

Objective Bayesian Two Sample Hypothesis Testing for Online Controlled Experiments

Alex Deng
Microsoft
Redmond, WA
alex deng@microsoft.com

Abstract

As A/B testing gains wider adoption in the industry, more people begin to realize the limitations of the traditional frequentist null hypothesis statistical testing (NHST). The large number of search results for the query “Bayesian A/B testing” shows just how much the interest in the Bayesian perspective is growing. In recent years there are also voices arguing that Bayesian A/B testing should replace frequentist NHST and is strictly superior in all aspects. Our goal here is to clarify the myth by looking at both advantages and issues of Bayesian methods. In particular, we propose an objective Bayesian A/B testing framework for which we hope to bring the best from Bayesian and frequentist methods together. Unlike traditional methods, this method requires the existence of historical A/B test data to objectively learn a prior. We have successfully applied this method to Bing, using thousands of experiments to establish the priors.

Categories and Subject Descriptors

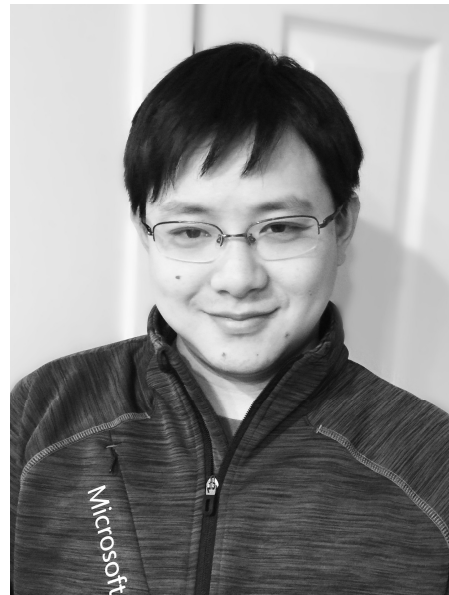
G.3 [Probability and Statistics]: Statistical Computing

Keywords

A/B testing; controlled experiments; Bayesian statistics; prior; objective Bayes; empirical Bayes; multiple testing; optional stopping

Short Bio

Alex Deng is a data and applied scientist on the Microsoft Analysis and Experimentation Team(A&E). The A&E team is responsible for one of the largest and best cutting-edge online experimentation system in the industry with the mission of accelerating innovation through trustworthy analysis and experimentation.



Deng finished his Ph.D. study in statistics at Stanford in 2010. Since then he has been a member of the A&E team and is currently leading a team focusing on methodological improvements of the experimentation platform and tackling the related engineering challenges. His works in this area have been published in proceedings of KDD, WWW, WSDM and JSM.