EDITORIAL!

klopotek.

Future systems needs for publishers to manage marketing or products and content in a digital world

Bayes' Theorem

$$P(H \mid E, c) = \frac{P(H \mid c) \; P(E \mid H, c)}{P(E \mid c)}$$

Despite the esoteric symbols, the idea – roughly speaking – is simple: The likelihood that something will happen can be plausibly estimated by how often it occurred in the past. Researchers are applying the idea to everything from gene studies to filtering e-mail. Source: http://news.cnet.com/2009-1001-984695.html

This article builds on our earlier article surrounding the use of Bayes theory in predicting the true result of product sales also available from our web site and some of the same theories are recounted here.

In today's world of publishing the proliferation of product types and new revenue models created by the availability new digital products and media channels has led to real challenges in testing and marketing new products and analyzing the effectiveness of adopting these new products and sales models.

These challenges have been created by a number of factors, such as the proliferation of the types of product and content available, the complex variation of ways to provide access to such content and products (for example, iPad and Android Apps) and the speed at which the market is being created for these new models and content types.

Traditionally, publishers have relied upon experimental A/B testing. It's a tried, trusted and tested way of learning what works and what does not. This theory is based on the principle of trying something, analyzing what worked and what did not, then trying something else and then comparing the results to see what is most effective.

While this works, it takes time and considerable resource. In today's world time is not a luxury one can always afford and the volume of data and factors affecting the models is such that it can become prohibitively expensive to attempt to analyze and compare using these traditional models.

This article describes the new statistical theories that provide a better and more manageable way forward, both in terms of accuracy of the results and in terms of speed and affordability.

Many companies in other industries are now adopting a "multi-armed bandit type" of experiment where:

- The goal is to find the best or most profitable action
- The randomization distribution can be updated as the experiment progresses in real time

The name "multi-armed bandit" describes a hypothetical experiment where you face several slot machines ("one-armed bandits") with potentially different expected payouts. You want to find the slot machine with the best payout rate, but you also want to maximize your winnings. The fundamental tension is between "exploiting" arms that have performed well in the past and "exploring" new or seemingly inferior arms in case they might perform even better.

Background

How bandits work
It does not matter if your models are physical or digital; this approach works for both physical and digital product, however, digital is much more precise and faster moving and therefore ecommerce models really benefit from the approach.

Several times per day, the software takes a fresh look at your "experimental" sales and new product models to see how each of the variations have performed, and we adjust the fraction of traffic, using algorithms, that each variation will receive going forward. A variation that appears to be doing well gets more traffic, and a variation that is clearly underperforming gets less. The adjustments the system then makes are based on statistical formulae that consider sample size and performance metrics together, so that you can be confident that we're adjusting for real performance differences and not just random chance. As the experiments progress, we learn more and more about the relative payoffs, and so do a better job in choosing good variations.

Benefits
Experiments based on multi-armed bandits are typically much more efficient than "classical" A-B experiments based on statistical-hypothesis testing. They're just as statistically valid, and in

many circumstances they can produce answers far more quickly. They're more efficient because they move traffic towards winning variations gradually, instead of forcing you to wait for a "final answer" at the end of any experiment. They're faster because samples that would have gone to obviously inferior variations can be assigned to potential winners. The extra data collected on the high-performing variations can help separate the "good" arms from the "best" ones more quickly.

Basically, bandits make experiments more efficient, so you can try more of them. You can also allocate a larger fraction of your traffic to your experiments, because traffic will be automatically steered to better performing pages.

Examples

A simple A/B test
Suppose you've got a conversion rate of 4% on your marketing progam. You experiment with a new version of the products or program that actually generates conversions 5% of the time. You don't know the true conversion rates of course, which is why you're experimenting, but let's suppose you'd like your experiment to be able to detect a 5% conversion rate as statistically significant with 95% probability. A standard power calculation1 tells you that you need 22,330 observations (11,165 in each arm) to have a 95% chance of detecting a .04 to .05 shift in conversion rates. Suppose you get 100 sales per day from the experiment, so the experiment will take 223 days to complete. In a standard experiment you wait 223 days, run the hypothesis test, and get your answer.

Now let's manage the 100 sales each day through the multi-armed bandit. On the first day about 50 sales are assigned to each arm, and we look at the results. We use Bayes' theorem to compute the probability that the variation is better than the original 2. One minus this number is the probability that the original is better. Let's suppose the original got really lucky on the first day, and it appears to have a 70% chance of being superior. Then we assign it 70% of the traffic on the second day, and the variation gets 30%. At the end of the second day we accumulate all the traffic we've seen so far (over both days), and have recomputed the probability that each arm is best. That gives us the serving weights for day 3. We repeat this process until a set of stopping rules has been satisfied (we'll say more about stopping rules below). The experiment finished in 66 days, so it saved you 157 days of testing and provided a much more accurate result.

We can also run the opposite experiment, where the original had a 5% success rate and the variation had 4%. The results were essentially symmetric. Again the bandit found the correct arm 482 times out of 500. The average time saved relative to the classical experiment was 171.8 days, and the average number of conversions saved was 98.7.

Stopping the experiment
By default, the system forces the bandit to run for at least two weeks. After that, it keeps track of two metrics.

The first is the probability that each variation beats the original. If we're 95% sure that a variation beats the original then the software declares that a winner has been found. Both the two-week minimum duration and the 95% confidence level can be adjusted by the user.

The second metric that we monitor is the "potential value remaining in the experiment", which is particularly useful when there are multiple arms. At any point in the experiment there is a "champion" arm believed to be the best. If the experiment ended "now", the champion is the arm you would choose. The "value remaining" in an experiment is the amount of increased conversion rate you could get by switching away from the champion. The whole point of experimenting is to search for this value. If you're 100% sure that the champion is the best arm, then there is no value remaining in the experiment, and thus no point in experimenting. But if you're only 70% sure that an arm is optimal, then there is a 30% chance that another arm is better, and we can use Bayes' rule to work out the distribution of how much better it is.

Ending an experiment based on the potential value remaining is nice because it handles ties well. For example, in an experiment with many arms, it can happen that two or more arms perform about the same, so it does not matter which is chosen. You wouldn't want to run the experiment until you found the optimal arm (because there are two optimal arms). You just want to run the experiment until you're sure that switching arms won't help you very much.

More complex experiments
The multi-armed bandit's edge over classical experiments increases as the experiments get more complicated. You probably have more than one idea for how to improve your products and sales models, so you probably have more than one variation that you'd like to test. Let's assume you have 5 variations plus the original. You're going to do a calculation where you compare the original to the largest variation, so we need to do some sort of adjustment to account for multiple comparisons. The Bonferroni correction is an easy adjustment, which can be implemented by dividing the significance level of the hypothesis test by the number of arms. Thus we do the standard power calculation with a significance level of .05 / (6 - 1), and find that we need 15,307 observations in each arm of the experiment. With 6 arms that's a total of 91,842 observations. At 100 visits per day the experiment would have to run for 919 days (over two and a half years). In real life it usually wouldn't make sense to run an experiment for that long, but we can still do the thought experiment as a simulation.