Week 3: Probability review
Decision making

In this exercise, you will see your first application of probability in a decision making model. Let's first describe the problem you will study. Suppose you are presented with a number of possible choices, some better than others, but you do not know beforehand how good options you haven't yet seen are. All the information you have is the quality of previous options you rejected, the quality of the option being offered to you right now, and the number of options available. To ground discussion, suppose you are selling your car for which you are going to receive J offers. Say you get tired after that and just sell it—this is a somewhat unrealistic assumption but let us live with it for a while. Offers will have of different values and if all J offers were presented to you upfront, you would sell the car to the highest bidder. Unfortunately, potential buyers make their offers in a random order and if the offers are not accepted on the spot, they withdraw their bid—presumably, they can find a different seller willing to accept their offer.

Denote the rank of the n-th offer as $X_n = 1, 2, \ldots J$. If $X_n = 1$, then the best offer was made in round n; $X_n = 2$ implies that the second-to-best offer was made in round n; so in general, $X_n = i$ means that the i-th best offer was made in round n. Since offers are made randomly, all possible $J!$ rank orderings are equally likely and the probability of the i-th best offer occurs in round n is $1/J$.

Your strategy for deciding which offer to accept is the following:

- you reject the first K offers. You are not an earnest seller during this phase, you are just probing the market;
- you select the L-th best offer out of these K, which we denote as X_0. Note that $L \leq K$;
- having rejected the first K offers, you now choose the next offer that exceeds X_0. That is, you choose the first X_n, $n > K$, for which $X_n < X_0$. Denote the accepted offer as $X^\star$;
- if you reach the last offer J, you become desperate and accept offer X_J.

A Simulate an individual experiment. Write a MATLAB function that takes as inputs the number of offers J, the number of rejected offers K, and the selection constant L. The function returns the rank of the accepted offer $X^\star$ and the round n in which it was accepted.

B Probability distribution of the rank of accepted offers. Start by fixing the selection constant to $L = 1$. Notice that the strategy now simplifies to selecting the first offer better than all of the first K offers. Also, fix the total number of offers to $J = 50$ and the number of rejected offers at $K = 30$. Using the function from part A, estimate the probability mass function of the rank of the selected offer $X^\star$. That is, estimate $\mathbb{P}[X^\star = j]$ for $j = 1, 2, \ldots, J$. Repeat the this procedure to display the pmf of $X^\star$ for $L = 2$ and $L = 5$. You should be pleasantly surprised that this decision strategy actually works reasonably well.

C Probability of accepting the best offer as a function of K. We are now interested in the probability of selecting the best offer, i.e., $\mathbb{P}[X^\star = 1]$. Fix the number of offers to $J = 50$. Start with $L = 1$ and estimate the probability $\mathbb{P}[X^\star = 1]$ as a function of the number of rejected offers K for K varying between 1 and $J - 1$. Repeat this procedure for $L = 2$ and $L = 5$. You should again be pleasantly surprised that the probability of selecting the best offer can be made quite high.

D Probability of accepting the last offer. Fix $L = 1$. We can think of the probability of selecting the last offer as the probability that the decision policy fails, since it drove you to a desperate decision. While we could use a simulation to estimate this probability it is possible (and therefore preferable) to evaluate it analytically.

Note that we seek the probability $\mathbb{P}[X^\star = X_J]$. So first, ask yourself: when do we select the last offer? We select X_J if all offers made after the K-th round are worst than X_0. Since X_0 is the best offer ($L = 1$) among the first K ones, we select will only select X_J if one of two things happened: (i) if the best offer was one of the first K ones—since in this case, we end up setting $X_0 = 1$ and rejecting all other offers; (ii) if the last offer is the best offer ($X_J = 1$) and the second to best offer is one of the first K ones ($X_0 = 2$). This argument is enough to let you establish that

$$\mathbb{P}[X^\star = X_J] = \frac{K}{J} + \frac{1}{J} \times \frac{K}{J-1}. \tag{1}$$

Explain.

E Probability of accepting best offer. The probability of selecting the best offer, $\mathbb{P}[X^\star = 1]$ can also be computed analytically when $L = 1$. Your goal is to prove that

$$\mathbb{P}[X^\star = 1] = \frac{K}{J} \sum_{i=K+1}^{J} \frac{1}{i-1}. \tag{2}$$

As is typically the case in complicated probability problems, the easiest way to obtain (2) is to use total probability to decompose the event of interest in simpler ones. So we suggest that you start by conditioning on $X_n = 1$ for each n, i.e., that you condition on the probability of the best offer being made in the n-th round. Then, compute the conditional probability

$$\mathbb{P}[X^\star = 1 \mid X_n = 1], \tag{3}$$

that is, the probability of selecting the best offer given that the best offer is made at round n. Think about what this probability is when $n \leq K$—i.e., when the best offer is made among the ones you rejected—and when $n > K$. Finally, you can use total probability to go from (3) to (2).

F Optimal number of rejected offers K. Again, fix $L = 1$. Note that the expression in (3) is a function of K. Let us then use it to find the optimal number of rejected offers K, i.e., the number K that maximizes our probability of selecting the best offer. To doing so, you should use the approximation

$$\sum_{i=K+1}^{J} \frac{1}{i-1} \approx \int_{i=K}^{J-1} \frac{1}{x} dx. \tag{4}$$

Using (4), establish that the optimal K is approximately J/e, where e is the base of natural logarithms. In other words, you should probe the market for roughly a third of the offers. Furthermore, show that for this value of K, the probability of selecting the best offer is approximately $1/e \approx 0.37$. This might explain your pleasant surprises in parts B and C. Does it?