

# A basic introduction to adaptive experiments

Ben Recht

November 19, 2025

One of the pesky parts of randomized experiments is setting the sample size. It's not desirable to constantly be experimenting on people. You want to have the smallest trials possible to ensure you satisfactorily determine the value of a treatment. What should this sample size be?

Let me try to pose a cost-benefit analysis. Assume we have a population of size  $N$  that might benefit from a treatment. From this population, we select a sample of size  $K$  uniformly at random. We run a randomized trial on this sample, and decide to accept or reject the treatment. We then apply this decision to the rest of the population. We record all their outcomes and compute the average effect across the entire population. Let's call that observed average effect  $E$ :

$$E = \frac{1}{N} \sum_{i=1}^N \text{outcome}_i(\text{assignment}_i) \quad (1)$$

To decide whether our aggregate outcome was good, we need to compare this average to something. One potential comparison is to think in the counterfactual manner of potential outcomes: what would have been the best single treatment to give to everyone if we had known all the outcomes in advance? That is, we imagine a world in which we made a blanket recommendation to the entire population to either receive the treatment or not, and then figure out which recommendation would have led to a better average outcome. The counterfactually optimal effect,  $E_c$  is then

$$E_c = \max_{T \in \{0,1\}} \frac{1}{N} \sum_{i=1}^N \text{outcome}_i(T) \quad (2)$$

We can relate the counterfactually optimal effect to the *average treatment effect* we studied in randomized trials.

$$ATE = \frac{1}{N} \sum_{i=1}^N \text{outcome}_i(1) - \text{outcome}_i(0) \quad (3)$$

The ATE is just the average outcome under treatment minus the average outcome under control. When the ATE is nonpositive, it is best not to recommend the treatment. That is, the counterfactually optimal treatment is 0. When the ATE is positive, it is best to recommend the treatment so the counterfactually optimal treatment is 1.

Now, what is the gap between  $E$  and  $E_c$ ? To bound this, let's switch to more formal notation. Let  $\Omega$  denote the population of units and  $S$  denote the sampled experimental group. For each

unit  $i \in \Omega$ , let the treatment assignment be denoted by  $z_i \in \{0, 1\}$  and the associated outcome by  $y_i(z) \in [0, 1]$ .

Let  $z_r(S)$  denote the treatment recommended after experimenting on  $S$ . If the optimal outcome in expression (2),  $z_*$ , is attained at  $z_* = 1$ , then the gap is

$$G = E - E_c = \frac{1}{N} \sum_{i=1}^N y_i(z_i) - y_i(1) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}[z_i = 0](y_i(0) - y_i(1)). \quad (4)$$

Similarly, if the counterfactually optimal treatment is when  $z_* = 0$ ,

$$G = \frac{1}{N} \sum_{i=1}^N \mathbf{1}[z_i = 1](y_i(1) - y_i(0)). \quad (5)$$

Hence, since  $\mathbb{E}[\mathbf{1}[z_i = z]] = \Pr[z_i = z]$  takes the same value for all indices  $i$ , the expected gap can be bounded in terms of the magnitude of the average treatment effect and the probability that the first unit does not receive the counterfactually optimal treatment. We have the identity

$$\mathbb{E}[G] = \Pr[z_1 \neq z_*] \Delta, \quad (6)$$

where  $\Delta$  is the magnitude of the average treatment effect ( $\Delta = |ATE|$ ).

To compute the expected value of this gap, it suffices to upper bound the probability that  $z_1 \neq z_*$ . We can do this with a simple union bound

$$\Pr[z_1 \neq z_*] \leq \Pr[1 \in S \text{ and } z_1 \neq z_*] + \Pr[1 \notin S \text{ and } z_r(S) \neq z_*] \quad (7)$$

$$= \frac{K}{2N} + \Pr[1 \notin S \text{ and } z_r(S) \neq z_*] \quad (8)$$

$$\leq \frac{K}{2N} + \Pr[z_r(S) \neq z_*]. \quad (9)$$

Hence, the expected gap between the welfare under our decision making and the counterfactual effect is

$$\mathbb{E}[G] \leq \left( \frac{K}{2N} + \Pr[z_r(S) \neq z_*] \right) \Delta. \quad (10)$$

It remains to find the sample size that makes this expected gap as small as possible.

The probability,  $p = \Pr[z_r(S) \neq z_*]$ , is closely connected to important concepts in statistical testing. In the world where the treatment is beneficial,  $p$  is the probability that your statistical test recommends not accepting the treatment. This is the probability of a false negative. In statistical jargon,  $1 - p$  is called the power of your test. Similarly, when the treatment is harmful,  $p$  is the probability of mistakenly recommending it. This is the probability of a false positive. It is more or less the probability of incorrectly rejecting the null hypothesis.

A good experimenter should set the sample size so their test has adequate power and also has a low false positive rate. What should the sample size be? Certainly the power increases with  $K$ , but how large should  $K$  be? If you knew the size of your effect in advance (you never do), you

could look up the appropriate sample size via a power calculation. To make the gap as small as possible, we need to find a bound for  $p$  in terms of  $K$  and then minimize the resulting expression.

Let's do just that. Let

$$\hat{\mu}_0 = \frac{2}{K} \sum_{i \in S} \mathbf{1}[z_i = 0] y_i(0) \quad (11)$$

$$\hat{\mu}_1 = \frac{2}{K} \sum_{i \in S} \mathbf{1}[z_i = 1] y_i(1). \quad (12)$$

Let's assume without loss of generality that  $z_\star = 1$ . The case when  $z_\star = 0$  follows from an identical analysis.  $z_r(S) = 0$  if  $\hat{\mu}_0 - \hat{\mu}_1 \geq 0$ . Hoeffding's inequality then lets us bound  $p$ . The expression  $\frac{1}{2}(\hat{\mu}_0 - \hat{\mu}_1)$  is a sum of  $K$  random variables drawn from a population of numbers with values between  $-1$  and  $1$ . Its mean is  $-\frac{\Delta}{2}$ . Hence, applying Hoeffding's inequality gives us the bound

$$\Pr[z_r(S) = 0] = \Pr[\hat{\mu}_0 - \hat{\mu}_1 \geq 0] \leq \exp\left(-\frac{K\Delta^2}{8}\right). \quad (13)$$

This in turn results in the following bound of the expected gap.

$$\mathbb{E}[G] \leq \left(\frac{K}{2N} + \exp\left(-\frac{K\Delta^2}{8}\right)\right) \Delta \quad (14)$$

Though  $K$  has to be an integer, we can be slightly unrigorous and treat it as a real number. Then, the gap (14) is minimized when

$$K = \frac{8}{\Delta^2} \log\left(\frac{N\Delta^2}{4}\right) \quad (15)$$

For people who have seen rules of thumb for power calculations before, this should look familiar. It corresponds to a power of at least

$$1 - \frac{4}{\Delta^2 N} \quad (16)$$

For most reasonable values of  $N$ , this is much larger than 80%. But note that the power is inversely proportional to the square of the effect size, just like in standard power calculators. With this choice of  $K$ , the corresponding bound on the gap is equal to

$$\mathbb{E}[G] \leq \frac{4}{N\Delta} \log\left(\frac{eN\Delta^2}{4}\right) \quad (17)$$

This expression says that as long as  $N$  is much larger than  $\Delta^{-1}$ , the procedure of decision-making by experiment achieves roughly the same average outcome as a clairvoyant who knew in advance the best treatment to apply.

It's worth noting that when  $\Delta$  is very small, the derived expression (17) is very large. Moreover, if you just choose to do nothing—not do any experiments nor apply the treatment—the gap is at most  $\Delta$ . If your treatment effect is marginal, there's no point in fretting about doing a huge

experiment. There are several cases where this might occur. For example, if  $N < 5.5\Delta^{-2}$ , then the right hand side of (17) is larger than  $\Delta$  for all values of  $K$ .

Similarly, optimal strategies in this framework call for randomized experiments far larger than those pragmatically applied in practice. If  $\Delta = 0.01$ , which is a common benefit observed in many medical trials, then the recommended experiment size for a population of 10 billion people is *one million people*. This means, by design, 500,000 people will receive a “suboptimal” assignment in the experimentation phase. Now, the resulting gap is  $10^{-6}$ . This could be more than you need. But it’s also worth questioning our current practices. If we were to only do an experiment with 60,000 people, a common pragmatic trial size, the gap would be 0.005. This is equal to the expected gap of flipping a coin and choosing whether to approve a treatment. This is one of my main critiques of the randomized trial paradigm: even when a trial is adequately powered to find an effect in the sample, it’s too small to make a decision about the net benefit for the general population.

We all know this, but sometimes we have to write these calculations down.

Unfortunately, this formula requires knowing the size of the treatment effect in advance of doing the experiment. That’s annoying. Statisticians and trialists know that this is just part of the elaborate theater you do when setting up an experiment. You guess the future to pick a sample size and hope for the best.

## 1 Adaptive Experimentation

Now, if you are willing to do more than one experiment, there’s a neat trick—first analyzed by Auer and Ortner [2010]—to get around needing to know the treatment effect. Start with a reasonably large sample size and do a randomized trial. If the observed gap between the treatment and control is 1, meaning everyone has the same outcome, stop experimenting and assign everyone to the best outcome. If the measured gap is smaller, then do another experiment. For this trial, make your sample size four times larger and test for a gap half the size of the previous experiment. If this again fails to distinguish treatment from control, repeat the process. Quadruple the sample size, halve your decision threshold. If you repeat this process of successively increasing your sample size, you are guaranteed to have the same gap you would have achieved if you knew the ATE in advance.

Let’s make the above paragraph precise.

---

### Algorithm 1 Successive Experimentation Algorithm

---

- 1: **for**  $m = 0, \dots, m_{\max}$  **do**
  - 2:     Run a randomized experiment with the next  $n_0 4^m$  units.
  - 3:     Let  $\hat{\mu}_j^{(m)}$  denote the mean of the group assigned to  $z = j$ .
  - 4:     **if**  $|\hat{\mu}_1^{(m)} - \hat{\mu}_0^{(m)}| \geq 2^{-m}$  **then**
  - 5:         Stop and apply the action with the larger mean to the remainder of the population.
  - 6:     **end if**
  - 7: **end for**
- 

This simple algorithm achieves the same scaling as we saw in the bound (17), without needing

to know or estimate the average treatment effect. We can formalize this result as a theorem.

**Theorem 1.1.** *If  $n_0 = 8 \log(N)$ , then the expected gap of Algorithm 1 is*

$$\mathbb{E}[G] \leq \frac{174 \log(N)}{N\Delta}. \tag{18}$$

This is the power of iterative experimentation. You can interpret this procedure as running increasingly large experiments until the confidence intervals around the treatment effect are bounded away from zero. The main difference from conventional experiments is that rather than having a p-value threshold of 0.05 and a power of 0.8, both error rates are set to  $1/(2N)$ . This is much smaller than convention, but it ensures we make good decisions for the general population.

This confidence interval interpretation provides a simple extension to the case where there are  $k$  possible treatment assignments, rather than 2. At each round, we'd run an RCT with  $k$  arms and eliminate all treatment options where the upper confidence bound of the treatment is below the highest lower confidence bound.

I'll also note that with a bit more complex schedule of the sample size and interval width, the bound on the gap (18) can be reduced, yielding a lower constant and replacing the  $\log(N)$  term with  $\log(N\Delta)$  to parallel the gap achievable when  $\Delta$  is known, equation (17). The optimal schedule grows the experimental group size slightly slower than  $4^m$ . For the full details of this schedule, the extension to  $k$  treatments, and a doubling scheme for the case when  $N$  is unknown, consult the original development of the algorithm by Auer and Ortner [2010].

Now, it's not always practical to run experiments this way, as multiple stages mean trials will take much longer. You have to stage experiments sequentially, but that's impractical when a single experiment takes multiple years. On the other hand, successive experimentation neuters the pedantic accusations that a study is underpowered.

## 2 Proof of Theorem 1.1

Like we did in the nonadaptive case, we have the following expression for the expected gap

$$\mathbb{E}[G] = \Pr[z_1 \neq z_*] \Delta. \tag{19}$$

Bounding this probability is a bit more tricky when we have a sequence of experiments. We take the following analysis strategy, adapted from Auer and Ortner [2010], bounding the probability by three cases.

To develop this bound, first define

$$m_0 = \min \left\{ m : 2^{-m} \leq \frac{\Delta}{2} \right\}. \tag{20}$$

$m_0$  would be the first round where the algorithm checks if the means are closer than  $\Delta/2$ . Since this is true, we have to have the inequality chain

$$2^{-m_0+1} \geq \frac{\Delta}{2} \geq 2^{-m_0} \tag{21}$$

Or, equivalently

$$2^{m_0} \leq \frac{4}{\Delta} \leq 2^{m_0+1}. \quad (22)$$

Since each experiment is a uniformly random sample from the remaining population, we can conceive of the samples being drawn in advance of all experimentation. That is, the population can be conceived of as a partition of sets defined by a random ordering

$$\Omega = S_0 \cup S_1, \cup, \dots, \cup S_\ell \quad (23)$$

where  $S_m$  would be the set of indices experimented on in round  $m$  of the algorithm. In this way, we can distinguish between cases in which unit 1 is assigned to  $1 - z_*$  and is in  $S_k$  for  $k < m_0$ , versus when it is in  $S_k$  for  $k > m_0$ .

We can now introduce the three key events whose probability we need to bound.  $\mathcal{E}_0$  is the event that the index 1 is assigned to  $S_k$  for some  $k \leq m_0$ .  $\mathcal{E}_1$  is the event that the algorithm terminates in a round with  $k < m_0$  and assigns the remaining units to  $1 - z_*$ .  $\mathcal{E}_2$  is the event that the algorithm makes it to round  $m_0$  but doesn't terminate in round  $m_0$ .  $\Pr[z_1 \neq z_*]$  is less than the sum of the probabilities of these three events:

$$\Pr[z_1 \neq z_*] = \Pr[(z_1 \neq z_*) \wedge \mathcal{E}_0] + \Pr[(z_1 \neq z_*) \wedge \neg \mathcal{E}_0] \quad (24)$$

$$\leq \Pr[\mathcal{E}_0] + \Pr[(z_1 \neq z_*) \wedge \neg \mathcal{E}_0] \quad (25)$$

$$\leq \Pr[\mathcal{E}_0] + \Pr[(\mathcal{E}_1 \vee \mathcal{E}_2) \wedge \neg \mathcal{E}_0] \quad (26)$$

$$\leq \Pr[\mathcal{E}_0] + \Pr[\mathcal{E}_1 \vee \mathcal{E}_2] \quad (27)$$

$$\leq \Pr[\mathcal{E}_0] + \Pr[\mathcal{E}_1] + \Pr[\mathcal{E}_2]. \quad (28)$$

The inequality (26) follows because for any unit not in the experimental groups of rounds 0 through  $m_0$ , they are assigned to  $1 - z_*$  if the algorithm incorrectly terminates before round  $m_0$  or they are assigned to  $1 - z_*$  after round  $m_0$ . The event  $\mathcal{E}_2$  contains these latter conditions.

Bounding  $\mathcal{E}_0$  just amounts to counting:

$$\Pr[\mathcal{E}_0] = \frac{n_0}{N} \sum_{k=0}^{m_0} 4^k = \frac{n_0 4^{m_0}}{N} \sum_{k=0}^{m_0} 4^{-k} \leq \frac{n_0 4^{m_0+1}}{3N}. \quad (29)$$

For  $\mathcal{E}_1$ , note that the probability that the algorithm does not terminate before round  $k$  and terminates in round  $k$  with the incorrect treatment is less than or equal to the probability that when we run an RCT on the set  $S_k$ ,  $\hat{\mu}_{z_*}^{(k)} \leq \hat{\mu}_{1-z_*}^{(k)} - 2^{-k}$ . Thus, by the union bound,

$$\Pr[\mathcal{E}_1] \leq \sum_{k=0}^{m_0} \Pr[\hat{\mu}_{1-z_*}^{(k)} - \hat{\mu}_{z_*}^{(k)} \geq 2^{-k}] \quad (30)$$

$$\leq \sum_{k=0}^{m_0} \exp\left(-\frac{n_0 4^k (2^{-k} + \Delta)^2}{8}\right) \quad (31)$$

$$\leq m_0 \exp\left(-\frac{n_0}{8}\right) \quad (32)$$

For  $\mathcal{E}_2$ , the calculation is simpler:

$$\Pr[\mathcal{E}_2] \leq \Pr[\hat{\mu}_{1-z_\star}^{(m_0)} - \hat{\mu}_{z_\star}^{(m_0)} \leq -2^{-m_0}] \quad (33)$$

$$\leq \exp\left(-\frac{n_0 4^{m_0} (\Delta - 2^{-m_0})^2}{8}\right) = \exp\left(-\frac{n_0}{8}\right) \quad (34)$$

Putting these bounds together gives

$$\Pr[z_1 \neq z_\star] \leq \frac{n_0 4^{m_0+1}}{3N} + (m_0 + 1) \exp\left(-\frac{n_0}{8}\right) \quad (35)$$

Choosing  $n_0 = 8 \log(N)$ , and applying the inequality (22) gives the bound

$$\mathbb{E}[G] \leq \frac{1}{N} \left( \frac{512 \log(N)}{3\Delta} + 3\Delta - \Delta \log_2(\Delta) \right) \leq \frac{174 \log(N)}{N\Delta}. \quad (36)$$

This completes the proof.

## References

Peter Auer and Ronald Ortner. Ucb revisited: Improved regret bounds for the stochastic multi-armed bandit problem. *Periodica Mathematica Hungarica*, 61(1-2):55–65, 2010. doi: <https://doi.org/10.1007/s10998-010-3055-6>.