Introduction

The primary mode of Internet advertising is via *search advertisements* and *display advertisements*. The basic metric for advertising effectiveness is the *conversion rate* that measures the percentage of last clicks that resulted in a purchase. The authors say that the entire path to purchase needs to be modeled as a *stochastic process* rather than the merely the last step. The authors illustrate this via a simple example where a customer clicks on a banner ad at t_1, then uses a search engine at t_2 and then goes ahead with the purchase. The basic problem with using the last click to model the conversion rate is that it overweights search ads and neglects the banner ad that drove the customer in the first place to a search engine. This correlated behavior between various time events are called *exciting effects*. In another example given, the standard way to take account of multiple clicks on a search engine that lead to a purchase is to attribute equal weights to each fo the clicks. In reality, the first click would have provided all the stimulation needed to click two other links before making a purchase.

Any model that tries to gets its hand around the path to purchase, should ideally look at all the time instants before the purchase. The data available should neither be treated as deterministic data nor should there be some kind of aggregation at a fixed point. Besides, the model needs to capture

- Different types of online advertisements have their distinct natures - Hence one needs to have a multivariate model
- Each customer is different and hence there is a need to incorporate individual heterogeneity in the model
- Model needs to have decaying effects of different types of advertisement clicks
- Clicks and purchasing data is *clumpy*. Poisson process doesn't work. History dependent intensity functions need to be used in the modeling exercise.

10,000 ft view

The model developed in the paper uses a hierarchical model to capture the heterogeneity and uses a multivariate point process to capture the purchase path. The model fits a conversion probability to the purchase path.The model also predicts individual consumers online behavior based on past behavioral data. The authors claim that

- They are the first to apply mutually exciting point process model in a marketing context
- They are the first to incorporate individual random effects into the mutually exciting point process
- They are the first to apply Bayesian inference using MCMC to mutually exciting point process.

Simulation of the point process is done by extending the thinning algorithms in Ogata.

Literature Review

- The effect of banner advertising on Internet purchasing - Uses a survival model. It is a univariate model
- Dynamic conversion rates at e-commerce sites - Model accumulative effects of website visits. It is a univariate model
- Modeling online browsing and path analysis using click-stream data - Markov model for webpage views
- Modeling browsing behavior at multiple websites - Use a bivariate distribution to model dependence on website visit durations across two websites
- Modeling multivariate distributions using copulas: Applications in marketing

- Original Hawkes process papers
- Ait Sahalia's paper of using exciting process to model contagion - Individual heterogeneity effects are not considered
- Modeling purchases as repeated events - Poisson process with intensity function as a function of previous path of the point process

Dataset and preliminary findings

- Historical purchase path available for a set of customers - 12,000 user Ids from April 1 to July 31, 2008
- The ad clicks are divided in to three categories - *search*, *display* and *other*
- A certain ad click leads to a conversion if it is succeeded by a purchase of the same individual within one day
- Search advertising leads display advertising in the dataset
- A random selection of 10,000 user Ids reveals that
  - Sequences involving repeating clicks on the same type of ads are very common
  - Significant portion of samples contain the behavior where one type of click is succeeded by another type of click
  - Display ads appear more likely to excite the other two types of ads than the other way around

Model Development

Basic Hawke's process form

$$\lambda_k(t|\mathcal{H}_t) = \mu_k + \sum_{i=1}^{K} \int_{-\infty}^{t} g_{jk}(t-u)dN_j(u), \quad \mu_k > 0$$

The common form of the response function is

$$g_{jk} = \alpha_{jk} e^{-\beta_{jk}\tau}, \quad \alpha_{jk} > 0, \beta_{jk} > 0$$

Proposed Model :

$$\lambda_k^i(t|\mathcal{H}_t^i) = \mu_k^i \exp(\psi_k N_k^i(t)) + \sum_{i=1}^{K} \sum_{l=1}^{N_j^i(t)} \alpha_{jk} \exp\left(-\beta_j \left(t - t_l^{j(i)}\right)\right) \quad , \mu_k > 0$$

$$\begin{aligned}
N^i(t)|\alpha, \beta, \psi, \mu^i &\sim \lambda^i(t|\mathcal{H}_t^i) \\
\mu^i|\theta_\mu, \Sigma_\mu &\sim \mathrm{MVN}_K(\theta_\mu, \Sigma_\mu) \\
\alpha_{jk} &\sim \Gamma(\overline{a}_\alpha, \overline{b}_\alpha) \\
\beta j &\sim \Gamma(\overline{a}_\beta, \overline{b}_\beta) \\
\psi &\sim \mathrm{MVN}_K(\theta_\psi, \Sigma_\psi) \\
\theta_\mu &\sim \mathrm{MVN}_K(\overline{\theta}_{\theta_\mu}, \overline{\Sigma}_{\theta_\mu}) \\
\Sigma_\mu &\sim \mathrm{IW}(\overline{S}^{-1}, \overline{\nu})
\end{aligned}$$

Benchmark Model :

$$\lambda_k^i(t|\mathcal{H}_t^i) = \mu_k^i$$

Use a logistic conversion model on the out-of-sample data as the benchmark model for comparison purpose

Assumptions :

- Intensity for each marginal process at any time depends on the entire history of all *marginal processes*
- baseline intensity for individual across type of ad clicks is modeled with a MVN prior
- α_{jj} capture *self-exciting effect*
- $\alpha_{jk}, j \neq k$ capture *mutually -exciting effect*
- $\beta_{jk} = \beta_j$: decay rates of exciting-effects of adclick type j on other adclick types is the same
- Each purchase can effect the subsequent purchase or effect on ads in a positive or negative way. Hence the use of multiplicative term $\exp(\psi_k N_k^i(t))$
- The intensity for the exciting process for a specific customer is a vector random process and encapsulates all the nonsystematic unobservables and idiosyncratic shocks
- Accumulative effects from the infinite past up to time $t = 0$ is zero

Estimation

MCMC algo is run to generate 50,000 iterations. These iterations are used to obtain parameter estimates. The findings from analyzing estimation results are

- Significant exciting effects exist between same types of ad clicks as well as between different types of ad clicks
- Compared to mutually exciting effects, self exciting effects are more salient
- Display ads tend to have greater exciting effects
- Clicking on ad and visiting firm's website increases the probability of purchase
- Past purchases dampens the prospect of future purchase
- Consumers more likely to click on banner ads have low conversion rates

Model comparison shows that Mutually exciting model is better than Self-exciting process, which in turn is better than Poisson process.

Model Applications

As a marketer, the one important metric that is important from this model is the *conversion probability*

$$CP_k(t; \mu^i, \alpha, \beta, \psi) = Pr\left(N_K^i(t_0 + t) - N_K^i(t_0) > 0 | N_K^i(t_0) - N_k^i(t_0-) = 1\right)$$

However one is not interested in a particular customer. Hence what is needed is *average conversion probability*, ACP

$$ACP_k(t) = E[CP_k(t; \mu^i, \alpha, \beta, \psi)|Data]$$

Since the above cannot be computed via closed form, a montecarlo simulation procedure is to computer average conversion probabilities for various adclick categories. The basic idea behind the simulation procedure is the accept-reject method. Given the parameters obtained from the mutually exciting process, 1 million purchase paths are simulated for a given starting ad click category.

The simulation results show that :

- Display advertisement effectiveness has been underestimated in the previous studies
- Comparing mutually exciting model to self exciting model shows that display, search and other ad effectiveness is under estimated
- Poisson model underestimates the adclick effectiveness more than self exciting or mutually exciting model

The author perform out-of-sample validation and conclude that mutual exciting model is better than self exciting model, which in turn is better than poisson process. 95% CI intervals capture the observed purchases per customer. DIC also shows that mutually exciting model is better than the other two benchmark models.

Discussion and Conclusion

- The authors test a Poisson count model and find it inferior in performance.
- Mutually exciting models better than restricted copula models
- One can generalize the model based on additional data availability - direct visits, ad exposure data
- Display ads are likely to stimulate visits even though they have low direct effect on purchase conversions
- The model enables one to predict future responses to different online ad formats