How to run an adaptive field experiment

Maximilian Kasy

December 2021

Is experimentation on humans ethical?

Deaton (2020):

> *Some of the RCTs done by western economists on extremely poor people [...] could not have been done on American subjects.*
>
> *It is particularly worrying if the research addresses questions in economics that appear to have no potential benefit for the subjects.*

Do our experiments have enough power?

Ioannidis et al. (2017):

> *We survey 159 empirical economics literatures that draw upon 64,076 estimates of economic parameters reported in more than 6,700 empirical studies. Half of the research areas have nearly 90% of their results under-powered. The median statistical power is 18%, or less.*

Are experimental sites systematically selected?

Andrews and Oster (2017):

> *[...] the selection of locations is often non-random in ways that may influence the results. [...] this concern is particularly acute when we think researchers select units based in part on their predictions for the treatment effect.*

Claim: Adaptive experimental designs can partially address these concerns

1. Ethics and **participant welfare**:

   Bandit algorithms are designed to maximize participant outcomes, by shifting to the best performing options at the right speed.

2. **Statistical power** and publication bias:

   Exploration Sampling (Kasy and Sautmann, 2021), is designed to maximize power for distinguishing the best policy, by focusing attention on competitors for the best option.

3. **Political economy**, site selection, and external validity:

   Related to the ethical concerns: Design experiments that maximize the stakeholders' goals (where appropriate). This might allow us to reduce site selectivity, by making experiments more widely acceptable.

What is adaptivity?

- Suppose your experiment takes place over time.
- Not all units are assigned to treatments at the same time.
- You can observe outcomes for some units
  before deciding on the treatment for later units.
- Then treatment assignment can depend on earlier outcomes,
  and thus be *adaptive*.

Why adaptivity?

- Using more information is always better than using less information, when making (treatment assignment) decisions.
- Suppose you want to
  1. **Help participants**
     $\Rightarrow$ Shift toward the best performing option.
  2. **Learn the best treatment**
     $\Rightarrow$ Shift toward best candidate options, to maximize power.
  3. **Estimate treatment effects**
     $\Rightarrow$ Shift toward treatment arms with higher variance.
- Adaptivity allows us to achieve better performance with smaller sample sizes.

When is adaptivity useful?

1. **Time till outcomes are realized**:
   - Seconds? (Clicks on a website.)
     Decades? (Alzheimer prevention.)
     Intermediate? (Many settings in economics.)
   - Even when outcomes take months, adaptivity can be quite feasible.
     Splitting the sample into a small number of waves already helps a lot.
     Surrogate outcomes (discussed later) can shorten the wait time.
2. **Sample size and effect sizes**:
   - Algorithms can adapt, if they can already learn something before the end of the experiment.
   - In very underpowered settings, the benefits of adaptivity are smaller.
3. **Technical feasibility**:
   - Need to create a pipeline:
     Outcome measurement - belief updating - treatment assignment.
   - With apps and mobile devices for fieldworkers, that is quite feasible, but requires some engineering.

Papers this talk is based on

- Kasy, M. and Sautmann, A. (2021).
  Adaptive treatment assignment in experiments for policy choice.
  Econometrica

- Caria, S., Gordon, G., Kasy, M., Osman, S., Quinn, S., and Teytelboym, A. (2021).
  An Adaptive Targeted Field Experiment:
  Job Search Assistance for Refugees in Jordan.
  Working paper.

- Kasy, M. and Teytelboym, A. (2021).
  Learning by matching.
  Revise and resubmit, Econometrics Journal.

- Kasy, M. and Teytelboym, A. (2020).
  Adaptive targeted disease testing.
  Oxford Review of Economic Policy.

Literature

- Statistical decision theory:
  Berger (1985),
  Robert (2007).
- Non-parametric Bayesian methods:
  Ghosh and Ramamoorthi (2003),
  Williams and Rasmussen (2006),
  Ghosal and Van der Vaart (2017).
- Stratification and re-randomization:
  Morgan and Rubin (2012),
  Athey and Imbens (2017).
- Adaptive designs in clinical trials:
  Berry (2006),
  FDA et al. (2018).
- Bandit problems:
  Weber et al. (1992),
  Bubeck and Cesa-Bianchi (2012),
  Russo et al. (2018).
- Regret bounds:
  Agrawal and Goyal (2012),
  Russo and Van Roy (2016).
- Best arm identification:
  Glynn and Juneja (2004),
  Bubeck et al. (2011),
  Russo (2016).
- Bayesian optimization:
  Powell and Ryzhov (2012),
  Frazier (2018).
- Reinforcement learning:
  Ghavamzadeh et al. (2015),
  Sutton and Barto (2018).
- Optimal taxation:
  Mirrlees (1971),
  Saez (2001),
  Chetty (2009),
  Saez and Stantcheva (2016).

Setup

- Waves $t = 1, \ldots, T$, sample sizes N_t.
- Treatment $D \in \{1, \ldots, k\}$, outcomes $Y \in [0, 1]$, covariate $X \in \{1, \ldots, n_x\}$.
- Potential outcomes Y^d.
- Repeated cross-sections:
  $(Y_{it}^1, \ldots, Y_{it}^k, X_{it})$ are i.i.d. across both i and t.
- Average potential outcomes:

$$\theta^{dx} = E[Y_{it}^d | X_{it} = x].$$

Adaptive targeted assignment

- The algorithms I will discuss are Bayesian.
- Given all the information available at the beginning of wave t, form posterior beliefs P_t over θ.
- Based on these beliefs, decide what share p_t^{dx} of stratum x will be assigned to treatment d in wave t.
- How you should to pick these assignment shares depends on the objective you try to maximize.

Bayesian updating

- In **simple** cases, posteriors are easy to calculate in closed form.
  - Example: Binary outcomes, no covariates.
  - Assume that $Y \in \{0,1\}$, $Y_t^d \sim Ber(\theta^d)$. Start with a uniform prior for θ on $[0,1]^k$.
  - Then the posterior for θ^d at time $t+1$ is a $Beta$ distribution with parameters

$$\alpha_t^d = 1 + T_t^d \cdot \bar{Y}_t^d, \qquad \beta_t^d = 1 + T_t^d \cdot (1 - \bar{Y}_t^d).$$

- In more complicated cases, simulate from the posterior using MCMC (more later).
  - For well chosen **hierarchical** priors:
  - θ^{dx} is estimated as a **weighted average** of the observed success rate for d in x and the observed success rates for d across all other strata.
  - The **weights** are determined **optimally** by the observed amount of heterogeneity across all strata as well as the available sample size in a given stratum.

Objective I: Participant welfare

- **Regret**: Difference in average outcomes from decision d versus the optimal decision,

$$\Delta^{dx} = \max_{d'} \theta^{d'x} - \theta^{dx}.$$

- Average in-sample regret:

$$\bar{R}_\theta(T) = \frac{1}{\sum_t N_t} \sum_{i,t} \Delta^{D_{it}X_{it}}.$$

- **Thompson sampling**
  - Old proposal by Thompson (1933).
  - Popular in online experimentation.
- Assign each treatment with probability equal to the posterior probability that it is optimal, given $X = x$ and given the information available at time t.

$$p_t^{dx} = P_t\left(d = \underset{d'}{\operatorname{argmax}}\ \theta^{d'x}\right).$$

Thompson sampling is efficient for participant welfare

- **Lower bound** (Lai and Robbins, 1985):
  Consider the Bandit problem with binary outcomes and any algorithm. Then

$$\liminf_{T\to\infty} \frac{T}{\log(T)} \bar{R}_\theta(T) \geq \sum_d \frac{\Delta^d}{kl(\theta^d, \theta^*)},$$

  where $kl(p, q) = p \cdot \log(p/q) + (1-p) \cdot \log((1-p)/(1-q))$.

- **Upper bound for Thompson sampling** (Agrawal and Goyal, 2012):
  Thompson sampling achieves this bound, i.e.,

$$\liminf_{T\to\infty} \frac{T}{\log(T)} \bar{R}_\theta(T) = \sum_d \frac{\Delta^d}{kl(\theta^d, \theta^*)}.$$

Mixed objective: Participant welfare and point estimates

- Suppose you care about both participant welfare, and precise point estimates / high power for all treatments.
- In Caria et al. (2020), we introduce **Tempered Thompson sampling**: Assign each treatment with probability equal to

$$\tilde{p}_t^{dx} = (1-\gamma) \cdot p_t^{dx} + \gamma / k.$$

  Compromise between full randomization and Thompson sampling.

Tempered Thompson trades off participant welfare and precision

We show in Caria et al. (2020):

- **In-sample regret** is (approximately) proportional to the share γ of observations fully randomized.
- The **variance** of average potential outcome estimators is proportional
  - to $\frac{1}{\gamma/k}$ for sub-optimal d,
  - to $\frac{1}{(1-\gamma)+\gamma/k}$ for conditionally optimal d.
- The variance of **treatment effect** estimators, comparing the conditional optimum to alternatives, is therefore decreasing in γ.
- An **optimal** choice of γ **trades off** regret and estimator variance.

Objective II: Policy choice

- Suppose you will **choose a policy** after the experiment, based on posterior beliefs,

$$d_T^* \in \underset{d}{\operatorname{argmax}}\ \hat{\theta}_T^d, \qquad \hat{\theta}_T^d = E_T[\theta^d].$$

- Evaluate experimental designs based on expected welfare (ex ante, given $\boldsymbol{\theta}$).
- Equivalently, expected **policy regret**

$$\Delta^{d_T^*} = \sum_d \Delta^d \cdot \mathbf{1}\left(d_T^* = d\right), \qquad \Delta^d = \max_{d'} \theta^{d'} - \theta^d.$$

- In Kasy and Sautmann (2021), we introduce **Exploration sampling**: Assign shares q_t^d of each wave to treatment d, where

$$q_t^d = S_t \cdot p_t^d \cdot (1 - p_t^d),$$

$$p_t^d = P_t\left(d = \underset{d'}{\operatorname{argmax}}\ \theta^{d'}\right), \qquad S_t = \frac{1}{\sum_d p_t^d \cdot (1 - p_t^d)}.$$

Exploration sampling is efficient for policy choice

- We show:
  - The posterior probability p_t^d that each treatment is optimal goes to 0 at the same rate for all sub-optimal treatments.
  - Posterior expected policy regret also goes to 0 at the same rate.
  - No other algorithm can achieve a faster rate.
- Key intuition of proof: Equalizing power.
  1. Suppose p_t^d goes to 0 at a faster rate for some d.
     Then exploration sampling stops assigning this d.
     This allows the other treatments to "catch up."
  2. Balancing the rate of convergence implies efficiency.

Aside: There were some errors in our original proof, but the main result holds up.

Inference

- Inference has to take into account adaptivity, in general.
- Example:
  - Flip a fair coin.
  - If head, flip again, else stop.
  - Probability distribution: 50% tail-stop, 25% head-tail, 25% head-head.
  - Expected share of heads?

$$.5 \cdot 0 + .25 \cdot .5 + .25 \cdot 1 = .375 \neq .5.$$

- But:
  1. **Bayesian** inference works without modification.
  2. **Randomization tests** can be modified to work in adaptive settings.
  3. **Standard inference** (e.g., t-tests) works under some conditions.

Bayesian inference

- The likelihood, and thus the posterior,
  are **not affected by adaptive** treatment assignment.

- Claim: The likelihood of $(D_1, \ldots, D_M, Y_1, \ldots, Y_M)$ equals $\prod_i P(Y_i|D_i, \boldsymbol{\theta})$,
  up to a constant that does not depend on $\boldsymbol{\theta}$.

- Proof: Denote $H_i = (D_1, \ldots, D_{i-1}, Y_1, \ldots, Y_{i-1})$. Then

$$
\begin{aligned}
P(D_1, \ldots, D_M, Y_1, \ldots, Y_M|\boldsymbol{\theta}) &= \prod_i P(D_i, Y_i|H_i, \boldsymbol{\theta}) \\
&= \prod_i P(D_i|H_i, \boldsymbol{\theta}) \cdot P(Y_i|D_i, H_i, \boldsymbol{\theta}) \\
&= \prod_i P(D_i|H_i) \cdot P(Y_i|D_i, \boldsymbol{\theta}).
\end{aligned}
$$

Randomization inference

- Strong null hypothesis: $Y_i^1 = \ldots = Y_i^k$.
- Under this null, it is easy to re-simulate the treatment assignment: Just let your assignment algorithm run with the data, switching out the treatments.
- Do this many times, re-calculate the test statistic each time.
- Take the $1 - \alpha$ quantile across simulations as critical value.
- This delivers **finite-sample exact inference** for any adaptive assignment scheme.

T-tests and F-tests

- As shown above, sample averages in treatment arms are, in general, biased, cf. Hadad et al. (2019).
- But: **Under some conditions**, the **bias is negligible** in large samples.
- In particular, suppose
  1. $\left(\sum_{i,t} \mathbf{1}(D_{it} = d)\right) / \tilde{N}_T^d \to^p 1$
  2. $\tilde{N}_T^d$ is non-random and goes to ∞.
- Then the standard law of large numbers and central limit theorem apply. T-tests can ignore adaptivity (Melfi and Page, 2000).
- This works for Exploration Sampling, Tempered Thompson sampling: Assignment shares are bounded away from 0.
- This does not work for typical Bandit algorithms (e.g. Thompson sampling): Assignment shares for sub-optimal treatments go to 0 too fast.

Data pipeline

Example pipeline for one wave (cf. Caria et al. 2020):

1. On a **central machine**, update the prior based on available data.
2. Calculate treatment assignment probabilities for each stratum.
3. Upload these to a **web-server**.
4. Field workers encounter participants,
   and enter participant covariates on a **mobile device**.
5. The mobile device assigns a treatment to participants,
   based on their covariates and the downloaded assignment probabilities.
6. A bit later, outcome data are collected,
   and transmitted to the central machine.

This is not too difficult, but it requires careful planning.
All steps should be **automated** for smooth implementation!

Surrogate outcomes

- We don't always observe the desired outcomes / measures of welfare quickly enough – or at all.
- Potential solution: **Surrogate** outcomes (Athey et al., 2019):
  - Suppose we want to maximize Y,
    but only observe other outcomes W,
    which satisfy the **surrogacy condition**

$$D \perp (Y^1, \ldots, Y^d) | W.$$

  - This holds **if all causal pathways** from D to Y **go through** W.
  - Let $\hat{y}(W) = E[Y|W]$, estimated from auxiliary data. Then

$$\begin{aligned} E[Y|D] &= E[E[Y|D, W]|D] \\ &= E[E[Y|W]|D] = E[\hat{y}(W)|D]. \end{aligned}$$

  - Implication: We can design algorithms that target maximization of $\hat{y}(W)$, and they will achieve the same objective.

Choice of prior

- One option: Informative prior, based on prior data or expert beliefs.
- I recommend instead: Default priors that are
  1. **Symmetric**: Start with exchangeable treatments, strata.
  2. **Hierarchical**: Model heterogeneity of effects across treatments, strata.
     Learn "hyper-parameters" (levels and degree of heterogeneity) from the data.
     $\Rightarrow$ Bayes estimates will be based on optimal partial pooling.
  3. **Diffuse**: Make your prior for the hyper-parameters uninformative.
- Example:

$$
\begin{aligned}
Y_{it}^{d} | (X_{it} = x, \theta^{dx}, \alpha^{d}, \beta^{d}) &\sim Ber(\theta^{dx}), \\
\theta^{dx} | (\alpha^{d}, \beta^{d}) &\sim Beta(\alpha^{d}, \beta^{d}), \\
(\alpha^{d}, \beta^{d}) &\sim \pi.
\end{aligned}
$$

MCMC sampling from the posterior

- For hierarchical models, posterior probabilities such as p_t^{dx} can be calculated by sampling from the posterior using **Markov Chain Monte Carlo**.
- General purpose Bayesian packages such as **Stan** make this easy:
  - Just specify your likelihood and prior.
  - The package takes care of the rest, using "Hamiltonian Monte Carlo."
- Alternatively, do it "**by hand**" (e.g. using our code):
  - Combine Gibbs sampling & Metropolis-Hasting.
  - Given the hyper-parameters, sample from closed-form posteriors for $\boldsymbol{\theta}$.
  - Given $\boldsymbol{\theta}$, sample hyper-parameters using Metropolis (accept/reject) steps.

The political economy of experimentation

- Experiments often involve some **conflict of interest**, that might prevent experimentation where it could be useful.
  - Academic experimenters: "We want to get estimates that we can publish."
  - Implementation partners: "We know what's best, so don't prevent us from helping our clients."
- Adaptive designs can **partially resolve** these conflicts
  1. Maintain controlled treatment assignment,
  2. but choose assignment probabilities to maximize stakeholder objectives.
- Conflicts can of course remain. e.g. Which outcomes to maximize? Choose carefully!

Conclusion

- Using adaptive designs in field experiments can have great **benefits**:
  1. More ethical, by helping participants as much as possible.
  2. Better power for a given sample size, by targeting policy learning.
  3. More acceptable to stakeholders, by aligning design with their objectives.
- Adaptive designs are practically **feasible**:
  We have implemented them in the field.
  E.g., labor market interventions for Syrian refugees in Jordan,
  and agricultural outreach for subsistence farmers in India.
- Implementation requires learning some new **tools**.
  - I have developed some software to facilitate implementation.
  - Interactive apps for treatment assignment, and source code for various designs.

`https://maxkasy.github.io/home/code-and-apps/`

Thank you!