“How Bayesian Persuasion can Help Reduce Illegal Parking and Other Socially Undesirable Behavior” by Penélope Hernández and Zvika Neeman

Han Wang

November 28, 2020

Outline

- An application of Bayesian persuasion: the city can allocate enforcement resourses across different neighborhoods and send messages to deter undesirable behavior, for example, illegal parking.
- Two approaches for addresssing this problem:
  1. recasting it as a linear programming problem
     - We can identify messages with the set of neighborhoods on which they achieve deterrence (sending recommendations of actions).
     - We only need to consider a certain type of allocation rule: the city allocates its enforcement resources proportionally to the deterrence thresholds.
  2. dividing it into two sub-problems: knapsack problem + Baysian persuasion
     - Each neighborhood receives the same message and their action depends on the posterior's mean (expected enforcement resources available).
     - Find the optimal “splitting” of the prior mean into a distribution of posterior's means.

Model

- A city with $N \geq 1$ different neighborhoods.
- The city determines the amount of enforcement resourses in each neighborhood out of the total amount of available resourses r.
  - The distribution of r is r_k, $k \in \{1, ..., K\}$, with probability π_k respectively.
  - We refer to k as the state of the world. The city and drivers have common prior. The city knows the realization of the state, but drivers do not.
- The city can send a message $m \in \{1, ..., M\}$ about the state of the world k.
  - "Public signals": The city sends the same message m to all neighborhoods.
  - A $K \times M$ matrix of information structure: The probability that the city sends message m in state k is $p_k(m) = Pr(m|k)$.
  - The unconditional probability of receiving message m is $Pr(m) = \sum_{k=1}^{K} p_k(m)\pi_k$.
  - The posterior belief upon receiving message m is $Pr(k|m) = \frac{p_k(m)\pi(k)}{\sum_{k'=1}^{K} p_{k'}(m)\pi(k')}$.
  - The expected resources available conditional on message m is $r(m) = \sum_{k=1}^{K} r_k Pr(k|m)$.
- Denote the amount of resources allocated to enforcement in neighborhood i in state k when the city sends the message m by $a_k^i(m)$.
  - $a_k^i(m) \equiv 0$, if message m is sent with probability zero in state k.
  - Resource constraint:
    In each state $k \in \{1, ..., K\}$, $\sum_{i=1}^{N} a_k^i(m) \leq r_k$ for every message m.
  - Deterrence threshold:
    Define $a^i(m) \equiv \sum_{k=1}^{K} a_k^i(m)Pr(k|m)$ as the expected resources in neighborhood i upon receiving message m.
    The measure of illegal parking in neighborhood i is a threshold function

$$q^i(a^i(m)) = \begin{cases} 1 & \text{if } a^i(m) < \tau^i \\ 0 & \text{if } a^i(m) \geq \tau^i \end{cases}$$

	m_1	m_2
π_1	1	0
π_2	p	$1-p$
π_3	0	1

(a) The information structure

π_1	$a_1^1(m_1)$	$a_1^2(m_1)$	$a_1^3(m_1)$	r_1
π_2	$a_2^1(m_1)$	$a_2^2(m_1)$	$a_2^3(m_1)$	r_2
	$a_2^1(m_2)$	$a_2^2(m_2)$	$a_2^3(m_2)$	
π_3	$a_3^1(m_2)$	$a_3^2(m_2)$	$a_3^3(m_2)$	r_3
	τ^1	τ^2	τ^3	

(b) The allocation of resources

- Denote the social disutility generated by illegal parking in neighborhood i by s^i.
- The city's objective is to allocate the amounts of enforcement resources and send messages with probabilities so as to minimize the expected social cost of illegal parking

$$\min_{\left\{a_k^i(m)\right\},\{p_k(m)\}} \sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{N} q^i\left(a^i(m)\right) s^i p_k(m)\pi_k$$

subject to the constraint of information structure and the resource constraint

(1) $p_k(m) \geq 0$ and $\sum_{m=1}^{M} p_k(m) = 1$ for every k

(2) In each state $k \in \{1, ..., K\}$, $\sum_{i=1}^{N} a_k^i(m) \leq r_k$ for every message m

First Approach – Linear Programming

- We can think messages as directly instructing receiver's action on whether illegal parking occurs in neighborhood i, i.e., $M = 2^{\{1,...,N\}}$, with $\emptyset \in M$.

- Denote the set of neighborhoods on which each message m achieves deterrence by $S(m) \subseteq \{1, ..., N\}$. Deterrence here means $i \in S(m) \iff a^i(m) \geq \tau^i$.

- **Proposition** Given $\{p_k(m)\}$ and an allocation $\{a^i_k(m)\}$ that satisfy feasibility constraints (1) and (2), the same probabilities together with the allocation $\{(a^i_k)^*(m)\}$ achieves equal or better deterrence than $\{a^i_k(m)\}$.

- Definition: $\{(a^i_k)^*(m)\}$ is a "ratio" rule.

  – For every state k, for every message m that is sent with a positive probability at k, and for every neighborhood $i \in S(m)$,

  $$\left(a^i_k\right)^*(m) = \frac{\tau^i}{\sum_{j \in S(m)} \tau^j} \cdot r_k$$

  – and for every location $i \notin S(m)$, or messages m that are sent with probability zero,

  $$\left(a^i_k\right)^*(m) = 0$$

- This proposition says given an information structure $\{p_k(m)\}$,
  $a^i(m) \equiv \sum_{k=1}^{K} Pr(k|m) a^i_k(m) \geq \tau^i \implies \sum_{k=1}^{K} Pr(k|m)(a^i_k)^*(m) \geq \tau^i$.

- Allocation according to this optimal ratio rule minimizes unnecessary waste of enforcement resources.

Recall:
$\min_{\left\{a^i_k(m)\right\},\{p_k(m)\}} \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{i=1}^{N} q^i\left(a^i(m)\right) s^i p_k(m)\pi_k$, s.t.,
(1) $p_k(m) \geq 0$ and $\sum_{m=1}^{M} p_k(m) = 1$ for every k
(2) In each state $k \in \{1, ..., K\}$, $\sum_{i=1}^{N} a^i_k(m) \leq r_k$ for every message m.

This problem can be rewritten as:
$\min_{\{p_k(m)\}} \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{i \in \{1,...,N\} \setminus S(m)} s^i p_k(m)\pi_k$, s.t.,
(1) $p_k(m) \geq 0$ and $\sum_{m=1}^{M} p_k(m) = 1$ for every k
(2^*) $\sum_{k=1}^{K} p_k(m)\pi(k) \left(a^i_k\right)^*(m) \geq \tau^i \sum_{k=1}^{K} p_k(m)\pi(k)$ for every m and $i \in S(m)$.

Note: $S(m)$ is determined through the condition (2^*).

$$i \in S(m) \iff \sum_{k=1}^{K} Pr(k|m)(a^i_k)^*(m) \geq \tau^i \iff \sum_{k=1}^{K} \frac{p_k(m)\pi_k}{\sum_{k=1}^{K} p_k(m)\pi_k}(a^i_k)^*(m) \geq \tau^i$$

We have a linear programming exercise with the size of input $K \times 2^N$. It is computationally hard with large N.

Second Approach – Splitting

- **Knapsack problem without persuasion**

  Define $D_S(r) = \begin{cases} \sum_{i\in\{1,\ldots,N\}} s^i & \text{if } r < \sum_{i\in S}\tau^i \\ \sum_{i\in\{1,\ldots,N\}\setminus S} s^i & \text{if } \sum_{i\in S}\tau^i \le r \end{cases}$ and $D(r) = \min_{S\subseteq\{1,\ldots,N\}} D_S(r)$.

  The city's payoff $D(r)$ is a non-increasing step function.

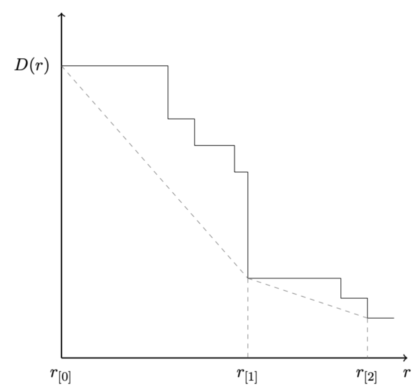

Figure 2: $D(r)$ and the convexification of $D(r)$ from below

- **Baysian persuasion problem**
- The goal is to split $E[r]$ into a distribution of $r(m)$, $m \in M$.
  - $E[r] = \sum_{m=1}^{M} Pr(m) \cdot r(m)$, implied by the fact that the mean of posteriors is prior.
  - By Blackwell & Girshick (1954), the distribution of posterior's mean is a mean-preserving contraction of (second-order stochastically donimated by) the prior distribution.
- Main take-away: Baysian plausibility implies that two messages may be not enough.

- An example
  - Suppose there are 3 states.
  - $(r_1, r_2, r_3) = (0, \frac{1}{2}, 1)$ with prior $(\pi_1, \pi_2, \pi_3) = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. $E[r] = \frac{1}{2}$.
  - Suppose the city has 2 neighborhoods with thresholds $\tau^1 = \frac{1}{2}$ and $\tau^2 = 1$ and social disutilities $s^1 = \frac{1}{4}$ and $s^2 = 1$.

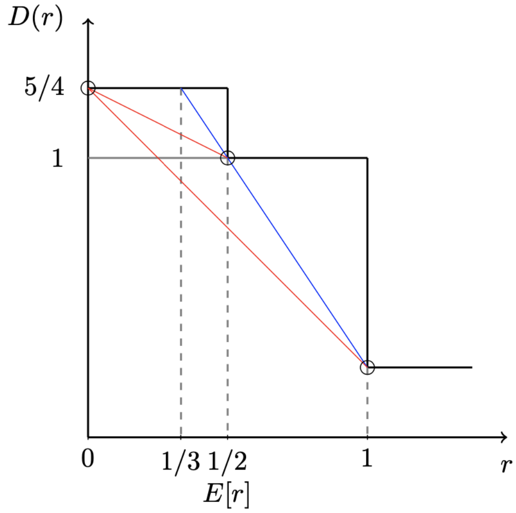

Figure 3: Three messages are better than two

  - The optimal deterrence with two messages L and H requires that $r(H) = 1$ and $r(L)$ as low as possible.

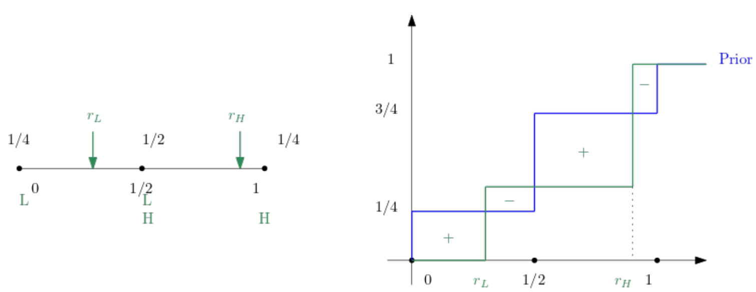

Figure 4: Distribution of posterior's mean

  - The lowest possible value of r_L depends on r_H and is given by $\max\left\{\frac{3r_H - 2}{8r_H - 5}, 0\right\}$.
    - $1/2 < r_H \leq 2/3$: $r_L < 1/2$ is unrestricted;
    - $2/3 < r_H < 1$: the lowest possible value of r_L is increasing in r_H;
    - $r_H = 1$: then the lowest possible value of r_L is $1/3$.
  - $r(L) = 1/3$. Because $E[r] = \sum_{m=1}^{M} Pr(m) \cdot r(m)$, $Pr(L) = 3/4$ and $Pr(H) = 1/4$.
  - With two messages: $\frac{3}{4} \cdot \frac{5}{4} + \frac{1}{4} \cdot \frac{1}{4} = 1$.
  - With three messages that reveals the state of the world: $\frac{1}{4} \cdot \frac{5}{4} + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot \frac{1}{4} = \frac{7}{8} < 1$.