$ than to the edge $x = a$ or $y = a$.

$y(x)$ with respect to x for x between 0 and $\frac{a}{2}$. This means that we need to calculate

$$A_{\mathcal{R}} = \int_{x=0}^{x^*} \frac{1}{2a}(a^2 - x^2)dx + \int_{x=x^*}^{\frac{a}{2}} \sqrt{a^2 - 2ax}dx\,.$$

The first integral is computed as

$$\begin{aligned}
\int_{x=0}^{x^*} \frac{1}{2a}(a^2 - x^2)dx &= \frac{1}{2a}\left(a^2 x - \frac{x^3}{3}\bigg|_0^{x^*}\right) = \frac{1}{2a}\left(a^2 x^* - \frac{1}{3}x^{*3}\right) \\
&= \frac{1}{2a}\left(a^3(-1+\sqrt{2}) - \frac{1}{3}a^3(-1+\sqrt{2})^3\right) \\
&= \frac{a^2}{2}(-1+\sqrt{2})\left(1 - \frac{1}{3}(-1+\sqrt{2})^2\right) \\
&= \frac{a^2}{6}(-1+\sqrt{2})\left(3 - (1 - 2\sqrt{2} + 2)\right) \\
&= \frac{a^2}{3}\sqrt{2}(-1+\sqrt{2})\,.
\end{aligned}$$

The second integral when we make the substitution $v = a^2 - 2ax$ is computed as

$$\begin{aligned}
\int_{x=x^*}^{\frac{a}{2}} \sqrt{a^2 - 2ax}dx &= \int_{v=a^2-2ax^*}^{0} \sqrt{v}\left(-\frac{dv}{2a}\right) \\
&= \left(\frac{1}{2a}\right)\left(\frac{2}{3}v^{3/2}\bigg|_0^{a^2-2ax^*}\right) = \frac{1}{3a}(a^2 - 2ax^*)^{3/2} \\
&= \frac{a^2}{3}(3 - 2\sqrt{2})^{3/2}\,,
\end{aligned}$$

when we use Equation 4 to simplify the expression $a^2 - 2ax^*$. To finish this problem we need quadrants, and then divide the result by the area of the square $4a^2$. When we do that we get

$$\frac{\sqrt{2}}{3}(-1+\sqrt{2}) + \frac{1}{3}(3-2\sqrt{2})^{3/2}\,.$$

Warning: This is a different result than given in the book. There they give

$$(-1+\sqrt{2})^2 + \frac{2}{3}(3-2\sqrt{2})^{3/2}\,.$$

If anyone sees an error in my logic or my algebra in the work here please let me know.

Snowflake

Lets assume that on iteration $k = 0$ (no iteration) we start with an equilateral triangle of side l. On the first iteration, $k = 1$, we take each side and attach another equilateral triangle of size $\frac{l}{3}$. On the second iteration, $k = 2$, we have segment lengths of $\frac{l}{3^2}$. In general, the length of each segment at iteration k is given by

$$l_k = \frac{l}{3^k} \quad \text{for} \quad k \geq 0\,.$$

Next let N_k be the number of segments at iteration k on this snowflake. Then from the initial triangle we have $N_0 = 3$. When $k = 1$ we have

$$N_1 = 4N_0 = 3 \cdot 4\,,$$

since each of the initial 3 sides ends up with 4 sides after the first iteration. After the second iteration each of the N_1 sides ends up with 4 sides so we get

$$N_2 = 4N_1 = 4^2 N_0 = 3 \cdot 4^2\,.$$

In general it looks like the number of sides with length l_k is then given by

$$N_k = 3 \cdot 4^k\,.$$

This makes it easy to calculate the perimeter P_k at iteration k of the structure generated in this manner. We have

$$P_k = N_k l_k = 3 \cdot 4^k \left(\frac{l}{3^k}\right) = 3l\left(\frac{4}{3}\right)^k \to +\infty\,,$$

as $k \to \infty$.

We now seek to determine the area of the snowflake. To do that we first need to evaluate the area of an equilateral triangle. To evaluate that we drop a vertical line from the apex of the original triangle to its base. Since this creates two right triangles with a hypotenuse of length l from the original triangle using the Pythagorean theorem gives

$$l^2 = h^2 + \left(\frac{l}{2}\right)^2\,,$$

as an equation for the height. Solving for h gives

Using this we find that the area of an equilateral triangle with length l is given by

$$A(l) = \frac{1}{2}\text{base} \cdot \text{height} = \frac{\sqrt{3}}{4} l\,.$$

This is A_0. Lets now derive the area of the snowflake after the first iteration. In that case we have the area of the original triangle *plus* the area of N_0 more equilateral triangles each with edge length $\frac{l}{3}$ giving a total area of

$$A_1 = A(l) + N_0 A\left(\frac{l}{3}\right) = A(l_0) + N_0 A(l_1)\,.$$

On the second iteration to each of the N_1 segments we add an equilateral triangle each with side length l_2 to give the total area A_2 of

$$A_2 = A(l_0) + N_0 A(l_1) + N_1 A(l_2)\,.$$

In general then the area after k iterations is given by

$$A_k = A(l) + \sum_{k=0}^{k-1} N_k A\left(l_{k+1}\right)\,.$$

when we take the limit of $k \to$ we get

$$\begin{aligned} A_\infty &= A(l) + \sum_{k=0}^{\infty} (3 \cdot 4^k)\left(\frac{\sqrt{3}}{4}\frac{l^2}{3^{2k+2}}\right) \\ &= A(l) + \frac{3\sqrt{3}}{4 \cdot 5} l^2 = \frac{2\sqrt{3}}{5} l^2\,, \end{aligned}$$

when we simplify. Notice that this is finite.

The doors

To start this problem it is easier to consider what will happen to a specific door, say door number 15. In that case the first person will open this door, the 3rd person will close it, the 5th person will open it again, and the 15th person will close it. These numbers are the pairs of factors of 15. For example $15 = 1 \cdot 15$ and $15 = 3 \cdot 5$. Thus for every door if it has an even number of factors (like 15 above) it will end *closed*. If a number has an odd number of factors like the number 9 with factors 1, 3, and 9 it will end *open*. Thus the number of open doors at the end of this process are the number of numbers n for $1 \le n \le 100$ that have an odd number of factors. These are the squares. These are the numbers

$$1\,, 4\,, 9\,, 16\,, 25\,, 36\,, 49\,, 64\,, 81\,, 100\,.$$

Since there are 10 of these numbers there will be 10 open doors and $100 - 10 = 90$ closed doors.

Two thirds of the average

From the problem specification we are asked to predict $\frac{2}{3}$ the average of n draws from a uniform distribution. Since as n gets large this sample statistics mean must tend to the population statistics or

$$\frac{2}{3}\overline{X} \to \frac{2}{3}(50) = 33.33\,.$$

Thus we expect this number (as an integer) to be 33. The trick to this problem is to recognize that if everyone follows this same set of reasoning, they will *all* pick this number meaning the average of the selections will be 33 and so $\frac{2}{3}$'s of this will be 22. Thus we should pick 22. This logic continues in that if everyone else follows this line of reasoning they will all pick 22 and thus the average will be 22 and we should pick

$$\frac{2}{3}(22) = 14.666 \approx 15\,.$$

If we continue this reasoning eventually we will pick 0. The sequence of number pick is given by

$$x_k = \text{round}\left(\frac{2}{3}x_{k-1}\right)\,,$$

for $k \geq 1$ with $x_0 = 50$. The sequence of numbers we get by iterating this function is

$$33\,,22\,,15\,,10\,,7\,,5\,,3\,,2\,,1\,.$$

These are computed in the MATLAB file `two_thirds_the_average.m`.

Compensation

The trick to this problem is to recall that an average can be computed in many different ways, one of which is recursively. Thus one of the quants creates a random number and adds it to his salary. He then tells this number another quant who adds his salary to that number and reports the sum to another quant. Each quant does this until each quant has added his salary to the "running sum". The first quant then takes this number, subtracts the random number he initially added in and then divide by the number of quants at the dinner. The resulting number is the average.

Einstein's brainteaser

We start this problem by reading the problem statement and then enumerating the various possibilities for each of the various unique items to each house

- colors: red, white, green, yellow, blue
- nationalities: English, Swedish, Danish, Norwegian, German
- drinks: tea, coffee, milk, beer, water
- pets: dog, bird, cats, fish, horses
- cigarette types: Pall Mall, Dunhill, Blends, Bluemaster, Prince

Since there are 5 houses and 5 items to get placed within each house without any constraints there are $(5!)^5 = 24883200000$ possible item placements. The statements given provide some information about the location of the above items that reduces the number of possible arrangements. In the `python` code `einsteins_brainteaser.py` I wrote a code to explicitly enumerate all possible item placements and to skip the ones that don't satisfy the given constraints. When we do that I'm finding multiple solutions to this problem. For example I find the following solutions

Color	Nationality	Drink	Pet	Cigarette
green	norwegian	coffee	bird	pallmall
blue	german	water	cat	prince
red	english	milk	horse	blends
yellow	danish	tea	fish	dunhill
white	swedish	beer	dog	bluemaster

Color	Nationality	Drink	Pet	Cigarette
green	norwegian	coffee	bird	pallmall
blue	german	water	fish	prince
red	english	milk	horse	blends
yellow	danish	tea	cat	dunhill
white	swedish	beer	dog	bluemaster

Color	Nationality	Drink	Pet	Cigarette
green	norwegian	coffee	bird	pallmall
blue	german	water	cat	prince
white	swedish	milk	dog	blends
red	english	beer	horse	bluemaster
yellow	danish	tea	fish	dunhill

Color	Nationality	Drink	Pet	Cigarette
green	norwegian	coffee	bird	pallmall
blue	german	water	cat	prince
white	swedish	milk	dog	blends
yellow	danish	tea	fish	dunhill
red	english	beer	horse	bluemaster

Color	Nationality	Drink	Pet	Cigarette
green	norwegian	coffee	bird	pallmall
blue	german	water	fish	prince
white	swedish	milk	dog	blends
yellow	danish	tea	cat	dunhill
red	english	beer	horse	bluemaster

Color	Nationality	Drink	Pet	Cigarette
green	norwegian	coffee	fish	blends
blue	german	water	cat	prince
yellow	swedish	milk	dog	dunhill
red	english	beer	horse	bluemaster
white	danish	tea	bird	pallmall

Color	Nationality	Drink	Pet	Cigarette
yellow	norwegian	water	cat	dunhill
blue	danish	tea	horse	blends
red	english	milk	bird	pallmall
green	german	coffee	fish	prince
white	swedish	beer	dog	bluemaster

The houses are listed from top to bottom. I spot checked that all of these solutions satisfy the given conditions of the problem and they seemed to. If anyone sees any solution that does not satisfy the constraints (or a bug in the code) please let me know.

Gender ratio

Assume that we have an arbitrary couple with no children. Then with probability $1/2$ they have a boy as their first child. With probability $(1/2)^2$ they have a girl and then a boy. With probability $(1/2)^3$ they have two girls and then a boy etc. Then the expected number of girls under this requirement is

$$\sum_{k=1}^{\infty}(k-1)\left(\frac{1}{2}\right)^k = \sum_{k=0}^{\infty}k\left(\frac{1}{2}\right)^{k+1} = 1\,.$$

This is an example of a *negative binomial* random variable. This means that on average under this policy we have 1 boy and 1 girl and thus the population ratio is *equal*. We evaluate sums like the above using the fact that

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}\,,$$

and then taking the derivative of both sides with respect to r to introduce a k coefficient into the summand and then simplifying.

Covering an chessboards with dominoes

A 64 square chessboard has 32 white squares and 32 black squares. When you remove two squares on any diagonal they must be of the same color either black or white. If you remove two black squares then there will be 30 black squares and 32 white squares. When placing a single domino will cover both a black and a white square we have to place at least 30 dominoes to cover all of the remaining black squares. This will leave 2 white squares uncovered. These remaining squares cannot be covered with dominoes.

Aircraft armor

Since the aircrafts that return were not shot down one would want to place the armor at the places where the returning aircrafts were *not* shot. This rather obvious conclusion and the interesting story behind it is discussed in [1].

Ages of three children

If we let the three children's ages be denoted by a, b, and c. The we assume that $a \leq b \leq c$, where equality can happen if two of the children are twins. The the information that the product of the ages is 36 means that that we can specify a few possible ages for the children. In Table 1 we present the possible ages a, b, and c that have 36 as a product. We also present the sum of the three ages in the fourth column. The census taker is now told that the sum

a	b	c	total
1	1	36	38
1	2	18	21
1	3	12	16
1	4	9	14
1	6	6	13
2	2	9	13
2	3	6	11
3	3	4	10

Table 1: The potential ages for the three children.

of the ages is the address of the house next door. If we look at the sums of the three ages in Table 1 we see that in most cases knowing the sum of the three ages is enough to determine which assignment of (a, b, c) is true. For example, if we are told that the sum of the three ages is 21 for example then the assignment of ages must be $(1, 2, 18)$. Since in fact the census taker *cannot* determine the ages from the sum of the three numbers. This information must not be enough and the potential possible sets of ages must be

$$(1\,,6\,,6) \quad \text{or} \quad (2\,,2\,,9)\,.$$

Since the censor taker is then informed that the oldest child is sleeping the second of the above two choices must be the children's ages.

Four Switches and a Lightbulb

For this problem we turn on switches one and two and wait some time, which if either of these were the switch that controlled the lightbulb would result in the light being on. After some time turn switch two off and turn switch three on. Then go into the room. Depending on how you find the bulb you can determine which switch works the light. If

- the bulb is warm and on then switch one works the light.
- the bulb is warm but off then switch two works the light.
- the bulb is cold but on then switch three works the light.

Turnover

Assume we select n cards from the 52 total that are available and place them in a *new* pile (the cards not selected will be in the *old* pile). Lets assume that x of these cards are from the 19 face up cards and thus $n - x$ of these cards are from the 33 face down cards (x is unknown to us). Then we can break up the 19 face up cards and the 33 face down cards depending on x and n in the new pile as follows

$$\begin{aligned} 19 &= x + (19 - x) \quad \text{face up} \\ 33 &= n - x + (33 - (n - x)) \quad \text{face down} \,. \end{aligned}$$

Where the above means that from the 19 original face up cards x are in the new pile and $19 - x$ are in the old pile. From the 33 original face down cards $n - x$ are in the new pile and $33 - (n - x)$ are in the old pile. If we next take the n cards that represent the new pile and then flip them all over we will now thus have

$$\begin{aligned} x &\quad \text{face down} \\ n - x &\quad \text{face up} \end{aligned}$$

To have the same number of face up cards in both piles would require

$$n - x = 19 - x \,,$$

since there are $19 - x$ face up cards in the old pile. The above equation is satisfied for *any* value of x as long as $n = 19$. Thus we can obtain the desired condition by taking any 19 cards and flipping them over.

Muddy faces

We are told that at least one child has mud on his face. To begin, consider the case where only one child has mud on his face. Then when that child looks around he will see no other children with mud on his face. He then knows that he is the only child with mud and should raise his hand. Now consider the case where there are two children A and B with mud on their faces. When one of the children, A, with mud on his face looks around in that case he will see another child, B, with mud on their face. He then reasons: If that is the only child with mud on their face they should raise their hand as argued in the one child case above. When B does *not* immediately raise their hand we reason that B must see someone else with mud on their face. This person must be A. Thus both A and B must raise their hands. Logic along these same lines indicate that for an arbitrary number of children with mud on their faces they will recognize this and raise their hands. One problem with this logic is that the children must make all of these conclusions at the same time. In the case with just two children with mud as argued above if between A and B, the child B realizes that he is one of the two children with mud and raises his hand before child A realizes this, then child A might conclude that they are in the one child with mud case where B is the only child with mud and not raise his hand. This behavior would break the logic above.

The Oracle at Delphi

At the end of each day we can be either long or flat (neutral). At the end of each day, if morning to get flat. If we are flat we have to decide whether we will buy the next morning or do nothing (stay flat). Let L_i be the maximum possible value of our profit at the end of day i if we are long then. Let N_i be the maximum possible value of our profit at the end of day i if we are flat at the end of day i. We will *end up long* at the end of day i if on day $i-1$ we

- are long and stay long or
- are flat and go long.

The optimal value for L_i is then given by which ever action is better or

$$L_i = \max(L_{i-1}(1+R_i), N_{i-1}(1+R_i)(1-k))\,.$$

We will *end up flat* at the end of day i if on day $i-1$ we

- are long and decide to sell or
- are flat and decide to do nothing.

The optimal value for N_i is then given by

$$N_i = \max((1-k)L_{i-1}, N_{i-1})\,.$$

We can then start these iterations from $N_0 = 1$ and $L_0 = 0$ and determine N_i and L_i for any i.

Miss Moneypenny

Lets assume that the procedure we will use to select our secretary is that we will interview some number (say m for $1 \le m \le n-1$) secretaries recording the *value*, v_m, of the best one observed during this initial screening. We then interview the remaining $n-m$ candidates and select the first one that has a value greater than the observed v_m or the last person interviewed if none do. We desire to compute the probability that we select the best secretary, or the person with the largest value, under this procedure. Let this event be denoted by B. Then this probability, $P(B)$, will obviously be a function of m. We will evaluate this probability by conditioning on the possible secretary we will select, say k for $m+1 \le k \le n$ under this procedure. Let the event that the best secretary is at location k be B_k and let S_k be the event that we select the kth secretary under this procedure. Then conditioning on the events B_k we have

$$P(B) = \sum_{k=m+1}^{n} P(S_k, B_k) = \sum_{k=m+1}^{n} P(S_k|B_k)P(B_k)\,. \tag{5}$$

Now $P(B_k) = \frac{1}{n}$, for all k, since a priori there is no expected location for the best candidate. Now lets evaluate $P(S_k|B_k)$ when $k = m+1$, or the probability that we select the $m+1$th person given that the $m+1$ is the best. Once we see the $m+1$st person they since that person is the best (over all people) they will have a score that is larger than the m initial people screened and will be picked according to the rules above. Thus we find that

Lets now consider how to evaluate $P(S_k|B_k)$ for $k = m+2$. Since the $k+2$nd person is the best over all secretaries we will only not select them if we mistakenly select the $k+1$st person instead. This we will not accept the $k+1$st person if they have a score that is less than the value v_m found during our initial screen. We can calculate this probability by observing that it is equal to the probability that when we sort the values of the first $m+1$ people the $m+1$st person is not largest in the sorted list. To not be the largest can happen in $\frac{m}{m+1}$ ways. Thus we find that

$$P(S_{k+2}|B_{k+2}) = \frac{m}{m+1} .$$

Lets now evaluate $P(S_{k+3}|B_{k+3})$. This is equal to the probability that we don't select either person $k+1$ or $k+2$ to be the secretary. This means that the people at spots $k+1$ and $k+2$ have value less than the largest of the first m people. This later event will happen when the most valuable person from the first $m+2$ comes from the first m people (and not the last two). This happens with probability $\frac{m}{m+2}$ and we have shown that

$$P(S_{k+3}|B_{k+3}) = \frac{m}{m+2} .$$

This pattern continues and we find

$$\begin{aligned} P(S_{n-1}|B_{n-1}) &= \frac{m}{n-2} \quad \text{and} \\ P(S_n|B_n) &= \frac{m}{n-1} , \end{aligned}$$

for the end points. Thus using these results and Equation 5 we have

$$\begin{aligned} P(B) &= \frac{1}{n} + \frac{m}{m+1}\left(\frac{1}{n}\right) + \frac{m}{m+2}\left(\frac{1}{n}\right) + \frac{m}{m+3}\left(\frac{1}{n}\right) + \cdots + \frac{m}{n-2}\left(\frac{1}{n}\right) + \frac{m}{n-1}\left(\frac{1}{n}\right) \\ &= \frac{m}{n}\left[\frac{1}{m} + \frac{1}{m+1} + \frac{1}{m+2} + \frac{1}{m+3} + \cdots + \frac{1}{n-2} + \frac{1}{n-1}\right] . \end{aligned}$$

This expression is valid for $1 \le m \le n-1$. This expression can be computed using the MATLAB/Octave function `eval_prob_best_selection_at_m.m`. Now for each value of n we want to pick the value of m such that $P(B)$ is largest. When $n = 2$, then m can only be 1, there is no decision to make and we have $P(B) = \frac{1}{2}$. If $n = 3$ then $m = 1$ or $m = 2$ and we find

$$\begin{aligned} P(B; n=3; m=1) &= \frac{1}{3}\left(1 + \frac{1}{2}\right) = \frac{1}{2} \\ P(B; n=3; m=2) &= \frac{2}{3}\left(\frac{1}{2}\right) = \frac{1}{3} , \end{aligned}$$

so we would take $m = 1$, to maximize our probability. If $n = 4$ we compute

$$\begin{aligned} P(B; n=4; m=1) &= \frac{1}{4}\left(1 + \frac{1}{2} + \frac{1}{3}\right) = 0.458 \\ P(B; n=4; m=2) &= \frac{2}{4}\left(\frac{1}{2} + \frac{1}{3}\right) = 0.416 \\ P(B; n=4; m=3) &= \frac{3}{4}\left(\frac{1}{3}\right) = 0.25 \end{aligned}$$

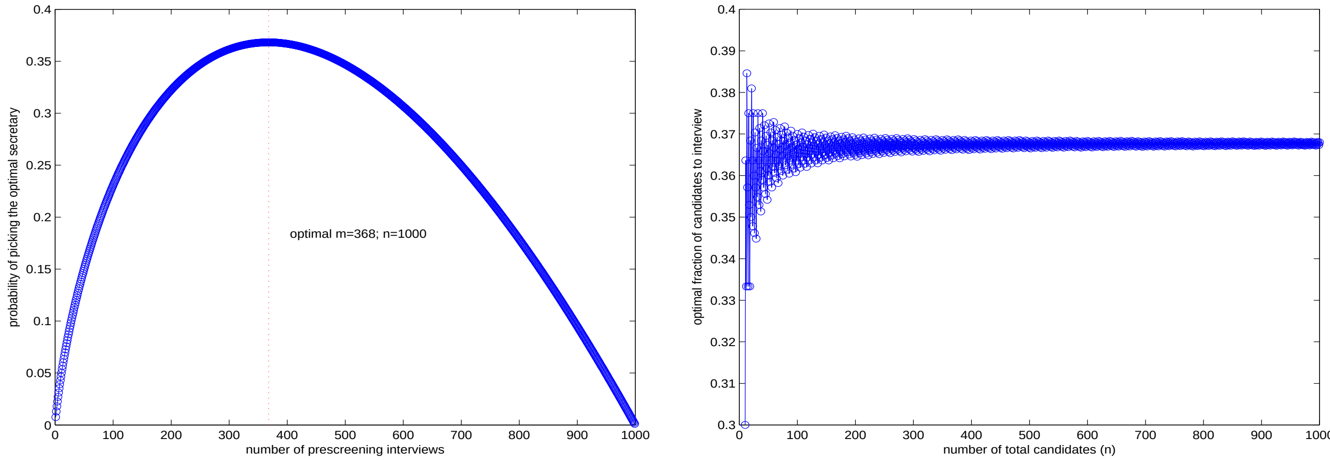

Figure 3: **Left:** For $n = 1000$ a plot of the probability that one selects the best (from n) secretary by interviewing the first m for $1 \leq m \leq n-1$ secretaries. **Right:** A plot of the *fraction* of the n secretaries (for $10 \leq n \leq 1000$) to screen before selecting the first one better than the ones initially interviewed. Notice the nice limiting value for this fraction.

Thus we would pick $m = 1$. In the MATLAB/Octave function `eval_prob_best_selection.m` given an input value for n we calculate all the possible probabilities of selecting the best secretary given that we screen the first m candidates for $1 \leq m \leq n-1$. It is interesting to calculate what the optimal *fraction* or $\frac{m_{\text{optimal}}}{n}$ of secretaries to interview before proceeding to select one. This experiment is performed in the MATLAB/Octave function `miss_moneypenny.m`. When that script is run it produces the results found in Figure 3. How the optimal fraction is greater than $\frac{1}{3}$.

Pirate puzzle

We will do a few iterations of this problem. Once you realize the general idea, starting with the last two pirates and working backwards to include more pirates, the general case is easier to reason. To begin with assume that we just have two pirates #9 and #10. In that case if #9 gives himself *any* gold then when he and #10 vote (assuming in this case that #10's vote carries more weight than #9 and he breaks any ties) pirate #10 will disagree with this allocation and vote #9 over. Thus the allocation must be 0 for pirate #9 and 1000 for pirate #10 or symbolically

$$(\#9, \#10) = (0, 1000)\,.$$

If there are three pirates then pirate #8 must create an allocation that will produced a majority vote test. Pirate #10 will never vote for any allocation that gives any gold to pirate #8 since he prefers the case with only two pirates where he gets all the gold. Pirate #9 will always vote for any allocation that #8 proposes since he has nothing to loose (if one wants to be sure of pirates #9's vote one could give him one gold piece). Thus pirate #8 will vote for the allocation that gives himself all the gold. This allocation is then

$$(\#8, \#9, \#10) = (1000, 0, 0)\,. \tag{6}$$

When we go to the four pirate case, pirate #7 must pick an allocation that gives himself the which #7 needs to produce to win the votes of his companions is one that when compared with allocation 6 will produce a majority vote in pirate #7's favor. If he gives both #9 and #10 one gold piece he can ensure their votes and get his desired allocation. This allocation then is

$$(\#7, \#8, \#9, \#10) = (998, 0, 1, 1)\,.$$

This pattern of deciding how to allocate the gold continues as we consider more and more pirates, where the allocation pirate #i must select must produce a better allocation for the remaining pirates as compared with the alternative that they would get if the reject #i's allocation.

Miscellaneous Interview Questions

What follows are a few miscellaneous interview questions I have seen over the years.

Bounding the distance between the mean and the median

Problem: Let X be a general one-dimensional random variable, $\tilde{\mu}$ its population median, μ its population mean, and σ^2 its population variance. Prove that

$$|\tilde{\mu} - \mu| \le \sigma\,.$$

Solution: We start with the following manipulation of the left-hand-side of the above

$$|\tilde{\mu} - \mu| = |E(\tilde{\mu} - X)| = \left|\int (\tilde{\mu} - x)p(x)dx\right| \le \int |\tilde{\mu} - x|p(x)dx\,. \tag{7}$$

Now recall that the median can be defined as the value that minimizes the L^1 error i.e. the median is the argument that obtains the minimum of the following function

$$E(c) = \int |c - x|p(x)dx\,.$$

Because of this, the integral $\int |\tilde{\mu} - x|p(x)dx$ gets *larger* (or stays the same) if we replace $\tilde{\mu}$ with the population mean μ or

$$\int |\tilde{\mu} - x|p(x)dx \le \int |\mu - x|p(x)dx\,. \tag{8}$$

Now use the Cauchy-Schwarz inequality in the form

$$|E(XY)| \le \sqrt{E[X^2]E[Y^2]}\,.$$

with the random variables $X = |\mu - X|$ and $Y = 1$ for then we have (in terms of integrals)

$$\left|\int |\mu - x|p(x)dx\right| = \int |\mu - x|p(x)dx \le \left(\int (\mu - x)^2 p(x)dx\right)^{1/2} = \sigma\,. \tag{9}$$

Thus using Equation 7 then 8 and finally 9 we have proven the desired result.

The expected number of loops when tying noodles together

Problem: A plate has N noodle in it. You randomly pick up two noodle ends and tie them together. You continue doing this until there are no more free noodle ends. What is the expected number of loops on the plate at the end?

Solution: Note that when we tie two noodle ends together we always eliminate a noodle from the total set of noodles are considering. This happens because

- If we happen to pick two ends that belong to the *same* noodle we create one loop and loose that individual noodle.
- If we happen to pick two noodle ends belonging to *different* noodles, when we tie these ends together we end up creating a single (longer) noodle but in this case no loops are created.

Let n be the number of noodles and l be the number of loops at any given time. Based on the above logic after we join two noodle ends we have the following two possible new values for the vector $\begin{bmatrix} n \\ l \end{bmatrix}$ depending on whether we attach two ends of the same noodle or not

$$\begin{bmatrix} n \\ l \end{bmatrix} \to \begin{bmatrix} n-1 \\ l+1 \end{bmatrix} \quad \text{with probability} \quad \frac{1}{n} \quad \text{or}$$
$$\begin{bmatrix} n \\ l \end{bmatrix} \to \begin{bmatrix} n-1 \\ l \end{bmatrix} \quad \text{with probability} \quad \frac{n-1}{n}\,.$$

To compute the expected number of loops let e_n be the expected number of loops given that we start with n noodles. Then from the above arguments we have that

$$e_n = (1 + e_{n-1})\left(\frac{1}{n}\right) + (0 + e_{n-1})\left(\frac{n-1}{n}\right)\,.$$

We can write the above in the simpler form

$$e_n - e_{n-1} = \frac{1}{n}\,.$$

We know that $e_1 = 1$ since only have one noodle in that case and will select it with probability one and tie its two ends. Letting $n = 2, 3, 4, \ldots$ into the above recursion relationship we get

$$\begin{aligned} e_2 &= 1 + \frac{1}{2} \\ e_3 &= 1 + \frac{1}{2} + \frac{1}{3} \\ e_4 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\,. \end{aligned}$$

In general it then looks like

$$e_n = \sum_{k=1}^{n} \frac{1}{k}\,.$$

References

[1] S. L. Savage. *The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty*. Wiley, June 2009.