Problem 1

Alice is trying to communicate with Bob, by sending a message (encoded in binary) across a channel.
(a) Suppose for this part that she sends only one bit (a 0 or 1), with equal probabilities. If she sends a 0, there is a 5% chance of an error occurring, resulting in Bob receiving a 1; if she sends a 1, there is a 10% chance of an error occurring, resulting in Bob receiving a 0. Given that Bob receives a 1, what is the probability that Alice actually sent a 1?
(b) To reduce the chance of miscommunication, Alice and Bob decide to use a repetition code. Again Alice wants to convey a 0 or a 1, but this time she repeats it two more times, so that she sends 000 to convey 0 and 111 to convey 1. Bob will decode the message by going with what the majority of the bits were. Assume that the error probabilities are as in (a), with error events for different bits independent of each other. Given that Bob receives 110, what is the probability that Alice intended to convey a 1?

Solution:

(a) Suppose event A represents: Alice sends 1, event A^c represents: Alice sends 0, event B represents: Bob receives 1 event B^c represents: Bob receives 0.

$$
\begin{aligned}
P(A|B) = & \frac{P(B|A)P(A)}{p(B)} \\
= & \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\
= & \frac{0.5 \cdot 0.9}{0.5 \cdot 0.05 + 0.5 \cdot 0.9} \\
\approx & \ 0.9474.
\end{aligned}
$$

(b) Suppose event A_1 represents Alice sends 111, event A_0 represents Alice sends 000, and event B represents Bob receives 110.

$$
\begin{aligned}
P(A_1|B) = & \frac{P(B|A_1)P(A_1)}{P(B)} \\
= & \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} \\
= & \frac{0.9^2 \cdot 0.1 \cdot 0.5}{0.9^2 \cdot 0.1 \cdot 0.5 + 0.05^2 \cdot 0.95 \cdot 0.5} \\
\approx & \ 0.9715.
\end{aligned}
$$

Problem 2

Fred decides to take a series of n tests, to diagnose whether he has a certain disease (any individual test is not perfectly reliable, so he hopes to reduce his uncertainty by taking multiple tests). Let D be the event that he has the disease, $p = P(D)$ be the prior probability that he has the disease, and $q = 1 - p$. Let T_j be the event that he tests positive on the jth test.

(a) Assume for this part that the test results are conditionally independent given Fred's disease status. Let $a = P(T_j \mid D)$ and $b = P(T_j \mid D^c)$, where a and b don't depend on the jth test. Find the posterior probability that Fred has the disease, given that he tests positive on all n of the n tests.

(b) Suppose that Fred tests positive on all n tests. However, some people have a certain gene that makes them always test positive. Let G be the event that Fred has the gene. Assume that $P(G) = 1/2$ and that D and G are independent. If Fred does not have the gene, then the test results are conditionally independent given his disease status. Let $a_0 = P(T_j \mid D, G^c)$ and $b_0 = P(T_j \mid D^c, G^c)$, where a_0 and b_0 don't depend on j. Find the posterior probability that Fred has the disease, given that he tests positive on all n of the tests.

Solution

(a) We need to calculate $P\left(D \mid \cap_{j=1}^n T_j\right)$. Use Bayes' formula, LOTP and conditional independence of T_j (if D is given) to obtain following

$$\begin{aligned}
P\left(D \mid \cap_{j=1}^n T_j\right) &= \frac{P(D)P\left(\cap_{j=1}^n T_j \mid D\right)}{P\left(\cap_{j=1}^n T_j\right)} = \frac{P(D)\prod_{j=1}^n P\left(T_j \mid D\right)}{P\left(\cap_{j=1}^n T_j \mid D\right)P(D) + P\left(\cap_{j=1}^n T_j \mid D^c\right)P\left(D^c\right)} \\
&= \frac{P(D)\prod_{j=1}^n P\left(T_j \mid D\right)}{P(D)\prod_{j=1}^n P\left(T_j \mid D\right) + P\left(D^c\right)\prod_{j=1}^n P\left(T_j \mid D^c\right)} \\
&= \frac{p \cdot a^n}{p \cdot a^n + (1-p)b^n}.
\end{aligned}$$

(b) Again, we need to calculate $P\left(D \mid \cap_{j=1}^n T_j\right)$. Same as in (a), obtain that is

$$P\left(D \mid \cap_{j=1}^n T_j\right) = \frac{P(D)P\left(\cap_{j=1}^n T_j \mid D\right)}{P(D)P\left(\cap_{j=1}^n T_j \mid D\right) + P\left(D^c\right)P\left(\cap_{j=1}^n T_j \mid D^c\right)}.$$

Use LOTP and independence of G and D to calculate

$$\begin{aligned}
P\left(\cap_{j=1}^n T_j \mid D\right) &= P\left(\cap_{j=1}^n T_j \mid D, G\right)P(G \mid D) + P\left(\cap_{j=1}^n T_j \mid D, G^c\right)P\left(G^c \mid D\right) \\
&= P\left(\cap_{j=1}^n T_j \mid D, G\right)P(G) + P\left(\cap_{j=1}^n T_j \mid D, G^c\right)P\left(G^c\right) \\
&= \frac{1}{2}P\left(\cap_{j=1}^n T_j \mid D, G\right) + \frac{1}{2}P\left(\cap_{j=1}^n T_j \mid D, G^c\right) \\
&= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \prod_{j=1}^n P\left(T_j \mid D, G^c\right) \\
&= \frac{1}{2} + \frac{1}{2}{a_0}^n.
\end{aligned}$$

Similarly we get that

$$P\left(\cap_{j=1}^n T_j \mid D^c\right) = \frac{1}{2} + \frac{1}{2}{b_0}^n.$$

Plug all these information in to obtain that

$$P\left(D \mid \cap_{j=1}^n T_j\right) = \frac{p\left(\frac{1}{2} + \frac{1}{2}{a_0}^n\right)}{p\left(\frac{1}{2} + \frac{1}{2}{a_0}^n\right) + (1-p)\left(\frac{1}{2} + \frac{1}{2}{b_0}^n\right)}.$$

Problem 3

We want to design a spam filter for email. A major strategy is to find phrases that are much more likely to appear in a spam email than in a no spam email. In that exercise,we only consider one such phrase:free money. More realistically, suppose that we have created a list of 100 words or phrases that are much more likely to be used in spam than in non-spam. Let W_j be the event that an email contains the jth word or phrase on the list. Let

$$p = P(spam), p_j = P(W_j|spam), r_j = P(W_j|not \quad spam)$$

where spam is shorthand for the event that the email is spam.
Assume that $W_1, ..., W_{100}$ are conditionally independent given that the email is spam, and also conditionally independent given that it is not spam. A method for classifying emails (or other objects) based on this kind of assumption is called a naive Bayes classifier. (Here naive refers to the fact that the conditional independence is a strong assumption, not to Bayes being naive. The assumption may or may not be realistic, but naive Bayes classifiers sometimes work well in practice even if the assumption is not realistic.)
Under this assumption we know, for example, that

$$P(W_1, W_2, W_3^c, W_4^c, ..., W_{100}^c|spam) = p_1p_2(1-p_3)(1-p_4)...(1-p_{100}).$$

Without the naive Bayes assumption, there would be vastly more statistical and computational difficulties since we would need to consider $2100 \approx 1.31030$ events of the form$A1 \cap A2... \cap A100$ with each A_j equal to either W_j orW_j^c . A new email has just arrived, and it include the 23rd, 64th, and 65th words or phrases on the list(but not the other 97). So we want to compute

$$P(spam|W_1^c, ..., W_{22}^c, W_{23}, W_{24}^c, ..., W_{63}^c, W_{64}, W_{65}, W_{66}^c, ..., W_{100}^c).$$

Note that we need to condition on all the evidence, not just the fact that $W_{23} \cap W_{64} \cap W_{65}$ occurred. Find the condition probability that the new email is spam (in terms of p and the p_j and r_j).

Solution:
Let W represents $W_1^c, ..., W_{22}^c, W_{23}, W_{24}^c, ..., W_{63}^c, W_{64}, W_{65}, W_{66}^c, ..., W_{100}^c$. Then using Baye's formula, we have:

$$p(spam|W) = \frac{p(spam) \cdot p(W|spam)}{p(spam) \cdot p(W|spam) + p(spam^c) \cdot p(W|spam^c)},$$

Because:
$p(W|spam) = p(W_1^c|spam) \cdot ... \cdot p(W_{22}^c|spam) \cdot p(W_{23}|spam) \cdot p(W_{24}^c|spam) \cdot ... \cdot p(W_{64}|spam) \cdot p(W_{65}|spam) \cdot p(W_{66}^c|spam) \cdot ... \cdot p(W_{100}^c|spam) = (1-p_1)...(1-p_{22})p_{23}(1-p_{24})...p_{64}p_{65}(1-p_{66})...(1-p_{100})$,

$p(W|spam^c) = p(W_1^c|spam^c) \cdot ... \cdot p(W_{22}^c|spam^c) \cdot p(W_{23}|spam^c) \cdot p(W_{24}^c|spam^c) \cdot ... \cdot p(W_{64}|spam^c) \cdot p(W_{65}|spam^c) \cdot p(W_{66}^c|spam^c) \cdot ... \cdot p(W_{100}^c|spam^c) = (1-r_1)...(1-r_{22})r_{23}(1-r_{24})...r_{64}r_{65}(1-r_{66})...(1-r_{100})$,
the equation above can be written as

$$\begin{aligned} &P(spam|W) \\ =&\frac{p(1-p_1)...(1-p_{22})p_{23}(1-p_{24})...p_{64}p_{65}(1-p_{66})(1-p_{100})}{p(1-p_1)...p_{23}(1-p_{24})...p_{65}(1-p_{66})(1-p_{100}) + (1-p)(1-r_1)...(1-r_{22})r_{23}(1-r_{24})...r_{64}r_{65}(1-r_{66})(1-r_{100})} \\ =&\frac{P}{P+Q}, \end{aligned}$$

where,
$P = p(1-p_1)...(1-p_{22})p_{23}(1-p_{24})...p_{64}p_{65}(1-p_{66})...(1-p_{100})$,
$Q = (1-p)(1-r_1)...(1-r_{22})r_{23}(1-r_{24})...r_{64}r_{65}(1-r_{66})...(1-r_{100})$.

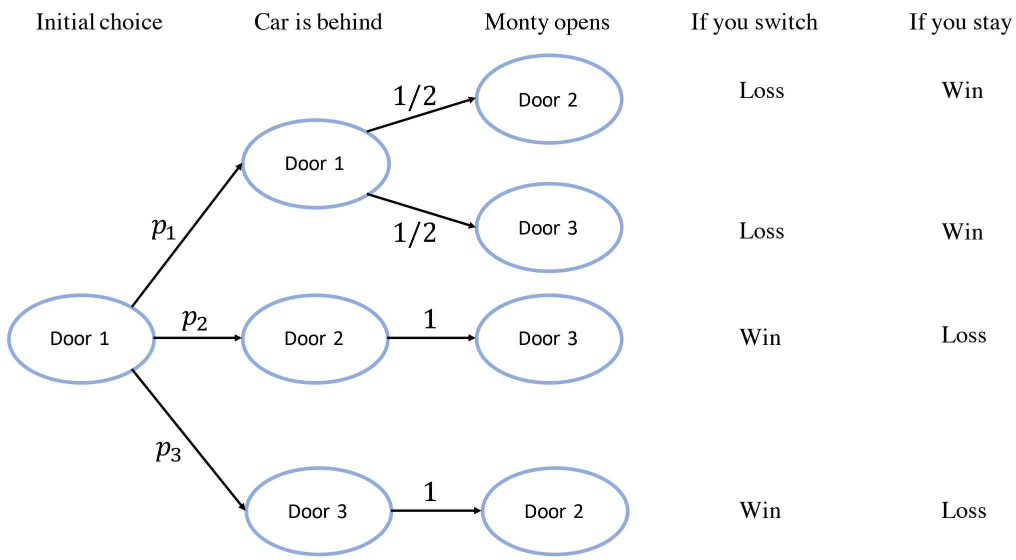
Problem 4

In Monty Hall problem, now suppose the car is not placed randomly with equal probability behind the three doors. Instead, the car is behind door one with probability p_1, behind door two with probability p_2, and behind door three with probability p_3. Here $p_1 + p_2 + p_3 = 1$ and $p_1 \geq p_2 \geq p_3 > 0$. You are to choose one of the three doors, after which Monty will open a door he knows to conceal a goat. Monty always chooses randomly with equal probability among his options in those cases where your initial choice is correct. What strategy should you follow?

Solution:
We define

1. $P_i^{\text{switch}}(\text{win})$ as the probability of winning if "choosing door i first and then switching".
2. $P_i^{\text{stay}}(\text{win})$ as the probability of winning if "choosing door i first and then sticking to initial choice".

When we choose door 1 first, outcomes are shown in the following diagram:



Therefore, we have

$$P_1^{\text{switch}}(\text{win}) = p_2 + p_3,$$
$$P_1^{\text{stay}}(\text{win}) = p_1.$$

Similarly, when we choose door 2 first,

$$P_2^{\text{switch}}(\text{win}) = p_1 + p_3,$$
$$P_2^{\text{stay}}(\text{win}) = p_2.$$

When we choose door 3 first,

$$P_3^{\text{switch}}(\text{win}) = p_1 + p_2,$$
$$P_3^{\text{stay}}(\text{win}) = p_3.$$

Remind that $p_1 \geq p_2 \geq p_3 > 0$. It's not difficult to find that $P_3^{\text{switch}}(\text{win})$ has the maximum winning probability. Hence, in this case, the optimal strategy should be: **choose door 3 first, then switch to the unopened door** after Monty opens some door. Intuitively, we are actually choosing the most unlikely door at the beginning, and then switch to the surviving door, which is the most likely one to be our target.

Problem 5

Consider the Monty Hall problem, except that Monty enjoys opening door 2 more than he enjoys opening door 3, and if he has a choice between opening these two doors, he opens door 2 with probability p, where $1/2 \leq p \leq 1$. To recap: there are three doors, behind one of which there is a car (which you want), and behind the other two of which there are goats (which you don't want). Initially, all possibilities are equally likely for where the car is. You choose a door, which for concreteness we assume is door 1.
(a) Find the unconditional probability that the strategy of always switching succeeds (unconditional in the sense that we do not condition on which of doors 2 or 3 Monty opens).
(b) Find the probability that the strategy of always switching succeeds, given that Monty opens door 2.
(c) Find the probability that the strategy of always switching succeeds, given that Monty opens door 3.

Solution

(a) Let C_j be the event that the car is hidden behind door j and let W be the event that we win using the switching strategy. Using the law of total probability, we can find the unconditional probability of winning in the same way as in class:

$$\begin{aligned} P(W) &= P\left(W \mid C_1\right) P\left(C_1\right)+P\left(W \mid C_2\right) P\left(C_2\right)+P\left(W \mid C_3\right) P\left(C_3\right) \\ &= 0 \cdot 1/3+1 \cdot 1/3+1 \cdot 1/3=2/3. \end{aligned}$$

(b) A tree method works well here (delete the paths which are no longer relevant after the conditioning, and reweight the remaining values by dividing by their sum), or we can use Bayes' rule and the law of total probability (as below).
Let D_i be the event that Monty opens Door i. Note that we are looking for $P\left(W \mid D_2\right)$, which is the same as $P\left(C_3 \mid D_2\right)$ as we first choose Door 1 and then switch to Door 3. By Bayes' rule and the law of total probability,

$$\begin{aligned} P\left(C_3 \mid D_2\right) &= \frac{P\left(D_2 \mid C_3\right) P\left(C_3\right)}{P\left(D_2\right)} \\ &= \frac{P\left(D_2 \mid C_3\right) P\left(C_3\right)}{P\left(D_2 \mid C_1\right) P\left(C_1\right)+P\left(D_2 \mid C_2\right) P\left(C_2\right)+P\left(D_2 \mid C_3\right) P\left(C_3\right)} \\ &= \frac{1 \cdot 1/3}{p \cdot 1/3+0 \cdot 1/3+1 \cdot 1/3} \\ &= \frac{1}{1+p}. \end{aligned}$$

(c) The structure of the problem is the same as part (b) (except for the condition that $p \geq 1/2$, which was no needed above). Imagine repainting doors 2 and 3, reversing which is called which. By part (b) with $1-p$ in place of p, $P\left(C_2 \mid D_3\right)=\frac{1}{1+(1-p)}=\frac{1}{2-p}$.

Problem 6

A/B testing is a form of randomized experiment that is used by many companies to learn about how customers will react to different treatments. For example, a company may want to see how users will respond to a new feature on their website (compared with how users respond to the current version of the website) or compare two different advertisements. As the name suggests, two different treatments, Treatment A and Treatment B, are being studied. Users arrive one by one, and upon arrival are randomly assigned to one of the two treatments. The trial for each user is classified as "success" (e.g., the user made a purchase) or "failure". The probability that the n-th user receives Treatment A is allowed to depend on the outcomes for the previous users. This set-up is known as a two-armed bandit. Many algorithms for how to randomize the treatment assignments have been studied. Here is an especially simple (but fickle) algorithm, called a "stay-with-a-winner" procedure:
(i) Randomly assign the first user to Treatment A or Treatment B, with equal probabilities.
(ii) If the trial for the n-th user is a success, stay with the same treatment for the $(n+1)$-st user; otherwise, switch to the other treatment for the $(n+1)$-st user.
Let a be the probability of success for Treatment A, and b be the probability of success for Treatment B. Assume that $a \neq b$, but that a and b are unknown (which is why the test is needed). Let p_n be the probability of success on the n-th trial and a_n be the probability that Treatment A is assigned on the n-th trial (using the above algorithm).
(a) Show that

$$p_n = (a-b)a_n + b,\ a_{n+1} = (a+b-1)a_n + 1 - b.$$

(b) Use the results from (a) to show that p_{n+1} satisfies the following recursive equation:

$$p_{n+1} = (a+b-1)p_n + a + b - 2ab.$$

(c) Use the result from (b) to find the long-run probability of success for this algorithm, $\lim_{n\to+\infty} p_n$, assuming that this limit exists.
Solution:
(a)

$$\begin{aligned}p_n =& P\{n\text{-th trial succeed}\}\\ =& P\{n\text{-th trial succeed}|\text{Treatment A is assigned on the } n\text{-th trial}\}P\{\text{Treatment A is assigned on the } n\text{-th trial}\}\\ &+ P\{n\text{-th trial succeed}|\text{Treatment B is assigned on the } n\text{-th trial}\}P\{\text{Treatment B is assigned on the } n\text{-th trial}\}\\ =& a\cdot a_n + b\cdot(1-a_n)\\ =& (a-b)a_n + b.\end{aligned}$$

$$\begin{aligned}a_{n+1} =& P\{\text{Treatment A is assigned on the } (n+1)\text{-th trial}\}\\ =& P\{\text{Treatment A is assigned on the } n\text{-th trial}\}P\{n\text{-th trial succeed}\}\\ &+ P\{\text{Treatment B is assigned on the } n\text{-th trial}\}P\{n\text{-th trial failed}\}\\ =& a_n a + (1-a_n)(1-b)\\ =& (a+b-1)a_n + 1 - b.\end{aligned}$$

(b)

$$\begin{aligned}
p_{n+1} =& P\{(n+1)\text{-th trial succeed}\} \\
=& P\{(n+1)\text{-th trial succeed}|\text{Treatment A is assigned on the } (n+1)\text{-th trial}\} \\
& \cdot P\{\text{Treatment A is assigned on the } (n+1)\text{-th trial}\} \\
& + P\{(n+1)\text{-th trial succeed}|\text{Treatment B is assigned on the } (n+1)\text{th trial}\} \\
& \cdot P\{\text{Treatment B is assigned on the } (n+1)\text{-th trial}\} \\
=& a a_{n+1} + b(1 - a_{n+1}) \\
=& (a-b)[(a+b-1)a_n + 1 - b] + b \\
=& (a+b-1)p_n + a + b - 2ab.
\end{aligned}$$

(c) It is denoted that $\lim\limits_{n\to+\infty} p_n = p$. Since the limitation exists, we have

$$p = (a+b-1)p + a + b - 2ab,$$

that is,

$$p = \frac{a+b-2ab}{2-a-b}.$$

Problem 7

(Optional: Challenging Problem) In a game hosted by Monty, you are presented with n identical doors, with $n \geq 3$. You select one, but do not open it. Monty now opens a door he knows to conceal a goat, and gives you the option of switching doors. After making your choice, Monty reveals another goat. He again gives you the option of switching. This process continues until only two doors remain (your current choice, and one other unopened door). You make your final choice, and receive whatever is behind your door. We assume throughout that Monty always chooses randomly from among the goat-concealing doors when more than one such door remains in play.

(a) What strategy maximizes your chances of success?

(b) Under such strategy, what is your winning probability?

Solution

See paper "Optimal strategies for the progressive monty hall problem."

(a) The optimal strategy for the progressive Monty Hall problem is to change doors once at the moment when only two doors remain in play.

(b) In the following we show

Let S be a given strategy and denote by a_n the probability of winning with S at the moment when n doors remain in play. We now make our initial door choice. If b_n denotes the probability of winning with S given that our current choice conceals a goat, and c_n denotes the probability of winning given that our current choice conceals the car, then a_n, b_n and c_n are related via the following equation:

$$a_n = \left(\frac{n-1}{n}\right) b_n + \left(\frac{1}{n}\right) c_n.$$

Let us assume, then, that our strategy calls for us to switch doors a total of k times with $k \leq n-2$. We will assume that we choose our new door randomly from the available options each time we switch. Denote by $\{m_i\}_{i=1}^{k}$ the number of doors remaining when we make the $k-i+1$-st switch. We have $3 \leq m_1 < m_2 < \cdots < m_k \leq n$.

For any integer j, the manner in which the probabilities b_j and c_j are related to b_{j-1} and c_{j-1} will depend on whether or not we switch at the moment when j doors remain. If we switch, then the probabilities are related in the manner described by the following equations:

$$\begin{aligned} b_n &= \left(\frac{n-3}{n-2}\right) b_{n-1} + \left(\frac{1}{n-2}\right) c_{n-1} \\ c_n &= b_{n-1}. \end{aligned}$$

If we do not switch, then the probabilities do not change. To see this, note that our probability of winning by sticking with our present door is equal to the probability that it conceals the car. A straightforward argument using Bayes' theorem shows that this probability can not change so long as we maintain this door as our selection. Consequently, our probability of winning can change only at those moments of the game when we decide to switch doors.

It follows that we have

$$b_{m_i} = \left(\frac{m_i - 3}{m_i - 2}\right) b_{m_{i-1}} + \left(\frac{1}{m_i - 2}\right) c_{m_{i-1}} \quad \text{and} \quad c_{m_i} = b_{m_{i-1}},$$

for all $1 \leq i \leq k$. For any subscript $j \neq m_i$ for any i, we have $b_j = b_{j-1}$ and $c_j = c_{j-1}$.

Problem 7 continued on next page...

To simplify the notation, we define $\beta_i = b_{m_i}$ and $q_i = m_i - 2$. This leads to

$$\beta_i = \left(\frac{q_i - 1}{q_i}\right)\beta_{i-1} + \left(\frac{1}{q_i}\right)\beta_{i-2}, \tag{1}$$

for $i = 1, 2, \ldots, k$. Note that we have the initial conditions $\beta_{-1} = 1$ and $\beta_0 = 0$. The probability of winning given there are n doors and a given set of k door changes is thus

$$a_n = \left(\frac{n-1}{n}\right)\beta_k + \left(\frac{1}{n}\right)\beta_{k-1}. \tag{2}$$

To solve (2), set $\gamma_i = \beta_i - \beta_{i-1}$. Subtracting β_{i-1} from both sides of equation (1) then leads to

$$\gamma_i = \frac{-\gamma_{i-1}}{q_i},$$

with $\gamma_0 = -1$. Thus $\gamma_1 = 1/q_1, \gamma_2 = -1/\left(q_1 q_2\right), \gamma_3 = 1/\left(q_1 q_2 q_3\right)$, and in general

$$\gamma_j = \frac{(-1)^{j+1}}{\prod_{i=1}^{j} q_i}.$$

Since $\beta_i = \gamma_i + \beta_{i-1}$,

$$\beta_1 = \frac{1}{q_1}, \quad \beta_2 = \frac{1}{q_1} - \frac{1}{q_1 q_2}, \quad \beta_3 = \frac{1}{q_1} - \frac{1}{q_1 q_2} + \frac{1}{q_1 q_2 q_3},$$

and in general

$$\beta_j = \frac{1}{q_1} - \frac{1}{q_1 q_2} + \frac{1}{q_1 q_2 q_3} - \cdots + \frac{(-1)^{j+1}}{\prod_{i=1}^{j} q_i}.$$

Finally, substitution back in (2) gives us

$$a_n = \frac{1}{q_1} - \frac{1}{q_1 q_2} + \frac{1}{q_1 q_2 q_3} - \cdots + \frac{(-1)^{k+1}}{\prod_{i=1}^{k} q_i} - \frac{(-1)^{k+1}}{n \prod_{i=1}^{k} q_i}.$$

Consider the following cases:

(1) If we never change doors, then $k = 0, \beta_0 = 0$ and $\beta_{-1} = 1$. It follows that $a_n = 1/n$.

(2) If we change doors exactly once then $k = 1$ and

$$a_n = \frac{1}{q_1} - \frac{1}{n q_1} = \frac{n-1}{n q_1}.$$

The probability of winning is maximized, given our constraints on q_1, by choosing $q_1 = 1$, which is equivalent to $m_1 = 3$. This corresponds to switching at the last possible moment. Again, this makes sense. If you are only going to switch one time you should do so after Monty opens his final door. This strategy gives a probability of winning $a_n = (n-1)/n$, which approaches one as n increases.

(3) If we change doors twice, $k = 2$ and

$$a_n = \frac{1}{q_1} - \frac{1}{q_1 q_2} + \frac{1}{n q_1 q_2} = \frac{n\left(q_2 - 1\right) + 1}{n q_1 q_2}.$$

This is clearly maximized by choosing $q_1 = 1$ and $q_2 = n - 2$. The optimal strategy of changing doors immediately then leaving the final change to the last moment has probability $a_n = \left(n^2 - 3n - 1\right) / \left(n^2 - 2n\right)$, which also approaches one as n increases. Note, however, that for a given number of doors n, the best one change winning probability is better than that for two changes.

Problem 7 continued on next page...

(4) If we change doors three times, then $k = 3$ and

$$a_n = \frac{1}{q_1} - \frac{1}{q_1 q_2} + \frac{1}{q_1 q_2 q_3} - \frac{1}{n q_1 q_2 q_3}.$$

Since $1/q_1$ is a common factor to all terms, we see that a_n is maximized by minimizing q_1. That is, we set $q_1 = 1$. Determining the appropriate values of q_2 and q_3 is trickier. The value of a_n increases with q_2 (suggesting that q_2 should be maximized), but decreases with q_3 (suggesting that q_3 should be minimized). We must balance these considerations with the fact that $q_2 < q_3$. This is accomplished by setting $q_3 = q_2 + 1$. With this substitution, it is straightforward to show that we should take $q_2 = n - 3$ and $q_3 = n - 2$. This corresponds to making the first two switches immediately, and then waiting until the end to make the third. This leads to a probability of success of

$$a_n = \frac{n^3 - 6n^2 + 9n - 1}{n(n-2)(n-3)}.$$

This approaches one as n increases, but is nonetheless a lower chance of success than in the two switch strategy.

Therefore, we begin by assuming that we change doors exactly k times, where $k > 1$. We have already shown that the $k = 1$ case (one switch at the end) is superior to the best $k = 2$ or 3 cases. For $k > 3$ and using the optimum strategy just proven, the probability of winning is

$$\begin{aligned} a_n =& 1 - \frac{1}{(n-k)} + \frac{1}{(n-k)(n-k+1)} - \cdots + \\ & \frac{(-1)^k}{(n-k)(n-k+1)\cdots(n-2)} - \frac{(-1)^k}{n(n-k)(n-k+1)\cdots(n-2)}. \end{aligned}$$

This is an alternating series where each term is strictly smaller in magnitude than its predecessor, and so has an upper bound

$$a_n < 1 - \frac{1}{(n-k)} + \frac{1}{(n-k)(n-k+1)} = 1 - \frac{1}{(n-k+1)} < 1 - \frac{1}{n}.$$

The optimum when changing doors more than once is an inferior strategy to changing doors at the last possible moment, completing the proof.