Date: Tuesday, April 23, 2019

# Overview of Bayesian Bandits

Two main types: 1) Proportions or rates
↳ CTR, conversion rate, dropout rate

2) Continous data (mean testing)
↳ session duration, user engagement, etc.

Main Strategy: a) Provide a prior for the metric of interest.

- Proportion: $U(0,1)$ (non-informative objective prior)
- Continous: $N(0,1)$

b) Run a for loop to allocate to differing arms:

In each iteration, we i) Choose an arm
ii) Observe reward r
iii) Update prior with observed reward r
↳ leads to a new posterior for the chosen arm.
(We use a conjugate family.)

iv) Simulate from each posterior to calculate which arm is optimal and repeat process.
↳ becomes the choice for step (i) after the first iteration.

* In application, any bandit algorithm is online:

The allocation of units (pulls of a slot machine in our example) at round t depends on what happened in previous rounds.