ST102/ST109 Outline solutions to Exercise 4

1. Let B_j denote the event that the prize is in Box j, for $j = 1, 2, \ldots, n$, and let N_2 denote the event that Monty does *not* open Box 2. Bayes' theorem tells us that the conditional probability that the prize is in Box 1 is:

$$P(B_1 \mid N_2) = \frac{P(N_2 \mid B_1)\, P(B_1)}{\sum\limits_{j=1}^{n} P(N_2 \mid B_j)\, P(B_j)}$$

$$= \frac{P(N_2 \mid B_1)\, P(B_1)}{P(N_2 \mid B_1)\, P(B_1) + P(N_2 \mid B_2)\, P(B_2) + \sum\limits_{j=3}^{n} P(N_2 \mid B_j)\, P(B_j)}.$$

According to the rules of the game, the probabilities to use here are as follows.

- $P(B_1) = P(B_2) = \cdots = P(B_n) = 1/n$, since all the boxes are initially equally likely to contain the prize.
- $P(N_2 \mid B_1) = 1/(n-1)$. If the prize is in Box 1, Monty has a choice, so he chooses at random which of the remaining boxes to leave unopened.
- $P(N_2 \mid B_2) = 1$. If the prize is in Box 2, Monty must leave it unopened.
- $P(N_2 \mid B_j) = 0$ for $j = 3, 4, \ldots, n$. If the prize is in any of Boxes 3 to n, Monty cannot open that box. This means that he *must* open all the other boxes not chosen by you, including Box 2.

Therefore, the conditional probabilities are:

$$P(B_1 \mid N_2) = \frac{1/(n-1) \times 1/n}{1/(n-1) \times 1/n + 1 \times 1/n + (n-2) \times (0 \times 1/n)} = \frac{1}{n}$$

and $P(B_2 \mid N_2) = (n-1)/n$.

2.* We require:

$$P(\text{second box empty} \mid \text{first box empty})$$

in the sampling process of placing k objects into n boxes. Each of the k objects can be placed into any of the n boxes, hence there are n^k ways of placing the objects. For the first box to be empty, the k objects must be placed among the other $n-1$ boxes. Therefore:

$$P(\text{first box empty}) = \frac{(n-1)^k}{n^k}.$$

Similarly, for the first two boxes to be empty:

$$P(\text{second box empty and first box empty}) = \frac{(n-2)^k}{n^k}.$$

The required conditional probability is:

$$
\begin{aligned}
P(\text{second box empty} \mid \text{first box empty}) &= \frac{P(\text{second box empty and first box empty})}{P(\text{first box empty})} \\
&= \frac{(n-2)^k/n^k}{(n-1)^k/n^k} \\
&= \left(\frac{n-2}{n-1}\right)^k \\
&= \left(1 - \frac{1}{n-1}\right)^k.
\end{aligned}
$$

3. Being told 'at least two of them are boys' in (a) could *incorrectly* be inferred as 'the two oldest children are boys' – which is the information in (b). Let:

   - N_2 denote the number of boys among the oldest two children
   - N_3 denote the number of boys among the three children
   - B_i denote the event that child i is a boy, for $i = 1, 2, 3$, where i denotes the order of birth.

   The event that at least two of the children are boys is:

   $$\{B_1 \cap B_2\} \cup \{B_1 \cap B_3\} \cup \{B_2 \cap B_3\} = \{N_3 \geq 2\}$$

   and the event that the two oldest children are boys is:

   $$\{B_1 \cap B_2\} = \{N_2 = 2\}.$$

   (a) Applying Bayes' theorem, we have:

   $$
   \begin{aligned}
   P(N_3 = 3 \mid N_3 \geq 2) &= \frac{P(N_3 \geq 2 \mid N_3 = 3)\, P(N_3 = 3)}{\sum_{i=0}^{3} P(N_3 \geq 2 \mid N_3 = i)\, P(N_3 = i)} \\
   &= \frac{1 \times 1/8}{0 \times 1/8 + 0 \times 3/8 + 1 \times 3/8 + 1 \times 1/8} \\
   &= \frac{1}{4}.
   \end{aligned}
   $$

   Note that $P(N_3 = i)$, for $i = 0, 1, 2, 3$ are *binomial probabilities* but these can be determined without explicitly recognising this since:

   - $P(N_3 = 0)$ means no boys (i.e. three girls) which, by independence, is $(1/2)^3 = 1/8$
   - $P(N_3 = 1)$ means one boy (i.e. two girls) which, by independence, is $3 \times (1/2)^3 = 3/8$, due to there being three combinations
   - $P(N_3 = 2)$ means two boys (i.e. one girl) which, by independence, is $3 \times (1/2)^3 = 3/8$, due to there being three combinations
   - $P(N_3 = 3)$ means three boys (i.e. no girls) which, by independence, is $(1/2)^3 = 1/8$.

(b) Applying Bayes' theorem, we have:

$$\begin{aligned}
P(N_3 = 3 \,|\, N_2 = 2) &= \frac{P(N_2 = 2 \,|\, N_3 = 3)\, P(N_3 = 3)}{\sum_{i=0}^{3} P(N_2 = 2 \,|\, N_3 = i)\, P(N_3 = i)} \\
&= \frac{1 \times 1/8}{0 \times 1/8 + 0 \times 3/8 + 1/3 \times 3/8 + 1 \times 1/8} \\
&= \frac{1}{2}.
\end{aligned}$$

How to explain this difference? Well, by choosing families with three children conditional on at least two boys, we are choosing from half of families since:

$$P(N_3 = 2) + P(N_3 = 3) = \frac{3}{8} + \frac{1}{8} = \frac{1}{2}.$$

By choosing families with three children conditional on the oldest two children being boys, we are choosing from a more restricted set. However, *both* criteria are satisfied for the case of three boys. Hence such families will seem more likely, not because there are much such families, rather because there are *relatively* more of them!

4.* Let J_i = Judge i, for $i = 1, 2, 3$. We note that judges vote independently, and that a (not) guilty vote can be correct or incorrect. For example, in (a), the three guilty votes could be either a correct verdict of guilty (noting 70% of defendants are guilty) or an incorrect verdict of guilty (noting that 30% of defendants are not guilty).

(a) We have:

$$\begin{aligned}
P(J_3 \text{ votes guilty} \,|\, J_1 \text{ and } J_2 \text{ vote guilty}) &= \frac{P(J_1,\ J_2 \text{ and } J_3 \text{ vote guilty})}{P(J_1 \text{ and } J_2 \text{ vote guilty})} \\
&= \frac{0.70 \times (0.85)^3 + 0.30 \times (0.25)^3}{0.70 \times (0.85)^2 + 0.30 \times (0.25)^2} \\
&= 0.8286.
\end{aligned}$$

(b) We have:

$$\begin{aligned}
&P(J_3 \text{ votes guilty} \,|\, \text{one of } J_1 \text{ and } J_2 \text{ votes guilty}) \\
&= \frac{2 \times 0.70 \times (0.85)^2 \times 0.15 + 2 \times 0.30 \times (0.25)^2 \times 0.75}{2 \times 0.70 \times 0.85 \times 0.15 + 2 \times 0.30 \times 0.25 \times 0.75} \\
&= 0.6180.
\end{aligned}$$

(c) We have:

$$\begin{aligned}
P(J_3 \text{ guilty} \,|\, J_1 \text{ and } J_2 \text{ vote not guilty}) &= \frac{0.70 \times 0.85 \times (0.15)^2 + 0.30 \times 0.25 \times (0.75)^2}{0.70 \times (0.15)^2 + 0.30 \times (0.75)^2} \\
&= 0.3012.
\end{aligned}$$