



INSTITUTE OF COGNITIVE SCIENCE
INTUIT DATA ENGINEERING AND ANALYTICS

Bachelor's Thesis

Enhanced A/B Testing

Andrea Suckro

May 26, 2015

First supervisor: Prof. Dr. Frank Jäkel
Second supervisor: Mita Mahadevan

Enhanced A/B Testing

basic ab testing what this thesis adds to the framework how this is used in costumer ...

Contents

1	Introduction	4
1.1	And a section	4
1.2	another one	4
2	Theory	5
2.1	Basic Assumptions and Setup	5
2.2	Best Bucket	5
2.2.1	Analytic	6
2.2.2	Approximation	6
2.2.3	Sampling	7
2.2.4	n Buckets	7
2.3	Best Assignment	8
2.3.1	Uniform	8
2.3.2	Random	8
2.3.3	Entropy	8
3	Methods	9
3.1	Data generation	10
3.2	Data analysis	10
3.3	Data representation	10
4	Analysis	11
4.1	And a section	12
4.2	another one	12
5	Discussion	13
	Appendices	I

Chapter 1

Introduction

1.1 And a section

1.2 another one

Chapter 2

Theory

2.1 Basic Assumptions and Setup

A test is set up with at least two buckets and a sampling rate that determines what percentage of all possible users should participate in the testing. If for a new user was decided that she takes part in the test the next part is assigning her to a bucket. Later in this chapter different mechanisms for this step will be discussed and evaluated. Now certain actions of the users are recorded.

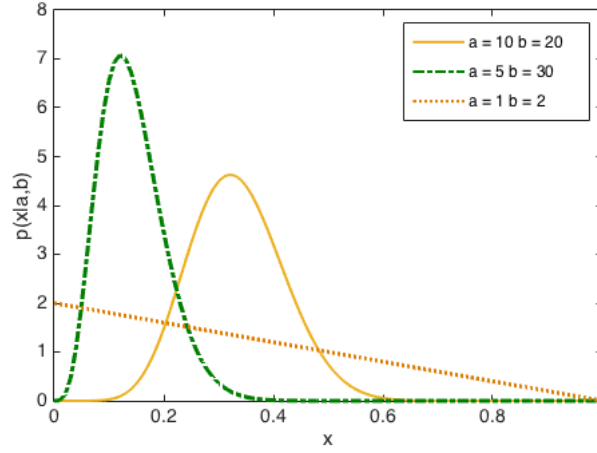
One is generating actions with a probability q_1 the other generates them with a probability q_2 . The prior distribution for all q is a Beta distribution, which is a convenient choice because it is a prior distribution for binomial proportions (the Beta distribution is the conjugate family for the binomial likelihood). We also assume no prior knowledge when creating the test case (uninformative prior $Beta(1, 1)$). N_1 customers get assigned to Bucket 1 and N_2 to Bucket 2. They generate k_1 and k_2 actions. The posterior distributions for each individual case is:

$$q = f(k|N) \propto f(N|k)f(k)$$
$$q_1 = Beta(k_1 + \alpha, N_1 - k_1 + \beta) = \frac{x^{k_1 + \alpha - 1}(1 - x)^{N_1 - k_1 + \beta - 1}}{B(k_1 + \alpha, N_1 - k_1 + \beta)}$$
$$q_2 = Beta(k_2 + \alpha, N_2 - k_2 + \beta) = \frac{x^{k_2 + \alpha - 1}(1 - x)^{N_2 - k_2 + \beta - 1}}{B(k_2 + \alpha, N_2 - k_2 + \beta)}$$

So each new assignment and each new event would directly alter the distributions parameters. *Questions: What is with returning users? Or users with a high action rate? Is every user a new one?* What is the posterior for $P(q_1 > q_2 | k_1, N_1, k_2, N_2)$ - that Bucket 1 performs better than Bucket 2 given the collected data? There are in principal two different ways to get an answer.

2.2 Best Bucket

Finding the best Bucket among the other test scenarios can be achieved in several ways. It follows a set of methods. Each will be explained first for the simple case with two Buckets and then expanded to the general setting with n different cases.



2.2.1 Analytic

In general for two random variables X, Y with corresponding probability density function f_X, f_Y and cumulative density function F_X, F_Y it holds:

$$\begin{aligned}
 P(X \geq Y) &= \iint_{[x \geq y]} f_X(x) f_Y(y) dy dx \\
 &= \int \int_{-\infty}^x f_X(x) f_Y(y) dy dx \\
 &= \int_{-\infty}^{\infty} f_X(x) F_Y(x) dx
 \end{aligned}$$

In the paper "Numerical Computation of Stochastic Inequality Probabilities" the author John D. Cook uses symmetries of the distribution to arrive at a set of equations that can be used to calculate the problem recursively.

$$\begin{aligned}
 g(k1, N1, k2, N2) &= P(X > Y) \\
 h(k1, N1, k2, N2) &= \frac{B(k1 + k2, N1 + N2)}{B(k1, N1)B(k2, N2)}
 \end{aligned}$$

From that one could calculate a base case for a small sample and then continue with:

$$\begin{aligned}
 g(k1 + 1, N1, k2, N2) &= g(k1, N1, k2, N2) + h(k1, N1, k2, N2)/k1 \\
 g(k1, N1 + 1, k2, N2) &= g(k1, N1, k2, N2) - h(k1, N1, k2, N2)/N1 \\
 g(k1, N1, k2 + 1, N2) &= g(k1, N1, k2, N2) - h(k1, N1, k2, N2)/k2 \\
 g(k1, N1, k2, N2 + 1) &= g(k1, N1, k2, N2) + h(k1, N1, k2, N2)/N2
 \end{aligned}$$

2.2.2 Approximation

For *sufficient large* samples one could again approximate the Beta Distribution with a normal distribution. This means that the following equation should be fulfilled for $B(k, N)$:

$$\frac{k+1}{k-1} \approx 0, \frac{N+1}{N-1} \approx 0$$

The Gaussian Distribution to a given Beta Distribution has the following shape:

$$B(k, N) \approx N \left(\frac{k}{k+N}, \sqrt{\frac{kN}{(k+N)^2(k+N+1)}} \right)$$

The inequality for this case then could be solved by:

$$\begin{aligned} P(X > Y) &= P(0 > Y - X) \\ &= P(0 > \mu_Y - \mu_X + (\sigma_X^2 + \sigma_Y^2)^{\frac{1}{2}} Z) \\ &= P \left(Z < \frac{\mu_X - \mu_Y}{(\sigma_X^2 + \sigma_Y^2)^{\frac{1}{2}}} \right) \\ &= \Phi \left(\frac{\mu_X - \mu_Y}{(\sigma_X^2 + \sigma_Y^2)^{\frac{1}{2}}} \right) \end{aligned}$$

I am not sure if this is really a smart way to do it but what would be maybe interesting is finding out which of the possibilities is the fastest for growing N .

2.2.3 Sampling

One could also sample from the different distributions and take the mean of the of the samples to determine the best distribution. The mean of a Beta distribution $B(k, N)$ is given by:

$$E[X] = \frac{k}{k+N}$$

Open question here: how exactly is the sample size related to the accuracy of the estimate?

2.2.4 n Buckets

The generalization would again depend on the chosen method to evaluate the inequality.

Analytic

I found examples for 3 Beta distributed random variables. But basically any formula would calculate the probability that one variable is greater than the others which results in $n - 1$ equations to solve the problem for n buckets. The same number of equations would occur if we would compute the inequality pairwise with the already presented formula.

Sampling

Additional Buckets would not change the underlying computation but would result in additional draws and new means to compare.

2.3 Best Assignment

It is not obvious what the best strategy is to assign a new user to a specific bucket. I will try to compare several intuitive approaches in this section. Interesting cases for the different strategies are: How fast does each strategy converge? What happens if there is no clear winner?

2.3.1 Uniform

2.3.2 Random

2.3.3 Entropy

The best assignment for a not yet tracked user would be a bucket that minimizes the entropy over the different buckets. The entropy of a Beta Distribution is given by:

$$\begin{aligned} h(q) &= E[-\ln(f(x; k, N))] \\ &= \int_0^1 -f(x; k, N) \ln(f(x; k, N)) dx \\ &= \ln(B(k, N)) - (k-1)\psi(k) - (N-1)\psi(N) + (k+N-2)\psi(k+N) \end{aligned}$$

Where ψ is the Digamma function. One could calculate the difference from $h(q_x)$ to $h(q'_x)$ where $x \in 1 \dots n$ number of Buckets and q'_x represents the distribution with the newly assigned user and assign the user the bucket where $\max(|h(q_x) - h(q'_x)|)$.

Methods

Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem
ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum
dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit
amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem
ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor
sit amet.Lorem ipsum dolor sit amet.

[illegible]

Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem
 ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum
 dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit
 amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem
 ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor
 sit amet.Lorem ipsum dolor sit amet.

[illegible][illegible]

sit amet.Lorem ipsum dolor sit amet.

Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.

Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.

Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.

Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.

foo

3.1 Data generation

3.2 Data analysis

3.3 Data representation

Analysis

[illegible]

Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem
ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum
dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit
amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem
ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor sit amet.Lorem ipsum dolor
sit amet.Lorem ipsum dolor sit amet.

[illegible][illegible][illegible]

[illegible][illegible][illegible][illegible]

4.1 And a section

4.2 another one

Discussion

13

[illegible]

List of Figures

Declaration of Authorship

I hereby certify that the work presented here is, to the best of my knowledge and belief, original and the result of my own investigations, except as acknowledged, and has not been submitted, either in part or whole, for a degree at this or any other university.

Osnabrück, May 26, 2015

