

Multi Armed Bandits

Zainab Danish, Andre Duarte, Spencer Smith, Derek Welborn

June 22, 2017

1 Introduction

Our application is rather simple: the page has a button and each time the button is clicked a random number between 0 and 1000 is generated. The factor is the caption on each page that we vary. The factor levels are:

- "Whats the pattern?" - V1
- "Can you figure it out?" - V2
- "It's just random!" - V3

The response variable is the amount of time spent on the site.

Our hypothesis is that a visitor may be more inclined to stay on the page longer clicking the button in order to try to understand whether there is a pattern in the number generation. Therefore, by knowing through the caption that it is random, the user may leave the page earlier.

2 Multi Armed Bandits

2.1 How it works

Multi-armed bandit experiments are an alternative approach to AB-testing. Whereas in AB-testing for a website you would send equal traffic to each variant, in a multi-arm bandit setting you would set aside a small percentage of time in which you would send equal traffic to each variant (the exploration phase), after which you would start redirecting majority of the traffic to the highest-performing variant. In a sense, the bandit learns to eliminate the inferior choices early on.

2.2 Strengths

Relative to standard hypothesis testing, bandits have the benefit of optimizing conversion rate. If for example, there is a major difference between say two variants of a webpage, under standard hypothesis testing conditions 50% of the users would be directed to the 'bad' variant which would result in a loss of users (and potentially money). Under the multi-armed bandit conditions, the traffic is 'greedily' redirected toward the best performing variant. Hence, an advantage of multi-armed bandits is high conversion rates – given that conversion is the basic objective – and minimization of potential loss. Multi-armed bandit experiments also tend to be faster at getting to an optimal condition, since it constantly updates the experiment.

2.3 Weaknesses

The greedy approach of the bandit, however, has its drawbacks. While under standard testing conditions it would not take too many trials to establish statistical significance because equal users are assigned to each condition, with multi-armed bandits a much larger number of trials is needed to establish statistical significance because a progressively smaller number of users get assigned to the low performing variant. In addition, since the process is greedy, a sub-optimal variant may be 'selected' as the winner instead of a better one due to randomness.

2.4 Sample Size under A/B testing

The parameters we chose for this experiment are listed as follows:

- Significance Level: 5% (set in Google Analytics)
- Power of test: 90% (we want to minimize type II error, so that we don't 'miss' when the variants are actually different)
- Effect Size: 0.1 standard deviation (we expect our variants to perform similarly, so the effect has to be small)
- Two-sided test

Based on the parameters listed above the sample size came out to be 2103 for each variant, 8412 in total. This number is very large, especially for an experiment that we are running within our class, and we won't have much traffic. The multi-armed bandit should be able to draw its conclusions with much less traffic than this.

3 Experiment Results and Conclusion

After 5 days of data collection, the experiment finished and found a winner: variant 3. We had 57 total user sessions, which is very little but apparently enough to draw the final conclusions. Google Analytics "chose" variant 3 as the best one. Indeed, 37 sessions were directed to this variant, while only 10 and 10 were directed to variants 1 and 2 respectively. Variant 3 has a 100% probability of outperforming variant 1, while variant 2 has 49.9% probability.

This result is not exactly surprising (we mainly expected variant 2 to perform poorly). However, it is very likely that we introduced significant bias to the experiment since many of the visits came from ourselves. This may have influenced the final result. In addition, most users in our sample knew that this was an experiment, and might have introduced demands characteristics (causing them to behave in a certain way).