Optimal Selling and Learning

Christopher R. Dance Onno R. Zoeter

July, 2011

Learning while selling

Contribution

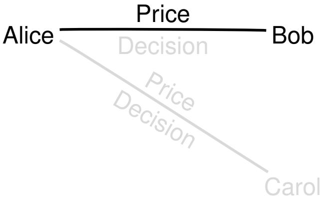
Learning by lottery: a way to use lotteries to learn more from each customer interaction.

Discussion

1. Learning from strategic agents requires more than machine learning.
2. How should strategy proof solutions be adapted if full rationality does not hold?

Learning while selling

0 cat for	0$
1 cat for	.25$



Contribution

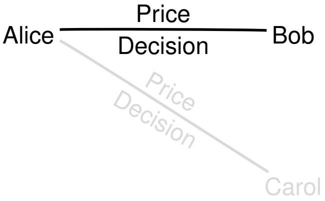
Learning by lottery: a way to use lotteries to learn more from each customer interaction.

Discussion

1. Learning from strategic agents requires more than machine learning.
2. How should strategy proof solutions be adapted if full rationality does not hold?

Learning while selling

0 cat for	0$
1 cat for	.25$



Contribution

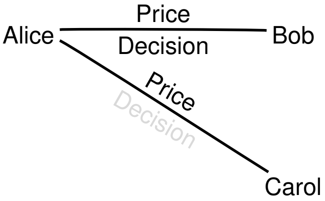
Learning by lottery: a way to use lotteries to learn more from each customer interaction.

Discussion

1. Learning from strategic agents requires more than machine learning.
2. How should strategy proof solutions be adapted if full rationality does not hold?

Learning while selling

0 cat for	0$
1 cat for	.50$



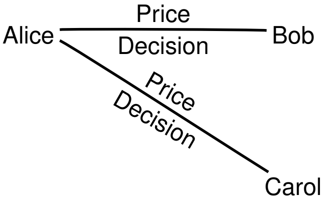
Contribution

Learning by lottery: a way to use lotteries to learn more from each customer interaction.

Discussion

1. Learning from strategic agents requires more than machine learning.
2. How should strategy proof solutions be adapted if full rationality does not hold?

0 cat for	0$
1 cat for	.50$



Contribution

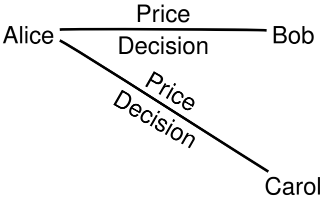
Learning by lottery: a way to use lotteries to learn more from each customer interaction.

Discussion

1. Learning from strategic agents requires more than machine learning.
2. How should strategy proof solutions be adapted if full rationality does not hold?

Learning while selling

0 cat for	0$
1 cat for	.50$



Contribution

Learning by lottery: a way to use lotteries to learn more from each customer interaction.

Discussion

1. Learning from strategic agents requires more than machine learning.
2. How should strategy proof solutions be adapted if full rationality does not hold?

Setting

- Buyer at t purchases the item if buying maximizes his utility

$$u_t := \underbrace{v_t}_{\text{valuation}} - \underbrace{p_t}_{\text{price}} \geq 0 \; .$$

- v_t is drawn from a valuation density $\mathbb{P}(v|\theta)$ with parameter θ.
- Production cost c.

For ease of exposition in first part of talk

- Buyers arrive once.
- Valuations are drawn iid.
- There is only one item to sell.

Optimal selling when θ is known

If seller has perfect market knowledge ($\mathbb{P}(v|\theta)$, and θ known) profit maximizing price is

$$p^* = \arg\max_p (p - c)\,\mathbb{P}(v > p)\ .$$

This is the standard setting in economics, we want to learn θ.

- When buyers come only once they will act *myopically*.
  E.g. No strategic rejections to attempt to lower future prices.
- Seller learns from buyers' decisions:
  $v_t \geq p_t$ if buyer bought,
  $v_t < p_t$ otherwise.

Seller's problem is a partially observable Markov decision problem (POMDP) where

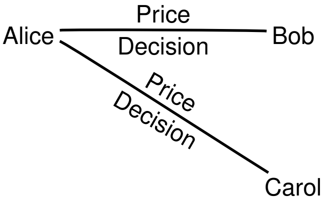
- actions are continuous, and
- Bayesian updates $p(\theta | v > p)$ might not have simple form.

Drawbacks of standard selling approach

- Optimal policy hard to compute,
- difficult updates, and
- little information from each customer: slow learning!

Lottery menu:

1. 0 cat for 0$
2. Role a dice
   - ⚀ – ⚁ 1 cat for .25$
   - ⚂ – ⚅ 0 cat for 0$
3. Role a dice
   - ⚀ – ⚃ 1 cat for .375$
   - ⚄ – ⚅ 0 cat for 0$
4. 1 cat for .50$



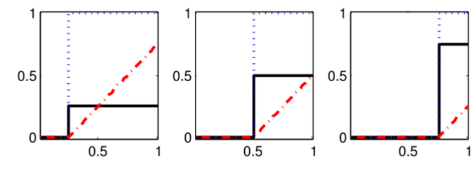
Schrödinger price experiment (SPE)

Superimpose price experiments using lotteries.

M options: get item with prob z_i, pay p_i when you get it.

Example (Standard censored price experiment (CPE))

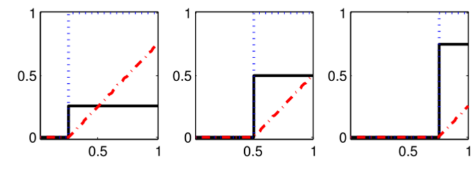
$p_{Mon} = .25, p_{Tue} = .5, p_{Wed} = .75$



	Prob buy	E[payment]
$v \in [0, .25]$	0	0
$v \in [.25, .5]$	$1/3$	$1/3 \times .25$
$v \in [.5, .75]$	$2/3$	$1/3 \times .25 + 1/3 \times .5$
$v \in [.75, 1]$	1	$1/3 \times .25 + 1/3 \times .5 + 1/3 \times .75$

Example (Standard censored price experiment (CPE))

$p_{Mon} = .25$, $p_{Tue} = .5$, $p_{Wed} = .75$

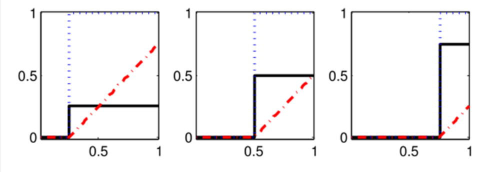


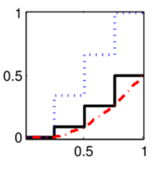
	Prob buy	E[payment]
$v \in [0, .25]$	0	0
$v \in [.25, .5]$	$1/3$	$1/3 \times .25$
$v \in [.5, .75]$	$2/3$	$1/3 \times .25 + 1/3 \times .5$
$v \in [.75, 1]$	1	$1/3 \times .25 + 1/3 \times .5 + 1/3 \times .75$

Using lotteries to extract more information

Example (Standard censored price experiment (CPE))

$p_{Mon} = .25,\ p_{Tue} = .5,\ p_{Wed} = .75$





	Prob buy z_i	E[payment] $z_i p_i$
$v \in [0, .25]$	0	0
$v \in [.25, .5]$	$1/3$	$1/3 \times .25$
$v \in [.5, .75]$	$2/3$	$1/3 \times .25 + 1/3 \times .5$
$v \in [.75, 1]$	1	$1/3 \times .25 + 1/3 \times .5 + 1/3 \times .75$

Schrödinger price experiment (SPE)

Superimpose price experiments using lotteries.

M options: get item with prob z_i, pay p_i when you get it.

Properties of the SPE

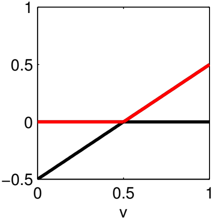
Lemma

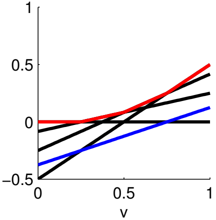
For iid valuations $v \in \mathbb{R}$, expected buyer surplus and revenue from CPE and SPE are the same.

Advantage: using the SPE the seller learns more: observations are intervals instead of half-spaces.

We have an insurance solution for risk averse agents. [See Aistats paper]

Mechanism functions





Fixed price, two options

1. Do not buy: $U(v) = 0$.
2. Buy: $U(v) = v - p$.

Utility maximizers

$$U(v) := \max\{0, v - p\} \ .$$

Menu of m lotteries: m options

$$U(v) := \max_i z_i(v - p_i) \ .$$

Definition (Mechanism function)

Rational agents never pick dominated choices.

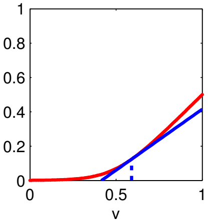
Mechanism function $U(v) := \max_i z_i(v - p_i)$ identifies selling mechanism.

- Derivative $U'(v) = z_i$
- Intercept with x-axis gives p_i

Idea

More options in the menu $\rightarrow$ higher observation resolution.

If $U(v)$ is *strictly* convex, for every v there is a unique best option.



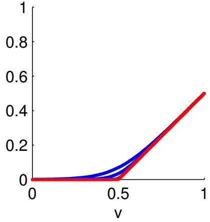
With fully rational buyers seller observes v perfectly
$\rightarrow$ simple updates for conjugate-exponential models.

An optimal selling mechanism

Any "strictly convexified" mechanism learns optimally. Consider

$$U(v) = \log\left(\sum_i e^{z_i(v-p_i)\kappa}\right)/\kappa \;.$$

As $\kappa \to \infty$ we can approximate any finite SPE arbitrarily closely.



In particular the optimal exploiting mechanism $U(v) = \max\{0, v - p^*\}$.

Lemma (Dance, Zoeter 2010)

There exist mechanisms that are arbitrarily close to the optimal mechanism yet for which it is strictly optimal for a buyer to identify his valuation.

Drawbacks of standard selling approach

- Optimal policy hard to compute,
- difficult updates, and
- little information from each customer: slow learning!

Hopeful Mechanism

- Myopic policy is optimal (optimal exploitation),
- v is observed perfectly, conjugate updates possible, and
- optimal exploration.

Let us discuss

1. Returning customers.
2. Robustness against boundedly rational agents.

Come and discuss our solutions for

1. Risk averse agents.
2. Non iid settings.
3. Multi-item setting.

If a buyer returns he can strategically signal lower value.

Lemma (Dance, Zoeter 2010)

It is optimal for buyers to signal true value if and only if price setting policies ensure that observations from buyer b do not affect lotteries offered to b at other times.

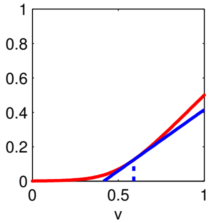
Folds

Our proposed solution: “n-fold”, or “jacknife” learning.
Offer to b is based on belief updated on all but b’s signals.

Assumption: value independent of past purchases.

This implies also that buying many times smallest z_i is not optimal, for it would be better to get your best option all those times.

Robustness



As $\kappa \to \infty$ buyers can lie at very little cost to themselves.

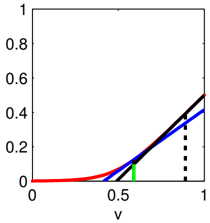
We are working on less extreme versions:

- Offer a finite set of options.
- Make sure that lying hurts.

Non-trivial

- Exploration - exploitation dilemma again.
- Difficult learning problem again.
- Anna Karenina principle.

Robustness



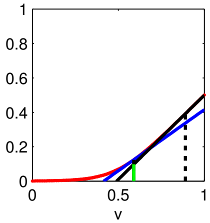
As $\kappa \to \infty$ buyers can lie at very little cost to themselves.

We are working on less extreme versions:

- Offer a finite set of options.
- Make sure that lying hurts.

Non-trivial

- Exploration - exploitation dilemma again.
- Difficult learning problem again.
- Anna Karenina principle.



As $\kappa \to \infty$ buyers can lie at very little cost to themselves.

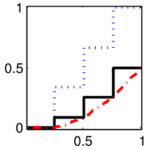
We are working on less extreme versions:

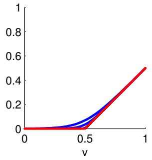
- Offer a finite set of options.
- Make sure that lying hurts.

Non-trivial

- Exploration - exploitation dilemma again.
- Difficult learning problem again.
- Anna Karenina principle.

Summary





Classical learning while selling:

- difficult POMDP problem and difficult belief updates,
- slow (only region information on valuations observed).

Learning-by-lottery

- fast (customer value observed perfectly),
- myopic policy learns optimally and exploits optimally.

Strategic behavior of returning customers can be avoided using folds.

Discussion

1. Learning from strategic agents requires more than machine learning.
2. Strategy proof solutions form a start, but might not be final answer: full rationality might not hold.

Questions?

{Chris.Dance,Onno.Zoeter}@xerox.com

Shameless plug: like good food, good wine, mountains, mechanism design & machine learning? Or know anyone that does?

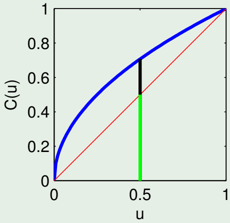
Come and discuss internship opportunities!

- **Problem:** Buyers are risk averse, so they might avoid lotteries.
- Economists model risk with concave increasing functions $C(u)$.
- If two options' utilities have PDFs $f_1(u)$, $f_2(u)$ *e.g.* lotteries
- Then the buyer selects option 1 according to

$$\mathbb{E}_{u \sim f_1} C(u) > \mathbb{E}_{u \sim f_2} C(u).$$

Example

For concave C would buyer prefer .5 or a 50-50 bet on 0 or 1?



Insurance prices

Insurance Prices

- Buyer declares valuations v with component v_i for item $i = 1, \ldots, n$.
- Buyer selects a lottery L with probabilities z and average price p.
- If the lottery outcome is i, they actually pay the insurance price p_i

$$p_i := p + v_i - z \cdot v, \text{ for } i = 1, \ldots, n.$$

Lemma (Dance and Zoeter, 2010)

- *Compare lotteries (z, p) and (z', p').*
- *For any increasing $C(u)$, insurance prices make buyers act risk neutrally*

$$\mathbb{E}_{i \sim z} C(v_i - p_i) > \mathbb{E}_{i \sim z'} C(v_i - p_i') \Leftrightarrow z \cdot v - p > z' \cdot v - p'.$$

Risk-averse buyers compare on the basis of

$$\sum_{i=1}^{n} z_i C \left(v_i - \underbrace{(p + v_i - z \cdot v)}_{=p_i} \right) = \sum_{i=1}^{n} z_i C \left(-p + z \cdot v \right)$$
$$= C \left(-p + z \cdot v \right) \geq C \left(-p' + z' \cdot v \right).$$

Since C is increasing, this is equivalent to comparing

$$-p + z \cdot v \geq -p' + z' \cdot v$$

which is exactly what risk-neutral buyers do. □

Extensions

1. Returning customers.
2. Robustness against boundedly rational agents.
3. Risk averse agents.
4. Generalization of the setting.
   Lotteries change likelihood term:
   works also for non iid $\mathbb{P}(v_{1:T}|\theta)$.
5. Multi-item setting.
   Optimal selling mechanism involves lotteries even if θ is known.
   Our results extend.

1. Returning customers.
2. Robustness against boundedly rational agents.
3. Risk averse agents.
4. Generalization of the setting.
   Lotteries change likelihood term:
   works also for non iid $\mathbb{P}(v_{1:T}|\theta)$.
5. Multi-item setting.
   Optimal selling mechanism involves lotteries even if θ is known.
   Our results extend.

1. Returning customers.
2. Robustness against boundedly rational agents.
3. Risk averse agents.
4. Generalization of the setting.
   Lotteries change likelihood term:
   works also for non iid $\mathbb{P}(v_{1:T}|\theta)$.
5. Multi-item setting.
   Optimal selling mechanism involves lotteries even if θ is known.
   Our results extend.

Questions?

{Chris.Dance, Onno.Zoeter}@xrce.xerox.com

1. **Fixed Price $F(p)$:**
   - Seller sets price p_i for item i.
   - If buyer selects i then she gets item i.
2. **Lottery $L(\{p_j, z_j\}_{j=1}^{n_j})$:**
   - Seller sets prices p_j and probabilities z_{ji}.
   - If buyer selects option j then she gets item i with probability z_{ji}.
3. **Random Price $R(\{p_j, q_j\}_{j=1}^{n_j})$:**
   - With probability q_j
     - Seller sets price p_{ji} for item i.
     - If buyer selects i then she gets item i.

1. **Fixed Price $F(p)$:**
   - Seller sets price p_i for item i.
   - If buyer selects i then she gets item i.
2. **Lottery $L(\{p_j, z_j\}_{j=1}^{n_j})$:**
   - Seller sets prices p_j and probabilities z_{ji}.
   - If buyer selects option j then she gets item i with probability z_{ji}.
3. **Random Price $R(\{p_j, q_j\}_{j=1}^{n_j})$:**
   - With probability q_j
     - Seller sets price p_{ji} for item i.
     - If buyer selects i then she gets item i.

1. **Fixed Price $F(p)$:**
   - Seller sets price p_i for item i.
   - If buyer selects i then she gets item i.
2. **Lottery $L(\{p_j, z_j\}_{j=1}^{n_j})$:**
   - Seller sets prices p_j and probabilities z_{ji}.
   - If buyer selects option j then she gets item i with probability z_{ji}.
3. **Random Price $R(\{p_j, q_j\}_{j=1}^{n_j})$:**
   - With probability q_j
     - Seller sets price p_{ji} for item i.
     - If buyer selects i then she gets item i.

Lemma (Myerson, 1981)

Let $L = \{p^j, z^j\}_{j=1}^{n_j}$ be a lottery for a single item. There is a random price process R and a fixed price process F such that revenue P satisfies

$$\mathbb{E}[P|L] = \mathbb{E}[P|R] \leq \mathbb{E}[P|F].$$

Proof.

1. Let R use prices v_i and probabilities $z_i - z_{i-1}$.
2. Consider a buyer with valuation $v \in [v_i, v_{i+1}]$.

$$\begin{aligned}
\mathbb{E}[P|R, v] &= \sum_{j=1}^{i} (z_j - z_{j-1}) \frac{p_j - p_{j-1}}{z_j - z_{j-1}} = \sum_{j=1}^{i} p_j - p_{j-1} = p_i \\
&= \mathbb{E}[P|L, v]. \\
\text{Yet } \mathbb{E}[P|R] &= \mathbb{E}_{j,v}[v_j\{v > v_j\}] \leq \max_j \mathbb{E}_v[v_j\{v > v_j\}]
\end{aligned}$$

which is the revenue of a fixed price process F with one of the prices v_j. □

Lotteries 101

Lemma (Myerson, 1981)

Let $L = \{p^j, z^j\}_{j=1}^{n_j}$ be a lottery for a single item. There is a random price process R and a fixed price process F such that revenue P satisfies

$$\mathbb{E}[P|L] = \mathbb{E}[P|R] \leq \mathbb{E}[P|F].$$

Proof.

1. Let R use prices v_i and probabilities $z_i - z_{i-1}$.
2. Consider a buyer with valuation $v \in [v_i, v_{i+1}]$.

$$\begin{aligned} \mathbb{E}[P|R, v] &= \sum_{j=1}^{i} (z_j - z_{j-1}) \frac{p_j - p_{j-1}}{z_j - z_{j-1}} = \sum_{j=1}^{i} p_j - p_{j-1} = p_i \\ &= \mathbb{E}[P|L, v]. \\ \text{Yet } \mathbb{E}[P|R] &= \mathbb{E}_{j,v}[v_j\{v > v_j\}] \leq \max_j \mathbb{E}_v[v_j\{v > v_j\}] \end{aligned}$$

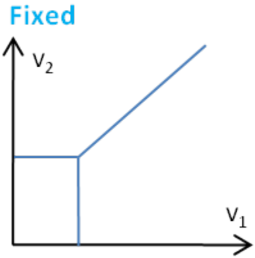
which is the revenue of a fixed price process F with one of the prices v_j.

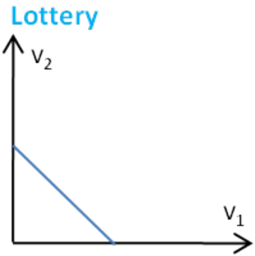
□

- For 2 items and valuation $v \in V$

$$\underbrace{\mathbb{E}[u(v)|F] = \max_i v_i - p_i}_{V \text{ cut into regions shown below}} \quad \underbrace{\mathbb{E}[u(v)|L] = \max_j q^j \cdot v - p^j}_{V \text{ cut into arbitrary convex regions}}$$

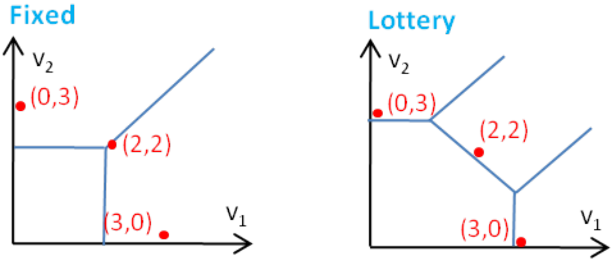
- You cannot make some lotteries from random price experiments
- Single item proof fails.





$u(v) = \max\{0, v_1 - p_1, v_2 - p_2\}$ $u(v) = \max\{0, ½ v_1 + ½ v_2 - p\}$

- Say valuations are uniform on $v \in \{(3,0),(0,3),(2,2)\}$.
- Revenues for the optimal F and L shown below are
  - $\mathbb{E}[P|F] = \frac{1}{3}(2+2+2) = \frac{6}{3}$
  - $\mathbb{E}[P|L] = \frac{1}{3}(3+3+2) = \frac{8}{3}$.



Theorem (Briest, Chawla, Kleinberg, Weinberg 2009)

For $n > 4$ items, lotteries make arbitrarily more revenue than fixed prices.

1. But surely people don't like to play lotteries?
2. Is it practical to compute optimal lotteries?