CSE 312 HW3

Rohan Mukherjee

April 19, 2023

1. (a) Let H = the coin is heads, and T = the coin is tails. $\mathbb{P}(G_1) = \mathbb{P}(G_1 \mid H) \cdot \mathbb{P}(H) + \mathbb{P}(G_1 \mid T) \cdot \mathbb{P}(T)$ since we either get heads or tails, and by the LTP. If we pull heads, then $5/15$ of the balls are gold, so we have a $1/3$ chance of pulling a gold ball, i.e. $\mathbb{P}(G_1 \mid H) = 1/3$. A coin flip is uniform, so $\mathbb{P}(H) = \mathbb{P}(T) = 1/2$. If we pull a tails, then $10/15$ of the balls are gold, so $\mathbb{P}(G_1 \mid T) = 2/3$, so we get $\mathbb{P}(G_1) = 1/2(1/3 + 2/3) = 1/2$. Since she draws the second ball with replacement independently of the first ball, G_1 and G_2 are conditionally independent (after choosing the urn), so we can perform the same calculation to get $\mathbb{P}(G_2 \mid H) = 1/3$, and $\mathbb{P}(G_2 \mid T) = 2/3$, so again $\mathbb{P}(G_2) = \mathbb{P}(G_2 \mid H) \cdot \mathbb{P}(H) + \mathbb{P}(G_2 \mid T) \cdot \mathbb{P}(T) = 1/3 \cdot 1/2 + 2/3 \cdot 1/2 = 1/2$.

   (b) We see that, since G_1, G_2 are conditionally independent after picking the urn, i.e. after the outcome of the coin is decided, (this justifies the second equality, I ran out of space),

   $$
   \begin{aligned}
   \mathbb{P}(G_1 \cap G_2) &= \mathbb{P}(G_1 \cap G_2 \mid H) \cdot \mathbb{P}(H) + \mathbb{P}(G_1 \cap G_2 \mid T) \cdot \mathbb{P}(T) \\
   &= \mathbb{P}(G_1 \mid H) \cdot \mathbb{P}(G_2 \mid H) \cdot \mathbb{P}(H) + \mathbb{P}(G_1 \mid T) \cdot \mathbb{P}(G_2 \mid T) \cdot \mathbb{P}(T) \\
   &= 1/3 \cdot 1/3 \cdot 1/2 + 2/3 \cdot 2/3 \cdot 1/2 \qquad \text{(See part (a))} \\
   &= 1/2 \cdot 5/9
   \end{aligned}
   $$

   (c) We see that

   $$
   \mathbb{P}(G_1 \mid G_2) = \frac{P(G_1 \cap G_2)}{\mathbb{P}(G_2)} = \frac{1/2 \cdot 5/9}{1/2} = 5/9
   $$

   from the numerous calculations above.

   (d) As $P(G_1 \mid G_2) = 5/9$ while $P(G_1) = 1/2$, we see that these events are NOT independent from the definition of independent probability.

   (e) My intuition for this problem is that if we know G_2 happens, i.e. that the second coin pulled was gold, it is at least a little more likely that the coin flip showed up tails, which in the end would make G_1 more likely. This is why $\mathbb{P}(G_1 \mid G_2) > \mathbb{P}(G_1)$.

2. (a) Let Q = you know the answer to the question. Clearly $\overline{Q}$ = you don't know the answer to the question. Finally, let C = you get the question wrong. Note that $\overline{C}$ = you get the question right. We are looking for $\mathbb{P}(\overline{C}) = 1 - \mathbb{P}(C)$. By the LTP, $\mathbb{P}(C) = \mathbb{P}(C \mid Q) \cdot \mathbb{P}(Q) + \mathbb{P}(C \mid \overline{Q}) \cdot \mathbb{P}(\overline{Q})$. If you know the answer to the question, then you get it right with 100% probability, so $\mathbb{P}(C \mid Q) = 0$. It stated in the problem that $\mathbb{P}(Q) = p$, and also that $\mathbb{P}(\overline{Q}) = 1 - p$. If you do not know the question, you have to guess randomly among 5 options–i.e., 4/5 of the time you get the question wrong. So $\mathbb{P}(C \mid \overline{Q}) = 4/5$. We conclude that

$$\mathbb{P}(C) = (1 - p) \cdot 4/5.$$

Which tells us that $\mathbb{P}(\overline{C}) = 1 - (1 - p) \cdot 4/5 = (5 - (4 - 4p))/5 = (1 + 4p)/5$.

(b) Using the notation from part (a), we are looking for $\mathbb{P}(Q \mid \overline{C})$. By Bayes rule,

$$\mathbb{P}(Q \mid \overline{C}) = \frac{\mathbb{P}(\overline{C} \mid Q) \cdot \mathbb{P}(Q)}{\mathbb{P}(\overline{C})}$$

If we know the answer to the question, we get it right with 100% probability, so $\mathbb{P}(\overline{C} \mid Q) = 1$. The problem states that $\mathbb{P}(Q) = p$. Using that $\mathbb{P}(\overline{C}) = (1 + 4p)/5$ from part (a), we get that

$$\mathbb{P}(Q \mid \overline{C}) = \frac{1 \cdot p}{(1 + 4p)/5} = \frac{5p}{4p + 1}$$

For a quick check, since $p \leq 1$, $5p/(4p + 1) \leq 1$, and as p gets small, the probability also gets small (if the probability that you know the answer to the question is smaller, this probability should probably be smaller too, since you will be guessing more often).

3. (a) Since both Xena's parents have type A blood, both her parents must have at least one A gene. Since Yvonne has type O blood, and the only way to have type O blood is to have genotype OO, it must be the case that both parents also carry an O gene. So both parents have genotype AO. By Bayes rule, we see that

$$\mathbb{P}(G_X = AO \mid Ph_X = A) = \frac{\mathbb{P}(Ph_X = A \mid G_X = AO)\mathbb{P}(G_X = AO)}{\mathbb{P}(Ph_X = A)}$$

If $G_X = AO$, then clearly $Ph_X = A$ (since it said that the A dominates the O), so $\mathbb{P}(Ph_X = A \mid G_X = AO) = 1$. The parents both have genotype AO, and since the problem states that they could give either gene down with equal probability, the multiset of all possible genotypes (where order first comes from the father, and then the mother) is $\{AO, OA, AA, OO\}$. We are looking for two of those outcomes, so $\mathbb{P}(G_X = AO) = 2/4 = 1/2$, since passing genes down is uniform. Finally, Xena's phenotype is uniquely determined by her genotype. 3 of the 4 possible ordered genotypes would give her a phenotype of A, so $\mathbb{P}(Ph_X = A) = 3/4$. We conclude that $\mathbb{P}(G_X = AO \mid Ph_X = A) = 1/2/3/4 = 2/3$, which is exactly what you would expect (i.e., her genotype could only be AA, AO, or OA, and in two of those cases she has an O).

   (b) Since Zachary has type O blood, he must have genotype OO. Therefore, we can restrict to the sample space where Zachary has genotype OO. We are now looking (in the restricted sample space) for $\mathbb{P}(G_C = OO \mid Ph_X = A)$. Since the mother has type A blood, she can either have genotype AO, or genotype AA. By the law of total probability, $\mathbb{P}(G_C = OO \mid Ph_X = A) = \mathbb{P}(G_C = OO \mid Ph_X = A \cap G_X = AO) \cdot \mathbb{P}(G_X = AO \mid Ph_X = A) + \mathbb{P}(G_C = OO \mid Ph_X = A \cap G_X = AA) \cdot \mathbb{P}(G_X = AA \mid Ph_X = A)$. If the mother has genotype AA she can't pass down an O gene, so the child can't have genotype OO, so $\mathbb{P}(G_C = OO \mid Ph_X = A \cap G_X = AA) = 0$. We found last time that $\mathbb{P}(G_X = AO \mid Ph_X = A) = 2/3$. In the case where the mother has type A blood and genotype AO, the multiset of all possible outcomes with the genes from Zach would be $\{AO, OO, AO, OO\}$. Since they give their genes uniformly, the child has genotype OO 1/2 of the time, so $\mathbb{P}(G_C = OO \mid Ph_X = A \cap G_X = AO) = 1/2$. Therefore, $\mathbb{P}(G_C = OO \mid Ph_X = A) = 2/3 \cdot 1/2 = 1/3$.

   (c) Finally, we are looking for $\mathbb{P}(G_X = AO \mid Ph_X = A \cap G_C = AO)$ (Note: the child MUST have genotype AO by the logic from the last problem, since he either has genotype AO, or OO, and only in the first case does he have phenotype A). By Bayes rule, this equals

$$\frac{\mathbb{P}(G_C = AO \cap G_X = AO \cap Ph_X = A)}{\mathbb{P}(G_C = AO \cap Ph_X = A)}$$

We notice that $(G_X = AO \cap Ph_X = A) = G_X = AO$, since if the mother has genotype AO, then she already has phenotype A. So the above equals

$$\frac{\mathbb{P}(G_C = AO \cap G_X = AO)}{\mathbb{P}(G_C = AO \mid Ph_X = A)\mathbb{P}(Ph_X = A)}$$
$$= \frac{\mathbb{P}(G_C = AO \mid G_X = AO) \cdot \mathbb{P}(G_X = AO)}{\mathbb{P}(G_C = AO \mid Ph_X = A)\mathbb{P}(Ph_X = A)}$$

By similar logic near the end of the last problem, if the mother has genotype AO, then the child has genotype AO 1/2 of the time (the other half of the time being OO), so $\mathbb{P}(G_C = AO \mid G_X = AO) = 1/2$. We found in part (a) that $\mathbb{P}(G_X = AO) = 1/2$. We also notice that, in the restricted space where the mother has genotype AO, the child can have either genotype AO, or genotype OO, so $\mathbb{P}(G_C = AO \mid Ph_X = A) = 1 - \mathbb{P}(G_C = OO \mid Ph_X = A) = 1 - 1/3 = 2/3$. We found in part (a) that $\mathbb{P}(Ph_X) = 3/4$. Putting it all together we get: $(1/2 \cdot 1/2)/(2/3 \cdot 3/4) = (1/2 \cdot 1/2)/(1/2) = 1/2$, as our final answer.

4. (a) The number of hands with exactly one ace is just going to be $4 \cdot \binom{48}{12}$, since it could have one of 4 aces, and then we have to exclude those 4 aces and pick 12 cards from the remaining 48 after exclusion. The total number of 13 hand cards is $\binom{52}{13}$, so since in this case the probability is uniform (i.e., we shuffle the deck uniformly at random), we see that $\mathbb{P}(A_1) = 4 \cdot \binom{48}{12}/\binom{52}{13}$. Since this argument didn't depend on it being the first deck we get that $\mathbb{P}(A_2) = 4 \cdot \binom{48}{12}/\binom{52}{13}$ as well (this works because we are working with NO other assumptions, we JUST want that a specific deck has precisely one ace). On the other hand, if we already know that A_1 has an ace, A_2 could have one of 3 aces, and then we would have to pick 12 cards from the remaining 52 - 13 - 3 (we can't pick from the aces), to get that the second deck, assuming it has one ace, could be one of $3 \cdot \binom{36}{12}$ decks. The total number of decks the second hand could be is just $\binom{39}{13}$ (since aces can be included, but none of the cards in the first hand can be included), we get that $\mathbb{P}(A_2 \mid A_1) = 3 \cdot \binom{36}{12}/\binom{39}{13}$. Plugging these into a calculator shows $\mathbb{P}(A_2 \mid A_1) \neq \mathbb{P}(A_2) = \mathbb{P}(A_1)$, so they are not independent.

(b) Continuing on with the calculation from before, assuming the first and second deck have precisely one ace, the third deck could have one of two aces, and then it needs 12 of the remaining 52 - 13 - 13 - 2 non-ace cards, which can be done in $\binom{24}{12}$ ways. The sample space, i.e. the number of possible third decks is just $\binom{26}{13}$, since an arbitrary deck could have none/more than one ace. We get that $\mathbb{P}(A_3 \mid A_1 \cap A_2) = 2 \cdot \binom{24}{12}/\binom{26}{13}$. Finally, note that if all 3 previous decks have exactly one ace, then the fourth deck has exactly one ace 100% of the time, i.e. with probability 1, so $\mathbb{P}(A_4 \mid A_1 \cap A_2 \cap A_3) = 1$. We conclude that

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdot \mathbb{P}(A_4 \mid A_1 \cap A_2 \cap A_3)$$
$$= 4 \cdot \binom{48}{12}/\binom{52}{13} \cdot 3 \cdot \binom{36}{12}/\binom{39}{13} \cdot 2 \cdot \binom{24}{12}/\binom{26}{13}$$

which is around 10.5%.

5. (a) Since the n profiles are shown uniformly at random (independent of what I think about them), this is just the probability that the most beautiful woman is in the 3rd position, which happens $1/n$ of the time (she could be in any of n places, and the event we want is her in 1 specific place).

(b) We can think of this as finding the most beautiful woman in the first $i-1$ spots. We are looking for the probability that the most beautiful woman among the first $i-1$ spots (inclusive) lies somewhere in spots $1, 2, \ldots, q-1$. Since the most beautiful woman can be in any of the $i-1$ spots uniformly at random, and our event has her in $q-1$ specific spots, the probability of this event happening is just $(q-1)/(i-1)$.

(c) In this case, we have to consider two possibilities:

First, if $q \leq i$, when the most beautiful woman is in position i, the probability that she will be matched with (i.e., $\mathbb{P}$(match using strategy | woman at position i)), is just the probability that the most beautiful woman among the first $q-1$ is also the most beautiful woman among the first $i-1$, since if this is the case the first woman we will see who is more beautiful than all the previous ones will be "the one". So, given that $q \leq i$, $\mathbb{P}$(match using strategy | woman at position i) $= (q-1)/(i-1)$, the probability we found last time.

Second, in the case where $q > i$, we will have already seen "the one", and so no woman we see next will ever be as beautiful as her. Since we cannot match with any woman that came before position q, we see that we have a probability of 0 of matching with "the one".

(d) Let $q > 1$. Since the most beautiful woman can be in any of the n spots, we see that (by the law of total probability),

$$\mathbb{P}(\text{match using strategy}) =$$

$$\sum_{i=1}^{n} \mathbb{P}(\text{match using strategy | woman at position i}) \cdot \mathbb{P}(\text{woman at position i})$$

We saw last time that if $i < q$, then we have no chances of matching with "the one", so the sum's first $q-1$ terms are just 0. Therefore, this sum equals

$$\sum_{i=q}^{n} \mathbb{P}(\text{match using strategy | woman at position i}) \cdot \mathbb{P}(\text{woman at position i})$$

Since the woman can be in any of the n spots uniformly at random, $\mathbb{P}$(woman at position i) $= \frac{1}{n}$. We showed in part (b) that for $i \geq q$, $\mathbb{P}$(match using strategy | woman at position i) $= \frac{q-1}{i-1}$. So this sum becomes

$$\sum_{i=q}^{n} \mathbb{P}(\text{match using strategy | woman at position i}) \cdot \mathbb{P}(\text{woman at position i})$$

$$= \sum_{i=q}^{n} \frac{q-1}{i-1} \cdot \frac{1}{n}$$

6

In the case where $q = 1$, we "reject" the first 0 people, and we accept the first woman who is more beautiful than the first 0 people. Since this will always just be the first woman, if $q = 1$ we have a $\frac{1}{n}$ chance of matching with the most beautiful woman (by the logic from part (a)).

(e) I plugged my probability into a program, and got that for $n = 100$, the best q is 38. For $n = 1000$, the best q is 369.