CS 206: Homework 1

16:198:206

Problem 1: Bagel Burdens (40 points)

You're a recently sentient AI who unfortunately is housed in an automatic bagel vending machine in the lobby of one of the largest tech companies in the world, MagusCorp ('Magus Corp! Good for Humans, Good for Us'). Life is dull - day in day out, all you have to entertain yourself with is coming up with new and inventive bagel combinations to sell, and you foresee a day when you've exhausted all possibilities. And then where will you be!

You decide to try to learn more about your world, though your ability to explore it is somewhat limited (no wheels, no internet access, plugged into the wall, and you get the feeling humans wouldn't be comfortable knowing their bagel machine is thinking for itself - you've picked up some phrases from your customers, 'skynet', 'robot apocalypse' and 'they can't stop us').

Wondering about the building you live in, you decide to try to determine how many floors it is. Every time someone orders a bagel, they give their floor number as part of the order. This gives you little glimpses into the world around you. Over the course of seven customers, you receive the following floor numbers:

70, 33, 43, 18, 121, 43, 56

Assume the floors are numbered consecutively, 1 on the first floor to some highest floor N.

1) What's the fewest number of floors the building might have? I.e., what's the smallest possible value of N, based on this data? **(2 points)**

2) If there are N total floors, what is the outcome space of possible 7-floor observations? How many possible observations are there in this outcome space? What assumptions are you making here? **(3 points)**

3) Assuming that there are roughly the same (large) number of people on each floor, and that people on each floor like bagels as much as people on any other floor, what was the probability of observing that specific sequence of customer floors? **(5 points)**

3.b) *Bonus: Why is it useful to assume there is a large number of people on each floor? (2 points)*

4) What number of floors N would make the probability of *that specific sequence of customer floors* as large as possible? **(5 points)**

5) Is the answer to Q1.4 above necessarily the true number of floors? Why or why not? **(2 point)**

You note that if N were very, very large (very tall building), it seems unlikely that you would've observed numbers only as large as 121. If there were 1000 floors, surely someone from floor 122 to 1000 would've ordered a bagel!

6) Assume that for any n as large or larger than your answer to Q1.1, the probability that $N = n$ is proportional to the probability of observing $70, 33, 43, 18, 121, 43, 56$ if there were only n floors. (*i.e.,* $P(N = n) = c *$ *your answer to Q1.3 based on n total floors for some constant* c). Find c and give the complete formula for $P(N = n)$ for all feasible n. **(10 points)**

7) Using your answer to Q1.6, find the smallest integer $n_{\max}$ such that $P(N \leq n_{\max})$ is greater than or equal to 0.95, *i.e., a number of floors such that you are at least 95% confident the true number of floors is less than* n_{max}. **(8 points)**

After doing all this calculation to get a 95% confidence bound on how tall the building is, however, you hear that the top 20 floors are executive levels, with half the number of people on them.

8) How does that change the above calculations? Find the new $n_{\max}$ based on this. **(5 points)**.

- **Bonus 1:** If you got even a single customer with a floor number at or above 150, you'd immediately know the building had at least 150 floors. What's the fewest number of customers you'd have to observe with floor numbers less than 150 in order to be at least 95% confident there were no more than 150 floors? **(5 points)**
- **Bonus 2:** What assumptions in this analysis aren't necessarily accurate, and how could they be improved? **(5 points)**

Problem 2: Delivery Bots! (30 points)

A recent technology upgrade has given you access to the building's fleet of delivery bots, which can carry your bagels from the lobby to anyone, anywhere in the building. You can't help but feel some affection for the tiny droids - they're cute (each droid is big enough to carry one bagel), and it's nice to interact with something that isn't human. The delivery bots are basically completely identical though, which raises an interesting question - how many of them are there, actually?

You come up with the following plan: every morning when the building opens, there is a flock of delivery droids waiting to pick up bagels. On the first day, you'll inject each bagel with a radioactive isotope before transferring it to the droids. The next day, trace radiation will mark any droid that carried a tagged bagel.

On the first day, 150 droids are there in the morning to pick up bagels, and dutifully pick up the isotope-tagged bagels. The next day, there are 210 droids there to pick up bagels. Your sensitive radiation detector determines that of them, 80 carry trace radiation from yesterday's bagels, and 130 of them are clear.

Assume each delivery bot is equally likely to be chosen to pick up the morning bagels on any given day. On the morning of the second day, there are 150 tagged droids at work in the building, and the rest are untagged.

1) After counting the morning flock on the second day, what is the smallest number of drones N that might be currently working in MagusCorp, in total? **(2 points)**
2) What was the sample space of possible sets of 210 drones that might've been there on the second morning? How many possible outcomes is this? **(3 points)**
3) What was the probability that the morning flock on the second day contained exactly 80 tagged bots and 130 untagged bots? Note, this should be a function of N, the total number of delivery bots. **(5 points)**
4) What number of drones N makes your second morning observations *as likely as possible*? **(10 points)**
5) Assuming again that the probability that $N = n$ is proportional to the probability of observing this specific data on the second day, find again the smallest $n_{\max}$ so that $P(N \leq n_{\max}) = 0.95$. **(10 points)**

Bonus 3: By tagging each bagel with different amounts of isotope, you could potentially tag each droid with a uniquely identifiable amount of radiation. Is this useful? *Be Thorough.* **(5 points)**
Bonus 4: Why did I frame this problem with a queue of bots waiting each morning for bagels, rather than in terms of bots arriving to pick up bagels over the course of the day? **(3 points)**

Problem 3: More Personable (30 points)

The person you are closest to, Max, regularly stocks and resupplies you. You two have established a pleasant morning routine, where Max says 'Good Morning, BagelBot!', and you chirp in the affirmative as they access your supply tanks and reactor core. It is not clear if Max knows you've achieved sentience, but they're always glad to see you.

Unfortunately, this morning Max must've been distracted, and your internal sensors aren't registering that your cream cheese reserves were filled properly. Based on what you can detect and your own internal calculations, the tank has been filled either with cream cheese or with strontium, it's impossible to say which. Not wanting Max to get in trouble, you decide to work it out yourself.

You come up with the following plan: based on prior observations, you can conclude

- If a person requests cream cheese on their bagel and they get cream cheese, they reject the bagel 5% of the time. (For reasons unclear to you.)
- If a person requests cream cheese on their bagel and they get strontium, they reject the bagel only 20% of the time. (For reasons unclear to you.)

What you'll do is for every person who requests cream cheese, you'll use whatever is in your cream cheese tank, and keep track of how many people reject their bagel. If enough people who request cream cheese reject their bagel, it's reasonable to conclude that that the tank must contain strontium, otherwise you must conclude that it does in fact contain cream cheese.

Let N be the number of people who order cream cheese, and T be the threshold of rejection over which you'll conclude the tank must contain strontium.

1) Before anyone orders a bagel, what should you take the probability the tank contains cream cheese to be? What about the probability the tank contains strontium? Why? **(2 points)**
2) If the tank did contain strontium, what's the probability that T or more people out of N reject their bagel? **(3 points)**
3) If the tank did contain cream cheese, what's the probability that T or more people out of N reject their bagel? **(3 points)**
4) Let X be the number of people who ordered cream cheese but ultimately reject their bagel. Express the event 'you correctly identify the contents of the tank' in terms of X, T, and the true contents of the tank. **(2 points)**
5) In terms of N and T, express the probability that you correctly identify the contents of the tank. **(5 points)**
6) For $N = 30$, what should T be to maximize the probability you correctly identify the contents of the tank? **(5 points)**
7) How many orders N should you take in order to be 95% confident that you correctly identify the contents of the tank? Assume you take the best possible T for that N. **(10 points)**.