Stat 110 Homework 2   Mark Grozen-Smith

1.

a) The possible ways to get a *yes* is if someone does drugs and draws a drug slip or if they don't do drugs and they draw a no-drug slip. The independence of the event determining if the person does drugs or not and the event determining which type of slip they draw allows us to multiply those probabilities together or each type of person. The dijointness of the two events — a druggie drawing a drug slip and a non druggie drawing a no-drug slip — allows us to add the two probabilities. Therefore $\boxed{P(y) = p * d + (1 - p) * (1 - d)}$

b) The best case scenario would be if $p = 0 \, or \, 1$, that would be equivalent to be asking a yes or no question to the individuals and we could get perfect information out of the responses. On the other hand, we have a severe lack of information regardless of individuals' responses if the probability of responding with no is the same as that of a yes (i.e. $y = P(y^c)$). To derive the probability p, we set $P(y) = P(y^c)$). Which is $p * d + (1 - p) * (1 - d) = (1 - p) * d + p * (1 - d)$ which results in $\boxed{d = \dfrac{1}{2}}$ as the worst case scenario.

c) Again, $y = p * d + \frac{1-p}{4}$ by the same argument as the end of part (a). Solving for d we find $\boxed{d = \dfrac{y}{p} - \dfrac{1-p}{4p}}$

2.

a) For $(A = D)$, everyone in D must be in A and everyone not in D must not be in A. Now let's assume D has N people. For each of those N people, they have a $1/2$ chance of being a part of A. For the rest of the people in C who are not in D, namely $100 - N$, there is a $1/2$ chance they are not in A. Since each person is independent of each other, we can just multiple all these numbers together. This leaves us with $(\frac{1}{2})^N (\frac{1}{2})^{100-N} = \boxed{(\dfrac{1}{2})^{100}}$.

b) For a single person, the only legal relationship possibilities are $AB, A^c B$, and $A^c B^c$. Each of these have a probability $(\frac{1}{4})$, and since they are all possible, though disjoint, for each person, we add them together. Thus, for each person, there is a $(\frac{3}{4})$ chance that they fit in with $(A \subseteq B)$. Since each person is independent of the others, the total $P(A \subseteq B) = \boxed{(\dfrac{3}{4})^{100}}$.

c) Each person has a $(\frac{1}{2})$ chance of not being friends with Alice, and the same for not being friends with Bob. Since all relationships are independent, this implies that there is a $(\frac{1}{4})$ probability that a person is friends with neither Bob nor Alice. Therefore, there is a $(\frac{3}{4})$ probability that a single person is friends with at least one of Alice or Bob. By independence, we can multiply this probability together 100 times to find the probability that all members of C are friends with Alice and/or Bob and we get $\boxed{(\frac{3}{4})^{100}}$.

3.

a) This statement assumes that the first two coins are already the same. The statement is true, given that the first two flips resulted the same, however there is only a .5 chance of that happening in the first place.

b) Each flip is independent and the coins are fair, so this problem can easily be solved using the nave definition of probability. How many ways can we have 3 heads divided by how many ways can we have at least 2? The option for all heads is just HHH, whereas for at least 2 heads we have HHH, THH, HTH, or HHT. Thus, the probability that all three tosses are heads, given at least 2 are heads, is $\boxed{1/4}$.

c) Since the coin flips are fair and independent, and we only need one more flip to result in heads, we can just take the probability that the final flip is heads. Which is $\boxed{1/2}$.

d) The difference here between this problem and the last is that this problem eliminates the options THH and HTH since we already know the first 2 tosses were heads.

4. This problem has confounding data. We don't care at all about the Democrats or the Others because if we see any of them, we just continue sampling. Thus, this problem boils down to "If we observe one person's political perspective, what is the probability of them being Green-Rainbow rather than Republican?" This would be just p_3 if $p_3 + p_4 = 1$, however this is not true. Our Pebble World^{TM} has lost some mass! We have to renormalize by dividing by $(p_3)(p_4)$. This leaves us with the probability of observing a Green-Rainbow before a Republican as $\boxed{\dfrac{(p_3)}{(p_3) + (p_2)}}$.

5.

a) I expect $P(B)$ to be equal 1/13 since, on average, we expect to draw an Ace on the 13th card. A that point there are 3 Aces left and 39 cards left total. Thus, the probability of drawing an Ace next is $\frac{3}{39} = \boxed{\frac{1}{13}}$

b) We know C_j occurred, therefore we know that $52 - j$ cards are left in the deck and three of them are aces. The locations of each of the remaining cards are all independent of each other and have equal probabilities so we know the $\boxed{P(B \mid C_j) = \dfrac{3}{52 - j}}$.

c) By the Law of Total Probabilities, $P(B)$ is the sum of all $P(B \mid C_j) * P(C_j)$ for j between 1 and 48 as those are all the possible positions for the first Ace to be in. By that logic, $P(B) = \sum_{j=1}^{48} \frac{3*P(C_j)}{52-j}$. To find $P(C_j)$, just think about what the probability is of seeing $j - 1$ non-Aces in a row and then one Ace. This gives $P(C_j) = \frac{\frac{48!}{(48-j-1)!}*4}{\frac{52!}{(52-j)!}}$. Combining these two, we

get $P(B) = \sum_{j=1}^{48} \frac{12}{52-j} * \frac{\frac{48!}{(48-j-1)!}}{\frac{52!}{(52-j)!}}$

d) Before we flip over any cards, every card has an equal probability of being an Ace, 1/13. Sure, as we flip over the Ace, the probability of any unflipped card being an Ace decreases, but every j-1 card we flip over (non-Ace) increases the probability of an unflipped cards being an Ace. Over all, these effects cancel out and the probability any remaining given card is an Ace stays at 1/13. This may seem unlikely for the j+1 card as we just drew an Ace as the j card. But this idea feels ok for the final card in the deck. Which is not intrinsically different from the j+1 card by symmetry. Thus, P(B) = 1/13.

6.

a) We are trying to find $\frac{P(d|T)}{P(d^c|T)}$ where d: individual has the disease and T: test results as positive. We have the sensitivity, $P(T \mid d)$, and the specificity, $P(T^c|d^c)$.

By Bayes Rule, $P(d \mid T) = \frac{P(T|d)P(d)}{P(T)}$ and $P(d^c \mid T) = \frac{P(T|d^c)P(d^c)}{P(T)}$.

If we divide these two, the $P(T)$ cancels and we get $\frac{P(d|T)}{P(d^c|T)} = \frac{P(T|d)P(d)}{P(T|d^c)P(d^c)}$

With the prior odds being $\frac{P(d)}{P(d^c)}$, the sensitivity, and $P(d \mid T^c) = 1-$ specificity, we see that the posterior odds of having the disease given a positive test is $\frac{priorOdds*sensitivity}{1-specificty}$

b) By the math shown above, if the disease is rare, $diseasePrior$ is low thus making the sensitivity factor drop out in comparison to the $(1 - specificity)$ term. Thus, this is mathematically proven. Intuitively, however, consider a disease that affects 1 in a million people. If a test is very sensitive, it will have a lot of false pasitives to make sure all the positives are caught, but we are not expecting a lot of positives. We are expecting very few, so it doesn't help us to make sure we catch every case. We want to be super sure that, if the test returns positive, we are certain the individual actually has the disease. Thus, the greater the specificity, the more accurate we can be with the positive predictive value.

7. If GG is the event that we have 2 girls, and G_c is the event that at least one child is a girl with characteristic c (which has probability p of occurring), what we are trying to find is $P(GG \mid G_c)$ which, by Bayes Rule, is $\frac{P(G_c|GG)P(GG)}{P(G_c)}$. We can say that $P(G_c \mid GG) = 1 - (1 - p)^2$ because that comes down to 1 minus the probability that you have 2 kids who do not exhibit that characteristic. We know $P(GG) = \frac{1}{4}$ by the simple coin toss situation. To find $P(G_c)$, we must think about $1 - P(G_c^c)$. Since the two children are independent $1 - P(G_c^c) = 1 - P(\text{neither child is a girl with C}) = 1 - P(\text{a child is not a girl with C})^2 = 1 - (1 - P(GC))^2$ where GC is the event that a child is a girl with C which we know has probability $\frac{p}{2}$. Therefore, $P(G_c) = 1 - (1 - \frac{p}{2})^2$ and $P(GG \mid G_c) = \frac{(1-(1-p)^2)\frac{1}{4}}{1-(1-\frac{p}{2})^2}$. This simplifies to $\frac{2p-p^2}{4p-p^2} = \frac{2-p}{4-p}$.