Model Development

The Advertiser's Objective

Let B represent the total budget for an advertising campaign. The advertiser spends B to purchase individual impressions in real-time through a second price auction. Let x_{ij} be a vector of attributes describing individual i in auction j. Then for auction j, the advertiser bids amount b_j, wins with probability $w(x_i, b_j)$, and pays amount $b_j' \leq b_j$. For each auction, the advertisers expected expenditure, $\mathbb{E}[\, e_j \mid x_{ij}, b_j \,]$, can be calculated as $e_j = w(x_{ij}, b_j) b_j'$.

Let m_{ij} represent an impression served to individual i as a result of the advertiser winning auction j. Let c_{ij} be the change in probability that individual i converts resulting from impression m_{ij} and a vector of individual attributes x_{ij}, such that

$$c_{ij} = \Delta \Pr(\, Conv_i \mid x_{ij}, m_{ij} \,). \tag{1}$$

At the time of each auction, we can express individual i's expected conversion probability as

$$\Pr(\, Conv_i \mid b_j, x_i \,) = c_{0i} + w(x_{ij}, b_j) c_{ij}, \tag{2}$$

where c_{0i} is the baseline probability of conversion for individual i, $w(x_{ij}, b_j)$ is the probability that the advertiser wins auction j, and c_{ij} is the impact of impression j on individual i's conversion probability.

Given the advertiser's goal to maximize conversions, we can express their objective function as:

$$\begin{aligned} &\max_{b_j} \sum_i \Pr(\, Conv_i \mid b_j, x_i \,) \\ &s.t. \sum_j b_j' \leq B \text{ and } b_j \geq 0 \; \forall \, j \end{aligned} \tag{3}$$

Because c_{0i} is not impacted by the firm's bidding decision, b_j, equation (3) reduces to:

$$\begin{aligned} &\max_{b_j} \sum_i w(x_{ij}, b_j) c_{ij} \\ &s.t. \sum_j b_j' \leq B \text{ and } b_j \geq 0 \; \forall \, j \end{aligned} \tag{4}$$

Parallels to the 0-1 Knapsack Problem

Assume that the advertiser knows the population defining when and to whom impressions can be served, how those impressions will impact conversion probabilities, and the price of each impression. That is:

Assumption 1: x_{ij} is known *a-priori* for all i and j.
Assumption 2: The form of c_{ij} is known *a-priori.*
Assumption 3: b_j' is known *a-priori*.

Under such a scenario, the advertiser can bid any amount $b_j > b_j'$, and win the auction with probability $w(x_{ij}, b_j) = 1$.

The firm's objective function in equation (4) is then a form of the standard 0-1 knapsack problem, where each auction, j, is an item, b_j' is the weight of the item, c_{ij} is the item's value, and B is the capacity of the knapsack. The firm must make a binary decision with regard to each auction by either bidding $b_j > b_j'$ to serve the impression and augment total expected conversions by c_{ij}, or bidding $b_j < b_j'$ to pass on the opportunity and leave expected conversions unchanged. The optimal course of action can be solved using the standard approaches.

Parallels to the Stochastic 0-1 Knapsack Problem

Relax Assumption 3, allowing the advertisers to instead know the probability of winning based on the bid amount b_j and the vector of attributes x_{ij}, such that:

Assumption 3': $w(x_{ij}, b_j)$ is known *a-priori*.

Further, assume that the amount paid equals the bid amount, such that:

Assumption 4: $b_j = b_j'$.

While this assumption violates the structure of second price auctions, the empirical distribution of the differences between amounts bid and paid is very near zero, as seen in Figure ____. Given this assumption, the firm's objective function in Equation (4) reduces to

$$\begin{aligned} &\max_{b_j} \sum_i w(x_{ij}, b_j) c_{ij} \\ &s.t. \sum_j b_j \leq B \text{ and } b_j \geq 0 \; \forall j \end{aligned} \tag{5}$$

and the firm's decision problem takes the form of the Stochastic Knapsack Problem.

Parallels to The Stochastic On-Line 0-1 Knapsack

Relax Assumption 1, such that the advertiser knows the distribution of individuals to whom an ad could be served and the probability with which each will be available. Each x_{ij} is revealed sequentially, and the advertiser must make a bidding decision before revealing the next x_{ij}.

Assumption 1': The joint distribution of x_{ij} is known.

Srinivas

Each bidding opportunity is characterized by a vector of features x_i (cookie id, location, webpage etc.). Prior probability of observing a vector x is assumed to be known.
Also, $\theta(x)$ - the conversion probability, $w(x, b)$- the probability of winning the auction for bid b, the number of bidding opportunities N are assumed to be known.

Decision variables: Amount to bid on each opportunity, b_i.
Objective function: Expected number of conversions on auctions won.
Constraints: Budget for bidding, B.

Assumption: The amount paid for an auction won equals the bid amount.

If the bid amounts are already decided and the only decision variables are whether to bid for an auction or not, then the problem can be modelled as the *online knapsack problem* where:

- Each bidding opportunity is an item
- The conversion probability is the value of the item
- Bid amount is the weight of the item

But there are two important distinctions:

- When an item is picked to be put in the knapsack, it can only be done with a certain probability (the probability of winning the auction for that bid amount)
- The random outcome (probability of winning the auction) is dependent on the decision variable (the amount to bid)

The dependency between the random outcome and the decision variable can be removed as follows:
Consider only a finite number of bid amounts possible for item i: b_{i1},b_{i2}, ..., b_{iK}.
An item i with value v_i can be modelled as a batch of items i_1, ..., i_K, all with value v_i but with different weight b_{i1}, ..., b_{iK}. Each of these items, if decided to be picked, can be put in the knapsack with probability p_{i1}, ..., p_{iK}.
An added constraint is that only one item in the batch can be picked.

Using this model, the problem can be formulated as a stochastic program with N stages of recourse.

Let x_{i1}, ..., x_{iK} be binary variables denoting if the corresponding item of batch i is picked. Note that $\sum_{k=1}^{K} x_{ik} = 1$. Let $OPT(N, B)$ denote the optimal objective value where N is the number of bidding opportunities and B is the budget for bidding. Then,

$$OPT(N, B) = \sum_{k=1}^{K} x_{ik}(p_{ik}(OPT(N-1, B) + \theta_i) + (1 - p_{ik})OPT(N-1, B))$$

Using the recursion above, the problem can be solved using dynamic programming.