A Bayesian Approach to AB Testing [Draft]

Introducing a Python Engine for Testing

Hugo Pibernat and Nuria Duran

April 10, 2014

1 Introduction

The main goal of our AB-Test is to decide which of two different versions of the same product performs better on a specific set of metrics. In this report we present a Bayesian approach to tackle this objective.

A Bayesian model includes a prior distribution which enables us to provide information about the parameters KPI_A and KPI_B governing our data *before* observing the data. The information included in the data is taken into account by means of the likelihood function. Combining both the prior and the likelihood function results in a posteriori distribution, a probability distribution of the parameters (the metric of each group, in this case) which will help us to decide which of the version performs better.

The (joint) posterior probability distribution is defined as follows:

$$P(\text{KPI}_A, \text{KPI}_B | data) \tag{1}$$

In order to compute this distribution, we will need the Bayes theorem, which states that

$$P(X|Y) = \frac{P(Y|X) \cdot P(X)}{P(Y)} \tag{2}$$

2 Binomial Variables: Modelling Conversion Rate

The history of a user (UH) looks as follows:

$$< 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1 > \tag{3}$$

where 0 represents a non-converted visit and 1 represents a converted visit. Hence,

$$\text{UH} \sim Bin(n, \theta) \tag{4}$$

where $n \in \mathbb{Z}^+$ and $\theta \in [0, 1]$ is the parameter we want to model with the posteriori distribution once given the priori distribution $P(\theta)$ and the likelihood function. We usually call this parameter Conversion Rate (CR).

2.1 Choosing the prior distribution

The prior distribution $P(\theta)$ enables us to include some knowledge in the final distribution of the KPI that we are modelling. There are several considerations when choosing the prior. However, for the sake of simplicity, in this case we will use the Beta distribution, which has a support of $[0, 1]$, exactly the range of values of CR:

$$Beta_{\alpha,\beta}(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)} \tag{5}$$

where α and β are the parameters of the distribution and B is the beta function, defined as follows:

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} B(\alpha,\beta) \tag{6}$$

A particularly useful set of parameters is $\alpha = \beta = 1$ which results in the uniform distribution. That is, a priori it is considered that all the values a parameter can take are equally likely to be taken.

2.2 Obtaining the posteriori distribution

To compute the posteriori distribution, apart from the priori distribution, we need the likelihood function, the probability mass function of the Binomial distribution in our case:

$$P(X = k) = \binom{n}{k} \theta^k (1-\theta)^{n-k} \tag{7}$$

In order to compute the posteriori distribution, we will use the Bayes theorem:

$$P(\theta|n,k) = \frac{P(n,k|\theta) \cdot P(\theta)}{P(n,k)} \tag{8}$$

We therefore need to obtain $P(n,k|\theta)$ and $P(\theta)$. As it will be seen later, the value of P(n,k) does not actually need to be computed.

First, we have

$$P(n,k|\theta) = \binom{n}{k} \theta^k (1-\theta)^{n-k} \tag{9}$$

Back to Equation 8 we now have that

$$P(\theta|n,k) =$$

$$= \frac{P(n,k|\theta) \cdot P(\theta)}{P(n,k)} =$$

$$= \frac{\binom{n}{k} \theta^k (1-\theta)^{n-k} \cdot \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}}{P(n,k)} = \tag{10}$$

$$= \frac{\binom{n}{k}}{P(n,k)B(\alpha,\beta)} \cdot \theta^{k+\alpha-1}(1-\theta)^{n-k+\beta-1}$$

We can rewrite the previous equation as:

$$P(\theta|n,k) = \lambda \cdot \theta^{k+\alpha-1}(1-\theta)^{n-k+\beta-1} \tag{11}$$

where

$$\lambda = \binom{n}{k} \frac{1}{P(n,k)B(\alpha,\beta)} \tag{12}$$

Hence, our posteriori distribution for the θ is:

$$P(\theta|n,k) = Beta_{n+\alpha,n-k+\beta}(\theta) \tag{13}$$

Or, in terms of sessions s, purchases p, and conversion rate R:

$$P(R|s,p) = Beta_{s+\alpha,s-p+\beta}(\theta) \tag{14}$$

2.3 Running an AB-Test on Conversion Rate

Let's focus now on obtaining the probability that one conversion rate group is higher than the other (although the result is easily generalisable to n test groups). That is:

$$P(R_A > R_B|\text{data}) = \iint_{R_A>R_B} P(R_A, R_B|\text{data}) \, dR_A \, dR_B \tag{15}$$

where $P(R_A, R_B|\text{data})$ is the joint distribution of both conversion rates. If we can assume independence of our tests groups[1], then the following holds:

$$P(R_A, R_B|\text{data}) =$$

$$= P(R_A|\text{data}_A) \cdot P(R_B|\text{data}_B) =$$

$$= \frac{P(\text{data}_A|R_A)P(R_A)}{P(\text{data}_A)} \cdot \frac{P(\text{data}_B|R_B)P(R_B)}{P(\text{data}_B)} = \tag{16}$$

$$= \frac{P(\text{data}_A|R_A)P(R_A)P(\text{data}_B|R_B)P(R_B)}{P(\text{data}_A)P(\text{data}_B)}$$

where $P(\text{data}_A|R_A)$, $P(R_A)$, on the one hand, and $P(\text{data}_B|R_B)$, $P(R_B)$, on the other hand, stand for the *likelihood function* and the *prior distribution*, and they can be obtained as explained above. Look in the *BayesianABTest.py* file for an implementation of this and the other methods explained in this paper. Or run some of the examples in *examples_ABTests.py* to see the method in action.

3 Gaussian Variables: Modelling Attempts per Level

When our data follow a Gaussian Distribution

$$X \sim N(\mu, \sigma^2)$$

the prior distribution becomes slightly more complicated. In this case we need to obtain a joint distribution of the form:

$$P(\mu, \sigma^2) \tag{17}$$

---

[1]Although it is quite difficult that with our sparse social graph any two events are really independent, this is still quite a general assumption in most of our tests.

From Gelman et al. [4], we know that it can be described as follows:

$$P(\mu, \sigma^2) = P(\sigma^2)P(\mu|\sigma^2) \tag{18}$$

where $P(\sigma^2)$ is an Inverse Gamma Distribution and $P(\mu|\sigma^2)$ is also a Gaussian Distribution.

Following Gelman's indications we can sample the posteriori distribution and then run the AB-Test using the code from *BayesianABTest.py*.

4 Log-Normal Variables: Modelling Spend

When, for instance, we are modelling user spend, data are positive. Therefore, a Gaussian (or Normal) distribution is not appropriate due to its $(-\infty, +\infty)$ support.

In those situations, a Log-Normal distribution might fit our data more accurately. A Log-Normal Distribution is such that if X is distributed log-normally, then $ln(X)$ is distributed normally:

$$\text{if } X \sim Log - N(\mu, \sigma^2) \text{ then } ln(X) \sim N(\mu, \sigma^2) \tag{19}$$

In a Log-Normal distribution the μ and σ^2 are not the mean and the variance. The mean, variance, median, and mode from a Log-Normal distribution are computed as follows:

- Mean: $e^{\mu + \frac{\sigma^2}{2}}$

- Variance: $(e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$

- Median: e^{μ}

- Mode: $e^{\mu - \sigma^2}$

An examplary code to sample the posteriori distribution can be found in *BayesianABTest.py*.

5 Modelling Spend per Session with a Two-Part Model Approach

When our data contain a substantial amount of zeroes, a log-normal distribution is not appropriate, as its support is strictly positive. In this situation, we can first model the zeros vs. not zeros with a Binomial distribution, then the non-zeroes with a Log normal distribution, and obtain our final posteriori distribution as the product of these two.

The following code runs an AB-Test on such a combination of variables: each set of variable values is randomly sampled, and then each random value from the first set is multiplied by one random value from the second set. Look at the file *BayesianABTest.py* for a possible implementation of this method.

6 References

1. Evan Miller. How Not To Run An A/B Test. Available at: `http://www.evanmiller.org/how-not-to-run-an-ab-test.html`

2. Chris Stucchio. *Analyzing conversion rates with Bayes Rule (Bayesian statistics tutorial).* Available at: `http://www.chrisstucchio.com/blog/2013/bayesian_analysis_conversion_rates.html`

3. Sergey Feldman. *Rich Relevance, Collection on Bayesian Testing.* Available at: `http://engineering.richrelevance.com/category/bayesian/`

4. Gelman, A. et al. *Bayesian Data Analysis*, 3rd Edition. CRC Press, 2013.