A BAYESIAN PERSUASION APPROACH TO RESTAURANT WAIT TIME REPORTING IN AN ONLINE SETTING

SPENCER COMPTON AND FRANKLYN WANG

Abstract. In this problem, we apply the Bayesian Persuasion model of [3] to restaurant wait time reporting, specifically in the context of apps like Yelp. A natural guess is that the optimal signalling scheme consists of reporting intervals of the wait time, like the result in [1]. We confirm that this is true in the case with one customer; however, previous work ([2]) shows that this is no longer true when there are two customers. In light of this result, we discretize the wait times, converting the problem to a problem in linear programming, and then we use learning theory to bound the loss of our signalling scheme. Finally, we propose a possible fix to the issue of lacking monotonic signalling structure. Specifically, we *regularize* the utilities of the customers, which causes the problem to fall into the model considered by [5].

‘

1. Introduction

In an increasingly digitized world, publishing of wait times by restaurants has become standard behavior to help customers make better informed choices. Consider the following, taken from Yelp.

While big data may help one predict the wait time accurately, we should first ask ourselves, are the companies incentivized to follow this method? This is not as easy as it appears at first glance: consider the following example.

1.1. An Illustrative Example.

Assume that we have a customer whose utility from going to McDonald's is 6.5 minus the wait time in minutes, and that the customer maximizes expected utility. Assume the wait time is equally likely to be 5 or 7 minutes. Then the restaurant does not want the customer to know the true wait time, because if the wait time is 7 minutes the customer will not go, but if the customer does not know the true value he will go every time, because his expected utility is positive.

Now assume that we have another customer whose utility from going to McDonald's is 5.5 minus the wait time in minutes, and assume the same utility as before. Then the restaurant does want the customer to know the true wait time, because the customer's expected utility is negative but the restaurant really wants the customer to know if the wait is five minutes. Intuitively, if telling you some information makes you come, there is no reason to dissuade you from coming by telling you more information.

In the second case, the optimal strategy is not to give the customer the true wait time. Instead, the customer should be told that the wait time is 5 minutes whenever it actually is 5 minutes, as well as just under 1/3 of the time when the wait time is 7 minutes. This way, someone visits just under 2/3 of the time, as opposed to 1/2 of the time. In a similar vein, the optimal behavior for a restaurant is to merely to find some time t such that the expected utility of the customer is zero given that they condition on the wait time being below t (follows from lemma 5.1).

2. Related Work

The study of sender-receiver games began with [1]. In Crawford's model, known as *cheap talk*, the sender can send signals to the receiver, but the signals do not carry strategic value. In our model, following [3], when the sender sends signals to the receiver, the sender must pre-commit to a signalling scheme. If player A chooses signalling schemes in "signalling space" then player B chooses best responses. The difference is that in cheap talk, the game is simultaneous; player A chooses possible responses to signals, whereas player B chooses some action given the signal. However, in Bayesian persuasion, the game is sequential; player B knows the signalling scheme that player A chose.

Another type of problem Bayesian persuasion has seen applications to is problems in teaching. For example, is "teaching to the test" a good policy? This question is addressed by [4]. If one has unmotivated students, then teaching to the test at least encourages the students to learn *something*, as opposed to nothing. However, if one has motivated students, not teaching to the test will cause the students to learn everything. In particular, the optimal solution is to induce a belief in the students that they are indifferent between studying a given piece of information and not studying a piece of information; the goal is to maximize the amount of information for which students are indifferent, and to tell the students that the rest will for sure not appear on the test.

In general, sender-receiver games can be solved by a method known as concavification, where we look at the smallest concave function which is weakly greater than the payoff function over all of signal-space. [7] shows that the optimal payoff is always the concavication evaluated at the prior. Nonetheless, signal space is big, and this intuition is not useful. It will, however, become useful when we discretize the space of wait times.

Bayesian Persuasion is generally tractable in a few contexts; one in which the action space is continuous, one in which the action space has size two, and one in which the payout to the receiver depends only on the mean of the posterior distribution. The last is somewhat tricky because in general a group of posteriors can be the outputs of the signals as long as the expected value of the posteriors is the prior. But a group of posterior means cannot be the outputs of the signals as long as the expected value of the posterior means is the prior mean! For a full analysis of this case, we direct the reader to [2].

3. Our Model

There is one restaurant, restaurant R. There are N customers $c_1, c_2, \dots c_n$, and customer c_i has utility function $u_i(t)$, where t is the wait time of the restaurant. If customer c_i knows that the restaurant's signalled wait time is X, customer c_i will go iff

$$E_{t \sim X}[u_i(t)] \geq 0.$$

We assume further that u_i satisfies a non-crossing condition; specifically, if $u_i(x) > u_j(x)$ for some value of x, then $u_i(x) > u_j(x)$ for all values of x. What this condition implies is that there exists an ordering of the customers $c_1, c_2, \dots c_n$ so that c_i going to the restaurant implies that $c_1, c_2, \dots c_{i-1}$ all go to the restaurant.

We assume that the restaurant signals via the Bayesian persuasion framework, as described in [3]. The restaurant and all customers know that the wait time is drawn from a prior distribution $t \sim T$. Furthermore, the restaurant knows the realized waiting time t_0. Assume that T is supported on a discrete subset of $\mathbb{R}^+$. There is a set of signals S. The *signalling function* $\varphi : T \to S$ maps the waiting time to a signal.

After receiving a signal s, each customer c_i then decides if

$$E_{t \sim X|s}[u_i(t)] \geq 0$$

and then decides whether or not to go to the restaurant. The utility of the restaurant comes solely where the customer goes to the restaurant or not; importantly, it does not depend on the wait time at all. In fact, we will assume that the utility of the restaurant is additive, and they gain one from a customer coming and zero otherwise.

We use a merging result of [3]. Specifically, signals matter insofar as they induce posteriors (the distribution of true wait times conditioned upon receiving a specific signal).

4. Relating Required Signals to Action Space

Lemma 4.1. *If there are k customers, there exists a Sender-Preferred Subgame Equilibrium where $|S| \leq k+1$.*

Proof. Consider the Sender-Preferred Subgame Equilibrium with the smallest possible signal space. Note that, as a result of the aforementioned non-crossing condition, there exists an ordering of the customers $c_1, c_2, \ldots c_n$ so that c_i going to the restaurant implies that $c_1, c_2, \ldots c_{i-1}$ all go to the restaurant. As a result, the only possible outcomes are where a prefix of this list go to the restaurant. Thus, there are at most $k+1$ possibilities (the lengths of all possible prefixes). In this sense, we say the customer action set is bounded by $k+1$. Assume, for sake of contradiction, that there exists two signals S_1 and S_2 such that they map to the same customer actions. We could create a new signal S' and use it whenever the original signalling strategy would use S_1 or S_2. Clearly it is still optimal for consumers to complete the same actions, and the sender does not have any decrease in utility. Since this causes the signal space to decrease, it causes a contradiction. By pigeonhole principle, such a conflict of two signals occurs whenever the signal space is larger than $k+1$. Thus, there exists a Sender-Preferred Subgame Equilibrium where $|S| \leq k+1$. □

This aids us in determining the structure of optimal signalling strategies and making computationally efficient algorithms to optimize a restaurant's signalling strategy.

5. Structure of Optimal Signalling Strategies

Often, an understanding the structure of equilibria makes designing computational solutions more approachable. We examine the structure of optimal signalling strategies for the special case where there is only one customer.

Lemma 5.1. *There exists a Sender-Preferred Subgame Equilibrium such that the support of S_1 contains no time strictly smaller than a time in the support of S_2.*

Proof. Consider an optimal signalling strategy in a Sender-Preferred Subgame Equilibrium which uses two signals S_1 and S_2. We know that there must exist an optimal signalling strategy using at most two signals by lemma 4.1. If the receiver does the same action for all the given signals, then we can trivially make the lemma true by always using S_1 and we will not decrease the sender's utility. Otherwise, without loss of generality let S_1 be the signal where the customer decides not to go to the restaurant, and S_2 be the signal where the customer decides to go to the restaurant. Suppose there exists a time i in the support of S_1 and j in the support of S_2 such that $i < j$. We can describe our signalling strategy in a way such that $P_{A,B}$ denotes the probability that the true wait time is A and we report signal B. We will modify our signalling strategies and denote the new probabilities in the form $P'_{A,B}$. We can set all probabilities in P' the same as P, except set $P'_{i,S_1} = P_{i,S_1} - \min(P_{i,S_1}, P_{j,S_2})$, $P'_{j,S_1} = P_{j,S_1} + \min(P_{i,S_1}, P_{j,S_2})$, $P'_{i,S_2} = P_{i,S_2} + \min(P_{i,S_1}, P_{j,S_2})$, and $P'_{j,S_2} = P_{j,S_2} - \min(P_{i,S_1}, P_{j,S_2})$ such that the utility of the sender does not decrease. This is because we are essentially swapping i and j between S_1 and S_2. Through this, the probability of using S_2 does not decrease and the expected wait time when the sender signals S_2 decreases. As such, the probability of the customer going does not decrease, and the utility of the sender does not decrease. Thus, using this exchange argument repeatedly applied, we show that our lemma holds. □

With this structure proven, computing a solution for the two-person scenario is computationally simple. We can simply binary search for the largest suffix of the waiting time distribution such that the customer goes to the restaurant. However, the same cannot be said for cases with more consumers. In cases with larger action sets, there exist examples within the Bayesian Persuasion model where there exists no optimal signalling strategy in the form of contiguous ranges [2].

6. Optimizing for N Potential Customers

Given that with a large number of customers, we no longer have a favorable structure of optimal solutions to exploit, we need a new approach. In this section, we make the assumption that utilities of customers are linear functions; by this, a customer chooses to go to the restaurant iff the expected wait time is below some threshold.

We utilize linear programming to determine an optimal signalling strategy to maximize utility for the sender. We will assume, like previously, that for each customer we know the cutoff expected wait time at which they will change whether or not they go to the restaurant. Thus, we know $x_0, x_1, \ldots x_N$ where x_i denotes the expected wait time such that if the expected wait time is $\leq x_i$ then i people will go. Additionally, let T_i denote the value of the i-th element of T, the set of possible wait times. Similarly, G_i denotes the probability that the wait time is T_i. All x_i, T_i, and G_i are constants in our linear program. Additionally, we will have states $S_0, S_1, \ldots, S_N$ where signal S_i denotes a signal where, upon receiving this, i customers are expected to come to the restaurant. Our signalling strategy will be described by a series of variables $p_{i,j}$ which represents the probability that signal i is used and the wait time is T_j. We will use variables h_i to denote the probability that we use signal i. Finally, we will use variables w_i to represent h_i multiplied by the weighted average of wait times conditioned up signal i being used. We use the following constraints:

$$\text{maximize}: \frac{1}{N} \times h_1 + \frac{2}{N} \times h_2 + \cdots + \frac{N}{N} \times h_N \tag{6.1}$$

$$\forall i \in [N] : h_i = \sum_{j=1}^{|T|} p_{i,j} \tag{6.2}$$

$$\forall j \in [|T|] : \sum_{i=0}^{N} p_{i,j} = G_j \tag{6.3}$$

$$\forall i \in [N], j \in [|T|] : p_{i,j} \geq 0 \tag{6.4}$$

$$\forall i \in [N] : w_i = \sum_{j=1}^{|T|} p_{i,j} \times v_j \tag{6.5}$$

$$\forall i \in [N] : w_i \geq x_i \times t_i \tag{6.6}$$

Lemma 6.1. *This linear program produces an optimal signalling strategy for the sender.*

Proof. The linear program works to maximize the expected number of customers that go to the restaurant (shown in equation 6.1). Equation 6.2 enforces that all $p_{i,j}$ conform with the chosen value of h_i, equation 6.3 enforces that all $p_{i,j}$ conform with the given value of G_j, and equation 6.4 enforces that all probabilities are non-negative. Equation 6.5 enforces the definition of all w_i. Finally, equation 6.6 enforces that for every signal i that is used with positive probability the average wait time is at least the minimum required for the customers to listen to the signal. As such, the linear program enforces the relevant constraints and optimizes for the sender's utility. □

While this linear program computes an optimal signalling strategy for the sender when all the customer's cutoff points are known, the case where we do not know all the cutoff points is interesting and applicable to the real-world as well.

We will utilize our linear program as a form of empirical risk minimization given a sample of customer cutoff points. We utilize Rademacher complexity to bound for generalization error to bound the error of our empirical risk minimization:

$$L_p(h) - L_s(h) \leq 2\,\mathrm{Rad}(F \circ S) + \epsilon$$

with high probability [6]

(The loss function is given by the probability the customer goes to the restaurant under a given signalling scheme, where in L_p the customer is randomly drawn from the population and in L_s the customer is randomly drawn from a sample of size m.) Since our error is in range $[0, 1]$, we use this version of Massart's Lemma to bound:

$$\mathrm{Rad}(F \circ S) \leq \sqrt{\frac{2 \log |H|}{m}}$$

[6]

Where $|H|$ is the size of our hypothesis class. Since the solutions to our linear program must fit inside a floating-point number, we will act as if it is discretized and we can store $|P|$ distinct values in range $[0, 1]$. If we have a hypothesis space of $|P|^{N|T|}$ (which contains the optimal hypothesis following lemma 4.1), then we apply our inequalities to get that we can have $\leq \varepsilon$ error if we use m samples where:

$$m \geq \frac{8N|T| \log |P|}{\varepsilon^2}$$

Which does not really help us, since originally we could simply sample all N customers and have an optimal signalling scheme. However, we can make the following observation:

Lemma 6.2. *There exists a signalling scheme in the hypothesis space where we only have* $\frac{1}{\varepsilon}$ *signals corresponding to actions where* $0, \varepsilon N, 2\varepsilon N, \ldots, N$ *customers come to the restaurant and the expected probability of a customer coming is at most* ε *less than the optimal signalling scheme with* $N+1$ *signals.*

Proof. Consider the optimal signalling with $N+1$ signals. In this scheme we know that signal S_i corresponds to a signal where i customers will come. We remap this to a signalling scheme S' which contains signals $S'_0, S'_{\varepsilon N}, S'_{2\varepsilon N}, \ldots, S'_N$. We will use the signal $S'_{i\varepsilon}$ whenever our original scheme would use any signals in range $[S_{i\varepsilon}, S_{(i+1)\varepsilon-1}]$. Note that, since we "round down" signals, customers will always still want to listen to our new signalling scheme S'. Let $\mathbb{P}[S_i]$ denote the probability of using signal S_i under the original scheme. The expected probability that a random customer comes under the old signalling scheme S was $\mathbb{P}[S_0] \times 0 + \mathbb{P}[S_1] \times \frac{1}{N} + \mathbb{P}[S_2] \times \frac{2}{N}, \ldots, \mathbb{P}[S_N] \times 1$. The expected probability that a random customer comes under the new scheme is $(\mathbb{P}[S_0] + \cdots + \mathbb{P}[S_{\varepsilon N-1}]) \times 0 + (\mathbb{P}[S_{\varepsilon N}] + \cdots + \mathbb{P}[S_{2\varepsilon N-1}]) \times \varepsilon + \cdots + (\mathbb{P}[S_{N-\varepsilon N} + \cdots + \mathbb{P}[S_N]) \times (1-\varepsilon)$. Since the coefficient of every $\mathbb{P}[S_i]$ decreases by at most ε, the total decrease is at most ε, and our statement is true. □

Thus, we can use a modified version of our linear program as a method of empirical risk minimization to optimize over the class where we have such $\frac{1}{\varepsilon}$ signals. We can simply remove all h_i in equation 6.1 where i is not a multiple of εN and is not N. Our new hypothesis space to find an optimal such signalling scheme is $|P|^{|T|/\varepsilon}$. Thus, we can use generalization error and Massart's lemma again to show that we can have $\leq \varepsilon$ error if we use m samples such that:

$$m \geq \frac{8|T| \log |P|}{\varepsilon^3}$$

If instead we use $\frac{2}{\varepsilon}$ signals (so that we are within $\varepsilon/2$ of the optimal signalling scheme with N signals by lemma 6.2), and use $m \geq \frac{4|T| \log |P|}{(\varepsilon/2)^3}$, we will be within ε error of the optimal signalling scheme by triangle inequality. Thus, we can do this and have a sampling complexity of $O(\frac{|T| \log |P|}{\varepsilon^3})$! In practice, $|T|$ and $\log |P|$ are effectively constants fairly limited in size, so this is a very strong bound.

7. Obtaining Monotonic Signal Structure

In [5], Mensch derives sufficient conditions (notably d-quasisupermodularity) for monotonic signal structure when the action space is continuous. One naive way to make the signal structure in this problem continuous is by allowing the customer to choose a probability p for going to the restaurant; however, this lacks interior optima. Instead, we can add a concave regularizer to the expected utility, which will yield interior optima. We have not fully explored this, but it would be interesting to attempt to verify d-quasisupermodularity for the regularized version of the function.

8. Conclusion and Applications

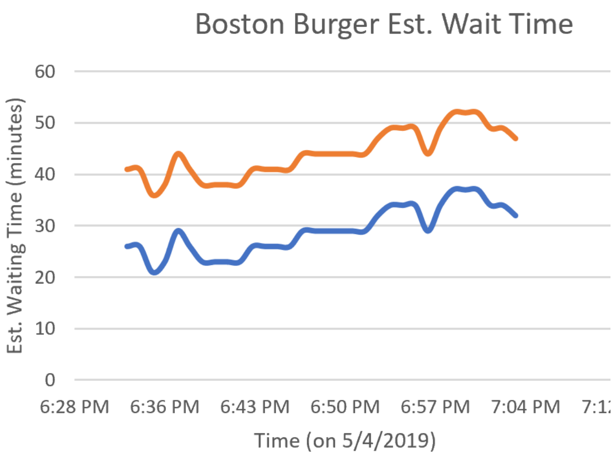
The optimal outcome in many cases is a contiguous signalling scheme, where for each signal s we tell the same thing to the customer regardless of where we are in that range. Assuming that the true prior of wait time is given by X_t, and the restaurant knows the exact wait time, and further assuming that X_t does not change substantially over time, we should expect that only the same ranges will be reported over and over again. For example, the following behavior should not happen:

When the true wait time is 15, report $[5, 25]$. When the true wait time is 25, report $[15, 35]$. This does not work, because $[5, 25]$ should intersects $[15, 35]$.

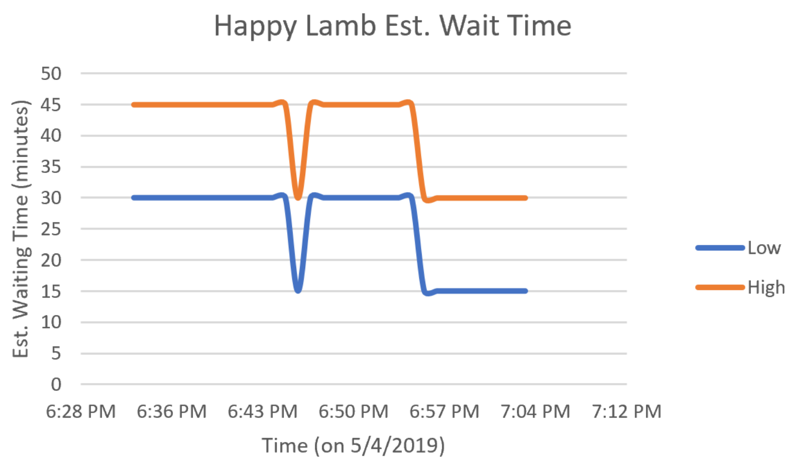
Now we ask, how does this work in the real world? Yelp provides wait time intervals (confidence not specified). Over a small period of time, how do their wait time estimates change?

We did an experiment for this. Below is evidence of qualitatively different behavior between two restaurants.

When we contacted restaurants who use Yelp's wait time estimation algorithm, we obtain that Yelp makes these estimates with machine learning. Yelp is given data by the restaurants, and then Yelp makes the predictions. Yelp's incentives are likely a mix of wanting to have accurate predictions, so that customers will use the app, as well as bringing in revenue for restaurants who use the app, so that they can acquire more restaurants as customers. Yelp declined to comment on their wait time estimation algorithm, stating that all relevant details are proprietary information.



(A)



(B)

Figure 1. a shows the wait time for Boston Burger and b shows the wait time for Happy Lamb. Happy Lamb seems to follow the theoretically optimal strategy, whereas Boston Burger appears to be reporting true wait times.

References

[1] Vincent P Crawford and Joel Sobel. Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451, 1982.

[2] Matthew Gentzkow and Emir Kamenica. A rothschild-stiglitz approach to bayesian persuasion. *American Economic Review*, 106(5):597–601, 2016.

[3] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[4] Edward P Lazear. Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics*, 121(3):1029–1061, 2006.

[5] Jeffrey Mensch. Monotone persuasion. *Available at SSRN 3265980*, 2018.

[6] Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

[7] Yun Wang. Bayesian persuasion with multiple receivers. *Available at SSRN 2625399*, 2013.