Group Project Two: Simulation

ECS 101: Introduction to Engineering and Computer Science (EECS Section)

Due: Friday, October 23, by end of day

Video Discussion: Please see the video before October 7 for important details.

Scenario: Imagine a beach that's 10km long, running east-west. A location on this beach refers to how far from the west end it is. For instance, the location 0km is at the west end, 10km is at the east end, and 5km is in the middle.

One company has already placed two coffee shops (or "cafés") along this beach. They're located at 4km and 8km. We are trying to determine where to place two cafés of our own. Call our cafés A1 and A2, and call our competitor's cafés B1 (the one at 4km) and B2 (the one at 8km).

Customers can come from anywhere along the beach. They will choose either A1, A2, B1, or B2 to buy their coffee. The probability that they select a café depends on:

- The distance they are from the café, and
- The price of the coffee at the café.

Your task is to figure out where to place each of your cafés, and how much money to charge for coffee at each of your cafés (it can be different at A1 and A2). The goal is to maximize your profit (total revenue minus total costs).

Data and Rules: There are four pieces of information that you need to work on this problem.

1. It costs you $2 to make a cup of coffee, so you should sell your coffee for more than this. The competitor sells coffee for $5 per cup at both B1 and B2. You can assume that the competitor won't change its prices in response to us.

2. Customers appear with uniform probability all across the beach. That means that a customer could appear anywhere from 0km to 10km (including at fractional points), and there's no point where customers are more likely or less likely to appear.

3. How does a customer choose a café? It is a two-step process.
    a. Compute the customer's <u>score</u> for each café. The score for café A1 is given as:

    score(A1) = (10 – (distance from customer to L(A1)) + 3*(6 – C(A1)),

    where L(A1) is the location of A1, and C(A1) is the cost of coffee at A1. Be careful to use the absolute value function for distance, because distance is not negative!

    For example, let's say that the customer appears at location 3km and that we put A1 at location 2km. Also, say that we charge $3.50 for coffee at A1. Then score(A1) = (10 – 1) + 3*(6 – 3.5). That simplifies to 9 + 7.5 = 16.5.

The scores for A2, B1, and B2 are computed the same way. Observe that the customer's score for a café gets higher as the distance for a customer gets smaller, and as the price of coffee at this café gets smaller.

b. Define total = score(A1) + score(A2) + score(B1) + score(B2). The probability that the customer chooses café X is simply score(X)/total.

4. How do you compute profit? You're going to randomly generate T trials (we specify T below). A trial generates a customer's starting point, finds the scores for each café, and chooses the customer's café as a result. If the customer chooses B1 or B2, you get no profit. If the customer chooses A1, then you get a profit of C(A1) – 2. If the customer chooses A2, then you get a profit of C(A2) – 2.

   After T trials, compute how much total profit you made, and divide by T. That's the average profit per customer you can expect, given your choice for L(A1), L(A2), C(A1), and C(A2).

Questions (Requirements):

Part 1: Create an Excel simulation for this problem, where you input L(A1), L(A2), C(A1), and C(A2), and where the simulation estimates the expected profit you make. You are allowed to put stores at the exact same location if you wish (for instance, A1 could be at 4km, just like B1). Prices can be any fraction you want.

Choose these four parameters however you wish, but your sheet must allow us to easily change the parameters (they should not be hard-coded in your formulas: make them easy to find and change). The Excel simulation should have 50,000 trials, and should output the expected profit, given the four parameters you specified. Again, make this expected profit very easy to find.

Part 2: What if we have only one café, A0, and our competitor has only one café, B0? We put A0 at location 0 and charge $4 for coffee. The competitor puts B0 at location 10 and charges $5 for coffee. Mathematically work out what our expected profit should be and verify that we get the same thing with your Excel simulation. (We've given hints in the video as how you would do this.)

Part 3: Now, code this same simulation in Python that you did in Excel for Part 1, and program it to run one million trials. The four parameters (L(A1), L(A1), C(A1), and C(A2)) should be stated at the beginning of your code so that we can easily change them. Choose the same values for the four parameters you selected in Part 1, and verify that you get (almost) the same profit that you got in that part.

Part 4: Finally, convert your code in Part 3 as a function. The <u>inputs</u> to your function should be L(A1), L(A2), C(A1), and C(A2). The <u>output</u> should be the expected profit. Write a program that explores several different combinations of those four parameters and tries to find the combination that maximizes profit. We are not expecting a program that will always return the best combination, but we want to see at least some thoughtful approach to finding a good combination of values.

Deliverables (What you turn in):

1. **Blackboard:**
    - A PDF or Word description of (a) how your algorithm works in Part 4, (b) your recommendation of the best values for L(A1), L(A2), C(A1), and C(A2) and the expected profit our company should get, and (c) a specific description of which team member performed which tasks. This should not exceed two pages, and one page is probably enough.

2. **GitHub:**
    - Excel file for Part 1.
    - Excel file for Part 2, and a PDF or Word document specifically explaining how you computed what the expected profit should be.
    - Python simulation code for Part 3.
    - Python code that maximizes profit in Part 4.

How you're being graded:

- Are the simulations (Excel & Python) correct in Parts 1-3? Did they follow the guidelines set forth above?
- Was the mathematical derivation in Part 2 correct? Did it show enough detail for me to follow it?
- Is there a plan for all of your team members to be involved in the project? Is it balanced and was it followed?

A perfectly executed project that does not involve your team members, giving them a chance to program or develop simulations in Excel, will not receive full credit. A team member that refuses to participate or is unresponsive to email communications will receive a lower grade than the remaining team members.

A CATME assessment will be filled out for this project after its completion to aid in this component of the grading.