STAT 30850 Final Report

Arjun Biddanda (abiddanda@uchicago.edu)
Joseph Marcus (jhmarcus@uchicago.edu)

March 15, 2016

Introduction

There are many contexts and applications where data is observed sequentially through time. For instance in high freqeuncy stock trading, investment firms have to make rapid descions in repsonse to new stock evaluations on micro-second timescales, or in A/B testing, techology companies often test the effect of varied advertisements on the "click behavior" of a user which is correlated with the effectiveness of the advertisement [CITE, CITE]. The setting in which hypotehsis testing must be performed on sequential streaming data is called "online testing". In online testing controlling the False Discovery Rate (FDR) at a given level has unique challenges as one doesn't observe all the data that could potentially be seen, later in the time series. Here we propose to use and implement a Bayesian model-based appraoch to control FDR in the online testing setting. We review and contrast our approach to previous commonly used heuristics / algorithims that are effectetive but conservative in online hyptohesis testing. We show our approach has higher power when compared to previous methods. Finally, we discuss future extensions and applications of our method.

Background

Broadly speaking, previous methods for controlling FDR in the online testing context use heuristics that increase or decrease the level at which one rejects a test depending on the number of previous discoveries made. Here we review three related, commonly used and well studied approaches to FDR control: α - investing, Levels Based on Number of Discoveries (LBOND), Levels Based on Recent Discoveries (LBORD) [CITE, CITE, CITE].

α-investing

Let:

t - be time
$w(t)$ - be a wealth function which changes through time
P_t - be a p-value output from an arbitrary test a time t
α - a global level that one would like to control FDR at
α_t - a time specific level

In alpha-investing ones defines a wealth function w. We imagine p-values are streaming in over time t which are provided by some abitrary test. We then proceed to run the α-investing procedure:

1. $w(t = 0) = \alpha$

2. At time t choose $\alpha_t \leq \frac{w(t-1)}{1+w(t-1)}$

3. Rejct the null hypothesis if $P_t \leq \alpha_t$
4. Define $w(t)$ as a function of $w(t-1)$

$$w(t) = \begin{cases} w(t-1) + \alpha & P_t \leq \alpha_t \\ w(t-1) - \frac{\alpha_t}{1-\alpha_t} & P_t > \alpha_t \end{cases}$$

5. Repeat the procedure starting back at (2) for each new data point in the time series.

As we can see above when we reject the null, the wealth function grows and when we fail to reject the null the wealth function decays. Specifically at time 0 we set the wealth function to a "global level" alpha. We then proceed to set a time specific α_t. We then reject or fail to reject the p-value P_t from time t and redefine our wealth function w depeding on what desicion was made. This ensures that the more discoveries we make the less stringent we are through time and recipricaly the fewer discoveries we make the more stingent we are through time. For instance if we fail to reject for many sequential time points the signals have to be very strong to overcome the current state of the wealth function.

LBOND / LBORD

Let:

t - be time
P_t - be a p-value output from an arbitrary test a time t
α - a global level that one would like to control FDR at
β_t - a time specific weight
D_t - count of discoveries made up to time t

In Levels Based on Number of Discoveries (LBOND) we define a series of weights β_t which all sum up to the global level α. We then set a time specific α_t equal to the the weight at time t times the max of 1 and the number of discoveries made up to the last time step D_{t-1}. We reject a p-value P_t if its less that α_t and add to our discovery count.

1. At time t set $\alpha_t = \beta_t \cdot max\{1, D_{t-1}\}$ where $\sum_{t=1}^{\infty} \beta_t = \alpha$
2. Reject if $P_t \leq \alpha_t$
3. If discovery add to D
4. Repeat

Levels Based on Recent Discoveries follows a similar appraoch but uses weights from the time when the last discovery was made.

TODO: Fill Description of LBORD

Methods

Here we propose to apply a Bayesian approach to FDR control in the online testing setting. Specifically we follow the work of Efron and model our streaming data as test statistics coming from a mixture model

Discussion

Conclusion

References