Problem 0: Readings
Read notes 2 and 3 from the course web site.

Problem 1: Bayes and von-Neumann
A casino has two coins, one of them is fair, the other is not. This can be modeled as follows:

\[
P(H|\text{fair}) = P(T|\text{fair}) = \frac{1}{2}
\]

\[
P(H|\text{biased}) = 1 - P(T|\text{biased}) = p
\]

Note that this is slightly different than the classical approach where one would express the probabilities as follows:

\[
P(H) = P(T) = \frac{1}{2} \text{ coin is fair}
\]

\[
P(H) = 1 - P(T) = p \text{ coin is biased}
\]

This is typical with Bayesian analysis when the event of being fair or biased can itself be considered as a probabilistic outcome rather than a given. With probability \(q\) (usually small), the casino will use the biased coin. You get to observe the result of the flip; however, you do not know which coin is actually being used.

(a) Knowing \(p\) and \(q\), a Bayesian approach is to compute the probability that a particular coin has been used after observing several outcomes. This should give an indication to which coin is more likely to have influenced the observation. Let’s do this for one outcome and two outcomes. Find:

\[
P(F|H) \quad P(F|T) \quad P(F|HH) \quad P(F|HT) \quad P(F|TH) \quad P(F|TT)
\]

where \(F\) stands for fair and \(B\) stands for biased. Explain why, for instance,

\[
P(F|H) \neq P(F|HH) + P(F|HT)
\]

Solution:

\[
P(F|HH) = \frac{P(HH|F)P(F)}{P(HH|F)P(F) + P(HH|B)P(B)} = \frac{0.25(1-q)}{0.25(1-q) + p^2q}
\]
\[ P(F | HT) = \frac{0.25(1 - q)}{0.25(1 - q) + p(1 - p)q} \]
\[ P(F | TH) = \frac{0.25(1 - q)}{0.25(1 - q) + p(1 - p)q} \]
\[ P(F | TT) = \frac{0.25(1 - q)}{0.25(1 - q) + (1 - p)^2q} \]
\[ P(B | - -) = 1 - P(F | - -) \]

\[ P(F | H) = P(F | \{HH, HT\}) = \frac{P(\{HH, HT\} | F)P(F)}{P(\{HH, HT\})} = \frac{P(HH)P(F | HH) + P(HT)P(F | HT)}{P(HH) + P(HT)} \]

or simply:
\[ P(F | H) = \frac{P(H | F)P(F)}{P(H | F)P(F) + P(H | B)P(B)} = \frac{0.5(1 - q)}{0.5(1 - q) + pq} \]

Of course, even though \( P(F | H) = P(F | \{HH, HT\}) \), this does not mean that \( P(F | H) = P(F | HH) + P(F | HT) \). For one thing, this sum may exceed 1; for instance, if we have a strong belief that the coin is fair regardless of the outcome, then both \( P(F | HH) \) and \( P(F | HT) \) are high. Take for example the case when \( p = 0 \).

(c) When \( p \) and \( q \) are unknown, von-Neumann suggests the following strategy: Let the casino produce two flips each time, regardless of which coin has been chosen. If the outcome of the two flips is \( HH \) or \( TT \), discard that outcome (casino will have to flip again); otherwise, if the outcome is \( HT \), interpret it as an \( H \), and if the outcome is \( TH \), interpret it as a \( T \).

<table>
<thead>
<tr>
<th>outcome</th>
<th>interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH</td>
<td>-</td>
</tr>
<tr>
<td>HT</td>
<td>( H )</td>
</tr>
<tr>
<td>TH</td>
<td>( T )</td>
</tr>
<tr>
<td>TT</td>
<td>-</td>
</tr>
</tbody>
</table>

Therefore, the probability of observing an \( H \) or a \( T \) (actually interpreting the observation as such) is given by:

\[ P(H) = P(HT | \{HT, TH\}) \]
\[ P(T) = P(TH | \{HT, TH\}) \]

Show that regardless of \( p \) and \( q \), \( P(H) = P(T) = 1/2 \). Do this in two ways:

- tree structure showing the outcomes of 2 coin flips
- using the law of total probability: \( P(HT) = \sum_i P(HT | A_i)P(A_i) \) where \( A_i \) ranges over the following events \( \{F, B\} \) (do the same for \( TH \)).
**Solution:** I will do the second approach:

\[
\begin{align*}
P(HT) &= P(HT|F)P(F) + P(HT|B)P(B) = 0.25(1 - q) + p(1 - p)q \\
P(TH) &= P(TH|F)P(F) + P(TH|B)P(B) = 0.25(1 - q) + (1 - p)pq
\end{align*}
\]

Since both are equal, it follows that

\[
P(HT|\{HT, TH\}) = P(TH|\{HT, TH\}) = 1/2
\]

**Problem 2: Plain old Monty Hall**

(a) Make sure you understand why it is always better to switch in the Monty Hall problem described in class. Think about the problem from different angles. Keep it active in your mind for some time. This part will not be graded.

(b) Here’s a convincing argument in favor of switching. We can simulate the game.

1. \( c \leftarrow 0 \) (this counts the number of wins)
2. \( n \leftarrow 0 \) (this counts the number of trials)
3. place the prize uniformly at random in one of the boxes, call this box \( x \)
4. choose a box uniformly at random, call this box \( y \)
5. decide that you will not switch no matter what (the simulation of the remainder of the game becomes irrelevant)
6. if \( x = y \) then \( c \leftarrow c + 1 \)
7. \( n \leftarrow n + 1 \)
8. print \( c/n \)
9. goto 3

It should be obvious that \( c/n \) will converge to 1/3 (why?), which is the probability of winning if we adopt a no switch strategy. The simulation works the same way regardless of the actions taken by the host. Why is this not a proof that switching is better?

**Solution:** The 1/3 probability is simply the probability that the prize is in the chosen box. We are ignoring all kinds of hints that we may get from the host. But even though 1/3 < 1/2, this does not mean that switching is better. It depends on the strategy of the host.

(c) Assume that Monty Hall does not know where the prize is. He will simply open one of the remaining boxes at random with equal probability. Consider the case where an empty box is revealed. Show that there is no benefit in switching.

**Solution:** Let us consider \( P(x = A|y = A, z = B, B \text{ is empty}) \). This is equal to (using Bayes’ rule):
\[ P(y = A, z = B, B \text{ is empty} | x = A) P(x = A) + P(y = A, z = B, B \text{ is empty} | x = B) P(x = B) + P(y = A, z = B, B \text{ is empty} | x = C) P(x = C) = 1/3 \cdot 1/2 \cdot 1/3 + 1/3 \cdot 1/2 \cdot 0 \cdot 1/3 + 1/3 \cdot 1/2 \cdot 1/3 = 1/2 \]

Therefore, there is no benefit in switching.

**Problem 3: Find the winning ball**

There are \( k \) bins, and \( kn \) balls. Only one ball is winning. The balls are randomly placed in the bins, in such a way that each bin receives exactly \( n \) balls. Define the following events for \( i = 1 \ldots k, m = 1 \ldots n \):

- \( B_i \): bin \( i \) contains the winning ball
- \( b_{i,m} \): picking the \( m^{th} \) ball from bin \( i \) reveals the winning one

At this point, we have the following:

\[ P(B_i) = \frac{1}{k} \]

\[ P(b_{i,1}|B_i) = \frac{1}{n} \]

Therefore,

\[ P(b_{i,1}) = P(b_{i,1}|B_i) P(B_i) + P(b_{i,1}|B_i^c) P(B_i^c) = \frac{1}{n \cdot k} + 0(1 - \frac{1}{k}) = \frac{1}{kn} \]

which is intuitive because \( kn \) is the total number of balls.

You are interested in finding the winning ball using two strategies (or more precisely, I am interested in making you do it).

**STRATEGY 1**

You are allowed to pick the balls one at a time from any bin you like until you find the winning one. Once a ball is picked, it is removed from the game.

(a) What is \( P(B_j|b_{i,1}^c) \) in English? Use Bayes’ rule to find \( P(B_j|b_{i,1}^c) \). The answer should be different for \( j = i \) and \( j \neq i \). Can you interpret the result?

**Solution:** \( P(B_j|b_{i,1}^c) \) is the probability that bin \( j \) has the winning ball given that on the first trial a wrong ball was picked from bin \( i \). This can be computed as follows:

\[ P(B_j|b_{i,1}^c) = \frac{P(b_{i,1}^c|B_j) P(B_j)}{P(b_{i,1}^c|B_i) P(B_i) + P(b_{i,1}^c|B_j) P(B_j)} \]

The denominator is equal to \((1 - 1/n)1/k + 1(1 - 1/k) = 1 - 1/(nk)\). The numerator will depend on whether \( i = j \) or not.
Case 1: \( i = j \), we get
\[
\frac{(1 - 1/n)1/k}{1 - 1/(nk)} = \frac{n - 1}{nk - 1}
\]
Case 2: \( i \neq j \), we get
\[
\frac{1(1/k)}{1 - 1/(nk)} = \frac{n}{nk - 1}
\]
So any bin other than bin \( i \) has a higher chance of having the winning ball.

(b) This is a general equality: \( P(X, Y | Z) = P(X | Z)P(Y | X, Z) \) (we have used it before). Prove this equality.

**Solution:**
\[
P(X, Y | Z) = \frac{P(X, Y, Z)}{P(Z)} = \frac{P(Z)P(X | Z)P(Y | X, Z)}{P(Z)} = P(X | Z)P(Y | X, Z)
\]

(c) Find \( P(B_j, b_{j,2}| b_{i,1}^c) \), for \( j = i \) and \( j \neq i \). This is the probability of selecting the winning ball from bin \( j \) after the removal of a wrong ball from bin \( i \). Use part (b) to compute this probability. Based on your result, suggest a search algorithm for the winning ball.

**Solution:**
\[
P(B_j, b_{j,2}| b_{i,1}^c) = P(B_j| b_{i,1}^c)P(b_{j,2}| B_j, b_{i,1}^c)
\]

Case 1: \( i = j \), we get
\[
\frac{n - 1}{nk - 1} \frac{1}{n - 1} = \frac{1}{nk - 1}
\]
Case 2: \( i \neq j \), we get
\[
\frac{n}{nk - 1} \frac{1}{n} = \frac{1}{nk - 1}
\]
Therefore, all bins have equal chance. Any order is good.

**STRATEGY 2**
You are allowed to pick the balls one at a time from any bin you like until you find the winning one. If a wrong ball is picked, it is returned to the same bin.

Repeat the work for this strategy. **Hint:** only part (c) above should change.

**Solution:**
\[
P(B_j, b_{j,2}| b_{i,1}^c) = P(B_j| b_{i,1}^c)P(b_{j,2}| B_j, b_{i,1}^c)
\]

Only \( P(b_{j,2}| B_j, b_{i,1}^c) \) is affected. Since we now return the ball to its bin:
Case 1: \( i = j \), we get
\[
\frac{n - 1}{nk - 1} \frac{1}{n} = \frac{n - 1}{nk - 1} < \frac{1}{nk - 1}
\]
Case 2: \( i \neq j \), we get
\[
\frac{n}{nk - 1} \frac{1}{n} = \frac{1}{nk - 1}
\]
Therefore, we are more likely to find the winning ball on our second trial if we pick a bin other than bin \(i\). In fact, a round robin strategy will be the best in this case.

PS: You might want to try playing with small examples first, e.g. when \(n = 2\) and \(k = 2\), to get a feel of what is going on.

**Problem 4: What is random?**

Consider 10 independent tosses of a fair coin. These are two possible outcomes:

\[
\text{HTHTHTHTHT} \quad \quad \text{HHHHHHHHHH}
\]

These two outcomes have the exact same probability of \((1/2)^{10}\), yet, one seems to better conform to randomness than the other. Is this a dilemma?

(a) Ignore the actual sequencing, i.e. the first outcome consists of 5 heads and 5 tails; and the second outcome consists of 10 heads. Use a binomial random variable to settle this dilemma.

**Solution:** The probability of getting \(k\) heads is given by the binomial mass function.

\[
P(k \text{ heads}) = b(k, n, p) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{10}{k} 1/2^{10}
\]

When \(k = 5\), this is \(\frac{10!}{5!5!} \cdot \frac{1}{2^{10}}\), and when \(k = 10\), this is \(1/2^{10}\). Therefore, \(k = 5\) is more likely.

(b) Now consider a geometric random variable \(X\) with parameter \(p = 1/2\). What is \(P(X > 1)\) and \(P(X > 10)\)? Compare the two and explain how they can settle this dilemma.

**Solution:** For a geometric random variable \(X\) with parameter \(p = 1/2\), \(P(X > 1) = 1 - P(X = 1) = 1 - p = 0.5\). \(P(X > 10) = 1 - \sum_{i=1}^{10} P(X = i) = 1 - p - p(1-p) - p(1-p)^2 - \ldots - p(1-p)^9 = 1 - p[1+(1-p)+(1-p)^2+\ldots+(1-p)^9] = 1 - p \cdot \frac{(1-p)^{10} - 1}{(1-p)-1} = (1-p)^{10} = 1/2^{10}\). Obviously, \(P(X > 1)\) is higher, which is consistent with our belief for \(HTHTHTHTHT\) that reveals a tails as soon as the next position, i.e. \(X = 2\).

(c) Mr. Bayes to the rescue: consider the possibility that the coin is biased with a probability \(p > 1/2\) for heads, and that \(P(\text{fair}) = P(\text{biased}) = 1/2\) (zero knowledge). Using a Bayesian approach, compute \(P(\text{fair} | \text{outcome})\) for each of the two outcomes. How does the Bayesian approach settle this dilemma?

**Solution:**

\[
P(\text{fair} | \text{HTHTHTHTHT}) = \frac{P(\text{HTHTHTHTHT} | \text{fair})P(\text{fair})}{P(\text{HTHTHTHTHT} | \text{fair})P(\text{fair}) + P(\text{HTHTHTHTHT} | \text{unfair})P(\text{unfair})}
\]
\[
\frac{1/2^{10} \cdot 1/2}{1/2^{10} \cdot 1/2 + p^{5}(1 - p)^{5} \cdot 1/2}
\]

Similarly,

\[
P(\text{fair}|\text{HHHHHHHHH}) = \frac{1/2^{10} \cdot 1/2}{1/2^{10} \cdot 1/2 + p^{10} \cdot 1/2}
\]

Since \( p > 1/2 \), \( p^{10} > p^{5}(1 - p)^{5} \); therefore, \( P(\text{fair}|\text{HTHTHTHTHT}) > P(\text{fair}|\text{HHHHHHHHHH}) \).