Causality recap

In machine learning, you’re typically a passive observer. As such, correlations are sufficient for most applications. However, once you start using machine learning to make decisions, you now have transitioned to caring about causal relations rather than correlations.

For example, you observe a room full of banana peels. A person walks in the dark and you observe whether they trip or not. You can take whether they tripped to learn about the number of banana peels, and you can predict whether or not they tripped based on the number of banana peels. You can predict either way because of the correlation. You can’t trip someone to make more banana peels.

This example may seem silly, but at its core, machine learning and data analytics is about discovering previously unknown relationships. Thus, it’s hard to which direction is the ‘silly’ direction when we try to determine causal directions. The distinction between correlation and causation is actually quite difficult to make in real life; keep in mind that we often use machine learning when causal relationships are difficult to determine. Last lecture, we covered paradigms for modeling causality: Rubin, Granger, and Pearl.

1 Auctions

We follow Easley and Kleinberg [2010].

In this section, we study the case where one seller is trying to sell one item to a set of buyers. A dual problem is where one buyer wants to purchase one item from multiple potential sellers. (These are known as procurement auctions.)

It’s good to know some of the terminology associated with auctions.

- First-price sealed-bid auctions.
- Second-price sealed-bid auctions, Vikrey auctions.
- Ascending-bid auctions, English auctions: the seller gradually raises the price until only one buyer remains.
- Descending-bid auctions, Dutch auctions: the seller gradually lowers the price until the first moment a buyer opts in.

Each buyer has a valuation (intrinsic value, true value) for the item for sale. Similarly the seller.

1.1 Known values

Suppose all the valuations are common knowledge. Then let \( x \) be the seller’s valuation, and let \( y \) be the maximum valuation among the buyers. If the price paid is \( p \), the seller gets \( p - x \) and the buyer gets \( y - p \), so the social welfare is \( y - x \). This is the surplus available from trade.

If the seller commits to selling at \( y - \epsilon \) and refuses to accept anything lower, then the buyer will buy it and all the surplus goes to the seller. If the buyer commits to buying at \( x + \epsilon \) and refuses to pay any more, then the seller will sell it and all the surplus goes to the buyer.

The surplus goes to the player who can most believably ‘commit’.

1.2 Independent, private values

Descending-bid is equivalent to first-price. (You learn nothing until the auction is over, other than no one has yet accepted the current price.)

Ascending-bid is equivalent to second-price. (You stay until your personal valuation. If you win, you pay the price of the second-to-last person who dropped out.)
1.2.1 Second price

If \( b_i \) is not the winning bid, player \( i \) gets payoff 0. If it is the winning bid, and \( b_j \) is the second-price bid, then the payoff is \( v_i - b_j \).

**Proposition 1.1** (Truthful bidding). In a sealed second-price auction, it is a dominant strategy for each bidder \( i \) to choose \( b_i = v_i \).

1.2.2 First price

If \( b_i \) is not the winning bid, player \( i \) gets payoff 0. If it is the winning bid, then the payoff is \( v_i - b_i \).

You ‘shade’ your bid slightly downward. There’s no equilibrium without further structure: what’s your belief of the bids of others.

Suppose there are 2 bidders, each with a private value \( U(0, 1) \). (The distribution is common knowledge; the values are not.)

A strategy is a mapping \( s(v) = b \) from valuations to non-negative bids. We assume that \( s(\cdot) \) is strictly increasing and differentiable, and \( s(v) \leq v \) for all \( v \). Note that this implies \( s(0) = 0 \). By symmetry, it is reasonable to assume that they use the same \( s \).

By assumption, the bidder with the higher valuation will win. If the bidder’s value is \( v_i \), then the probability he has the higher valuation is \( v_i \). Thus, his expected payoff is:

\[
g(v_i) = v_i(v_i - s(v_i))
\]

If \( s \) is an equilibrium strategy, then there is no incentive to deviate if the other bidder is also using \( s \).

Revelation principle: Suppose \( i \)’s competitor is using \( s \). Then \( i \) should never bid above \( s(1) \). Thus, every bid should be between \( s(0) = 0 \) and \( s(1) \). So, if \( i \) wants to deviate to a different strategy \( s' \), then would want to bid \( s'(v_i) \). However, if they still use \( s \) but report a different valuation, they could bid \( s'(v_i) \) by simply reporting valuation \( s^{-1}(s'(v_i)) \). Every deviation can be simulated with a change in valuation.

The revelation principle here says we can view deviations in strategy as deviations in valuations for the same strategy. In more general mechanism design problems, the revelation principle allows us to restrict ourselves to mechanisms in which users truthfully report their value. (If they had incentive to deviate, we can build that deviation into the mechanism.)

Thus, the no-deviation condition can be written for all ‘fake’ valuations \( v \):

\[
v_i(v_i - s(v_i)) \geq v(v_i - s(v))
\]

We want the maximizer of \( v(v_i - s(v)) \) to be \( v_i \). The first derivative is:

\[
v_i - s(v) - vs'(v)
\]

Thus, optimality states this is 0 at \( v = v_i \):

\[
s'(v_i) = 1 - \frac{s(v_i)}{v_i}
\]

We also have \( s(0) = 0 \). The solution to this differential equation is \( s(v_i) = v_i/2 \).

More generally, with \( n \) bidders in this setting, the optimal \( s(v_i) = \frac{n-1}{n}v_i \).

More generally, if the distribution has CDF \( F \) and PDF \( f \), expected payoff is:

\[
F(v)^{n-1}(v_i - s(v))
\]

Then:

\[
s'(v_i) = (n-1)\frac{f(v_i)v_i - f(v_i)s(v_i)}{F(v_i)}
\]

Revenue equivalence: in this particular setting with uniform distributions, seller gets the same revenue in both auctions.
1.2.3 All-pay

If \( b_i \) is not the winning bid, player \( i \) gets payoff \(-b_i\). If it is the winning bid, then the payoff is \( v_i - b_i \).

Political lobbying. Still has ‘shading’.

The formula for expected payoff is:

\[
v^{n-1}(v_i - s(v)) + (1 - v^{n-1})(-s(v))
\]

The condition is now:

\[
v^n_i - s(v_i) \geq v^{n-1}v_i - s(v)
\]

Thus: \( s'(v_i) = (n-1)v^{n-1}_i \) and \( s(v) = \frac{n-1}{n}v^n_i \). This is now exponential in the user’s valuation. This also has the same expected revenue as before. More revenue equivalence.

1.2.4 Reserve prices

Oftentimes, you can set a reserve price \( r \). If the highest bid is above \( r \), then the auction goes through; otherwise nothing happens. In the first price auction, it proceeds normally. In the second price auction, the winner pays either the reserve price \( r \) or the second highest bid; whichever is higher. The reserve price \( r \) is treated like a simulated bid.

The type of analysis we did can be repeated to find optimal reserve prices (which are rarely the seller’s valuation of the item).

1.3 Common values

Suppose everyone has a common valuation \( v \), but only observes some noisy estimate \( v_i = v + \epsilon_i \) where \( E\epsilon_i = 0 \).

Suppose then every bidder bid \( b_i = v_i \). The winner has high probability of losing value then, since, conditioned on them winning, it is very likely \( v_i > v \). This is the winner’s curse.

References