Learning Valuation Distributions from Strategic Buyers

Chris Dance and Onno Zoeter
Xerox Research Centre Europe
6 Chemin de Maupertuis
38240 Meylan France
{Chris.Dance,Onno.Zoeter}@xrce.xerox.com

February 10, 2011

Abstract

While learning customers' valuation distributions systematically was infeasible until not too long ago, due to the digitization of many shopping environments this has become increasingly possible – at least from a technology point of view. Although it is possible to construct a data stream suitable for machine learning and active price experimentation approaches, a successful use of these is hindered at least by (i) Censored observations: a customer's buying decision only reflects which of the available options was most appealing at purchase time. This constitutes only a very weak signal of the latent vector of customer valuations for products. (ii) Strategic buyers: returning customers might strategically reject certain offers to signal low valuations in an attempt to influence the seller's learning algorithm and obtain cheaper future prices.

The possible benefit for learning customers' valuations is arguably clear in all situations, but a series of recent papers on multi-item sales setting has made an extra case for the need of valuation distributions. For single product selling it is an old result [1] that revenue can't be improved by offering lotteries instead of fixed prices. It has been a long standing assumption that this result extends to multi-item selling. In [2] however it is shown that for two or more items revenue can be increased using lotteries tailored to the valuation distribution. In [3] an elegant demonstration shows that for four or more items the increase in expected revenue can be arbitrarily large.

We show [4] that problem (i) above can be solved completely by using lotteries for learning in the context of fully informed and rational buyers. The basic idea is to offer instead of buy/don't buy options a menu of lotteries where higher probabilities of obtaining the item are combined with higher prices that need to be paid when the items are obtained. By assuming that buyers will maximize their surplus, a seller can deduce regions of valuations for which each menu option is optimal. We give a constructive proof that there exist superpositions (using lotteries) of traditional price experiments that give the exact same expected seller revenue and buyer surplus yet gives increased information to the seller. We also show that there are suitable sequences of lotteries with an infinite number of options that are optimal learning strategies but can get, at the same time, arbitrarily close to an optimal exploiting strategy.

So whereas lotteries can't increase revenue for a single item in the case of perfect knowledge of valuation distributions, they *can* increase revenue if the valuation distribution needs to be learned. Our results extend to the multi-item case so offer an opportunity to learn high-dimensional valuation distributions.

Problem (ii) is studied in some detail in the economics literature and goes back at least to [5]. We present a characterization of the sequential selling mechanism that is resistant to strategic buyer signaling when selling to a group of buyers. The "if" part in the proof says that if customers' signals do not have an impact on their own future prices they have no incentive to strategically reject offers. Using analogies of n-fold cross validation and jack-knife techniques from machine

learning we obtain mechanisms that allow to explore and exploit optimally from a group of strategic agents.

Learning from strategic data sources such as from the buyers in the setting above requires us to consider a richer setting than the standard (active) learning problem since we need to ensure that even if the final system is optimally exploited good things still happen. For the multi-item selling setting the above described approach presents a solution that perfectly and optimally solves both the weak signal and strategic behavior problems. However it relies heavily on the assumption that buyers are fully informed and fully rational.

If there are elements of friction in the final application (imperfect information, irrational buyers, etc.) an incorrect assumption that buyers are perfectly rational can lead the system to draw the wrong conclusions from observations with possibly disastrous results. It will also leave opportunities on the table against myopic buyers since the system is then being overly conservative in an effort to be resistant to exploits from a perfectly rational buyer.

Arguably all results should start with the assumption of full rationality and refine from there. Relaxing this assumption is difficult. This is due to the 'Anna Karenina principle' which derives from quote from Tolstoy's book: "Happy families are all alike; every unhappy family is unhappy in its own way." Rationality is easy to describe and unique, however irrationality can manifest itself in many different ways, each requiring a different optimal response from the seller.

We conclude with a discussion of possible models of irrationality, what they imply for the selling mechanisms outlined above, and sketch directions for future research.

References

[1] J. Riley and R. Zeckhauser. Optimal selling strategies: When to haggle, when to hold firm. *Quart. J. Econ.*, pages 267–289, 1983.

[2] J. Thanassoulis. Haggling over substitutes. *J. Economic Theory*, 117(2):217–245, August 2004.

[3] P. Briest, S. Chawla, R. Kleinberg, and S. Weinberg. Pricing randomized allocations. In *21st SODA*, 2010.

[4] C. Dance and O. Zoeter. Optimal and robust price experimentation: Learning by lottery. Under review, 2010.

[5] N. Stokey. Intertemporal price discrimination. quart. *Quart. J.Econom.*, 93(3):355–371, 1979.

Topic multi-agents, online learning, mechanism design, discrete choice.

Preference oral

Presenter Onno Zoeter