

# MKTG 596: Adaptive Experimental Design Methods

Lalit Jain

April 19, 2023

All solutions must be written using L<sup>A</sup>T<sub>E</sub>X. Any code for a problem should be included. Please put those in an appendix section for the solution. I recommend using [overleaf.com](https://overleaf.com) for typesetting your solutions.

## 1 Bayesian Hypothesis Testing

We assume that we are given two hypothesis,  $H_0$  and  $H_1$  and that with some prior probability  $\pi_0$  that  $H_0$  is True and with some probability  $\pi_1$ , that  $H_1$  is true. In addition, we assume that at each time we observe data  $X_s$  drawn from the true distribution and let  $\mathbf{X}_t$  be the data up to time  $t$

We define the *Posterior-Odds*

$$PO_t = \frac{\mathbb{P}(H_1|\mathbf{X}_t)}{\mathbb{P}(H_0|\mathbf{X}_t)}$$

and the Bayes Factor  $\Lambda_t = \frac{\mathbb{P}(\mathbf{X}_t|H_1)}{\mathbb{P}(\mathbf{X}_t|H_0)}$

1. Show that  $PO_t = \frac{\pi_1}{\pi_0} \Lambda_t$ .

In sequential Bayesian Hypothesis Testing, we fix a threshold  $\alpha$  and continue to collect data until some time  $\tau$  where  $PO_\tau \geq \alpha$ , in which case we accept  $H_1$ , or if it drops under  $1/\alpha$  in which case we accept  $H_0$ . Let this stopping time be denoted  $\tau$ . The quantity  $\mathbb{P}(H_0|PO_\tau = \alpha)$  has a natural interpretation as the (*Bayesian*) *False Discovery Rate*, i.e. the probability of being  $H_0$  though we conclude  $H_1$  of this procedure (make sure to contrast it to the one from class!). We will see that  $\mathbb{P}(H_0|PO_\tau = \alpha) = 1/(\alpha + 1)$

2. Define  $\alpha' = \alpha \times \pi_0 / \pi_1$ . Show that the claim is equivalent to the statement that  $\frac{\mathbb{P}(\Lambda_\tau = \alpha' | H_1)}{\mathbb{P}(\Lambda_\tau = \alpha' | H_0)} = \alpha'$ . In Bayesian language, this is known as *calibration*.
3. Now use Wald's Ratio Identity to prove the previous. Hint: See [TNB14] - it pops up in the proof of the SPRT as well. If you are a bit confused, try to prove it at a fixed time before worrying about the anytime version.
4. Verify this result empirically. Construct a setting where  $H_0 = N(0, 1)$  and  $H_1 = N(\mu, 1)$  (play around with the value of  $\mu$ ). Assign  $\pi_0 = \pi_1 = .5$ . Run the procedure we described above for a large range of values and empirically check the result.

Due to this result, it's often claimed that Bayesian methods are immune to early stopping. Why?

## 2 UCB

Consider the following algorithm for the multi-armed bandit problem.

### Algorithm 1: UCB

**Input:** Time horizon  $T$ , 1-subGaussian arm distributions  $P_1, \dots, P_n$  with unknown means  $\mu_1, \dots, \mu_n$

**Initialize:** Pull each arm once. At any time let  $T_i(t)$  denote the number of times  $i$  has been pulled at time  $t$  and let  $T_i = T_i(T)$ .

**for:**  $t = 1, \dots, T$

Choose  $I_t = \arg \max_{i=1, \dots, n} \hat{\mu}_{i,t} + \sqrt{\frac{2 \log(nT^2)}{T_i(t)}}$

Observe  $X_{I_t,t} \sim P_{I_t}$ .  $T_{I_t}(t+1) \leftarrow T_{I_t}(t) + 1$ , update  $\hat{\mu}_{i,t+1}$

In the following exercises, we will compute the regret of the UCB algorithm and show it matches the regret bound from lecture. Without loss of generality, assume that the best arm is  $\mu_1$ . For any  $i \in [n]$ , define the *sub-optimality gap*  $\Delta_i = \mu_1 - \mu_i$ .

1. (Warm-up). Show that  $R_T = \sum_{i=1}^n \Delta_i \mathbb{E}[T_i]$ .

2. Consider the event

$$\mathcal{E} = \bigcap_{i \in [n]} \bigcap_{s \leq T} \left\{ |\hat{\mu}_{i,s} - \mu_i| \leq \sqrt{\frac{2 \log(2nT^2)}{s}} \right\}.$$

Show that  $\mathbb{P}(\mathcal{E}) \geq 1 - \frac{1}{T}$ .

3. Conditioned on event  $\mathcal{E}$ , show that  $T_i < \frac{8 \log(nT^2)}{\Delta_i^2}$  for  $i \neq 1$ .

4. Show that  $\mathbb{E}[T_i] \leq \frac{8 \log(nT^2)}{\Delta_i^2} + 1$ . When  $n \leq T$ , conclude by showing that  $R_T \leq \sum_{i=1}^n \left( \frac{24 \log(T)}{\Delta_i} + \Delta_i \right)$ .

## 3 Doubling Trick

Assume we have access to a fixed-horizon algorithm  $\mathcal{A}$  that depends on known the time-horizon  $T$  and the regret  $\mathcal{A}$  incurs is given by  $f(T)$ . In this exercise, we discuss a strategy to turn  $\mathcal{A}$  into an anytime algorithm.

1. Suppose that  $f(T) \leq \sqrt{T}$ . Come up with a strategy so that at anytime  $t$  the cumulative regret is bounded by  $C\sqrt{t}$  for some constant  $C$ .
2. Assume that  $f(T) \log(t)$ . Come up with a strategy so that at anytime  $t$  the cumulative regret is bounded by  $C \log t$  for some constant  $C$ .

## 4 Implementation of MAB

Implement Successive Elimination from Lecture using a confidence interval referred to as  $c(t)$  (specified below). Consider a setting with  $K = 10$  arms, with arm  $i$  being  $N(\mu_i, 1)$  where  $\mu_i = 1 - 1/i^\alpha$ . Run all combinations of the following settings:

1.  $\alpha \in \{1/2, 1\}$
2.  $c(t)$  the anytime Chernoff bound  $c(t) = O\left(\sqrt{\frac{2 \log(2Kt^2/\delta)}{t}}\right)$ , and  $c(t)$  the MSPRT with  $\rho = 1$ .

For each combination, run 20 repetitions (you may want to look into `multiprocessing` in Python). Report

- The stopping time of the algorithm with a 95% confidence interval.
- The regret at the stopping time of the algorithm with a 95% confidence interval.
- A plot of the total regret over time of each run (do them all on one plot with a lower transparency value), the average regret over the runs and confidence intervals.

Interpret your results. What would you recommend in practice?

## 5 Implementation of Sequential Testing

Consider the Hypothesis Test  $X_1, X_2, \dots \sim N(\mu, 1)$  where  $H_0 : \mu = 0, H_1 : \mu = \Delta$  where  $\Delta \in \{.01, .1, .5\}$

1. Implement the SPRT with  $\alpha = .05, \beta = .05$ . Do 100 repetitions, in half of which the null is True in half of which the alternative is True. What is your average stopping time, what is your empirical Type 1 error and what is your Power? Compare to a fixed horizon test. On average how much time did you save?
2. Now implement the MSPRT and formulate a hypothesis test. Try several different values of  $\rho$  and contrast them. Compare to the results from the previous exercise.

## References

- [TNB14] Alexander Tartakovsky, Igor Nikiforov, and Michele Basseville. *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, 2014.