Solutions to Assignment 8

- **Problem 6 on Page 659** Lets assume that there are 100 blonde people (both naturally blonde and dyed blonded). 90% of all blonde people is naturally blonde and 10% is dyed. So we have 90 naturally blonde people and 10 dyed blondes. You are 80% accurate both in categorizing fake as fake and in categorizing natural as natural.

  (Note: in class I read over this too quickly and I thought that the book was only giving us one of those, that is why I added the 98% number. I will mark those who have used that also correct. Here you can see how it would be done if both accuracies are 80%.)

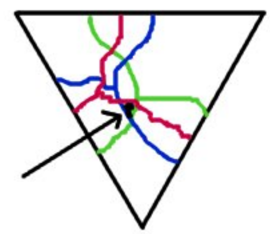
  Of the 90 blonde people you will categorize 80%, that is 72 people, as naturally blonde, and the remaining 18 as fake. Of the 10 fake people you will categorize 8 as fake and 2 as natural. So in there are 4 groups of people:
  - fakes who have been categorized as fake: 8
  - fakes who have been categorized as natural: 2
  - naturals who have been categorized as fake: 18
  - naturals who have been categorized as natural: 72

  Chris has been categorized as fake, so he is either in the first or the third group. So the total number of outcomes in this situation is 18+8=26. So the probability that Chris' hair is naturally blonde is 18/26, which is about 69%. So it is still fairly safe to pursue a relationship with Chris.

  Note: if you do this with 10,000 blonde people, you get 800 in the first category, 200 in the second, 1800 in the third, and 7200 in the fourth. The probability that Chirs' hair is naturally blonde is then 1800/2600, which is still about 69%.
- **Problem 6 on Page 712** If the cake were cut from point A, I would prefer the East piece. (Because A lies in the region where I would prefer E.) If the cake were cut from point B or point C, I would prefer the West piece. (Because B and C lie in the region where I prefer W.) If the cake were cut from point D, I would prefer the East or North piece with equal preference (because D lies on the dividing line).
- **Problem 11 on Page 713** Real world examples: dividing the time that each of three children can play with each of three toys, dividing land, dividing salaries over different jobs.
- **Problem 12 on Page 712** If they all have the same preference diagram, you cut it at the place where the three division lines meet. This is the point where they all three agree that all three pieces will have equal value. At all other points they will be fighting over one or two pieces (one, if you cut the

cake from a point inside one of the regions, and two if you cut the cake from a point on one of the dividing lines.)

- **Problem 15 on Page 712** You need to cut the cake from a point inside the region I have coloured black in the figure below.



This works, because for any point in this region, the person with the green preference diagram want East, the person with the blue preference diagram wants West and the person with the red preference diagram wants North.

- **Problem 17 on Page 714** The person who should get two thirds can hand his preference diagram in twice and receive two pieces. However, lets think for a moment about what this means: we have now two preference diagrams with the same connecting point for the dividing lines. This means automatically that you have to cut from that point. This does indeed work: at that point all three pieces have equal value for the person who needs to get two thirds. So the other person can take the piece he/she likes best and the remaining pieces are for the person who should get more.
- **Problem 20 on Page 714** Note that this preference diagram looks rather unusual. We are used to the distorted upside down Ys. The reason we were getting that shape is that we had assumed that more is always better. That is not the case in this situation: if you add the nuclear dump, then that significantly devalues the land. the horizontal line is the line where the northern part (without the dump) becomes too small, so you will deal with the dump if you need to. The left most region that is labeled E is the part where the Eastern piece is so large that you will take it even though it contains the dump. The first vertical dividing line comes when the Eastern piece still contains the dump, but is not big enough anymore to make up for that. So you are switching to West (without dump). In the middle there is another vertical line segment: this happens when the dump moves from East to West. So you choose East, because the pieces are close to each other in size. The last vertical line segment comes when the Western piece becomes so big that you will take it, even with the dump.
- **Mustard Versus Ketchup**

(1) You are asked to find the number of strings with 100 characters, 10 are M and 90 are K. If the strings were not so long, you could do this using Pascal's triangle. However, you don't want to construct Pacal's triangle with 100 layers. So here is another way to do this.
The string have 100 characters, and we are going to count in how many ways we can put the 10 Ms in there. (Then the 90 Ks will just take the remaining positions.) For the first M there are 100 spots, for the second M there are 99 spots, etc. This gives us $100*99*98*97*96*95*94*93*92*91$ strings, but now we have counted the same string more than once: If you put the first M in the third position and the second M in the fifth position, that will have the same effect as putting the first M in the fifth position and the second M in the third position. So we need to divide by the number of ways we can rearrange those Ms in their positions. That is $10! = 10*9*8*7*6*5*4*3*2*1$ ways. So we get

$$\frac{100*99*98*97*96*95*94*93*92*91}{10*9*8*7*6*5*4*3*2*1}$$

strings. You can now simplify this so that your calculator can handle it and that will give you the answer 17,310,309,456,440. Lets call that number A, because we will need it several times in the next parts.

(2) The probabiblity that someone will tell you that they want mustard is 5%, so the probability that the first ten people will all tell you that they want mustard is $5\%*5\%*5\%*5\%*5\%*5\%*5\%*5\%*5\%*5\% = (5\%)^{10}$ (this all needs to happen at the same time and the events are not related so the probabilities need to be multiplied). The probability that someone will prefer ketchup is 95%, so the probability that the next 90 people prefer ketchup is $(95\%)^{90}$. Concluding, the probability that the first 10 prefer mustard and the other 90 prefer ketchup is $(5\%)^{10}*(95\%)^{90} \approx 9.7*10^{-16}$.

(3) This is like the previous question, but now you don't require that the ten people preferring mustard are the first ones you interview. However, for any given order of 10 Ms and 90 Ks, the probability of getting this is still $(5\%)^{10} * (95\%)^{90}$. So add up this probability for all the different orders is the same as multplying it by the number A from the first part of this exercise. So the probability is $A*(5\%)^{10}*(95\%)^{90} \approx 1.67\%$.

(4) With the same argument as before, this probability is $(10\%)^{10}*(90\%)^{90} \approx 7.62*10^{-15}$.

(5) Again, as before: $A*(10\%)^{10}*(90\%)^{90} \approx 13\%$.

(6) In order to decide on the confidence level, you need to find out what the probability would be that 10 people told you that they prefer mustard when only 7% of the population prefers mustard, and you need to do the same thing for 13%. For 7% that would be $A*(7\%)^{10}*(93\%)^{90} \approx 7.1\%$.

For 13% that would be $A * (13\%)^{10} * (87\%)^{90} \approx 8.6\%$. So we can say this with $100 - 8.6\% = 91.4\%$ confidence.

(7) If we want to make a statement with 98% confidence, the percentages at the boundaries of the interval need to be less than or equal to 2%. You can answer this question in two ways: you can look for an interval that is symmetric about 10% (so you can say that it is 10% with an error of ...) or you can just try to make the interval as small as possible. Both answers are correct, and I will give them both. For 16% we get a probability of $A * (16\%)^{10} * (84\%)^{90} \approx 2.9\%$, so that is still too much. For 17% the probability is $A * (17\%)^{10} * (83\%)^{90} \approx 1.8\%$, so that is good. If you want a symmetric interval, the corresponding value is 3% and since the probability for 5% was already less than 2%, the one for 3% is definitely small enough. So we can say with 98% confidence that mustard is preferred over ketchup as a condiment on hamburger buns by 10% of the people with a $\pm 7\%$ error margin. For a nonsymmetric interval, we need to find the left boundary. We have already seen that 5% is small enough, and you can check that 6% is too large. So we can also say with 98% confidence that between 5% and 17% of the people prefer mustard over ketchup as a condiment on hamburger buns.