

# Adaptive Sampling Designs

This explains some of the advantages of using adaptive sampling designs for more efficient statistical experiments, Our interest in them is for controlled clinical trials, though they are used for many other problems as well. This page also discusses our work in developing new algorithms to help create and analyze such experiments.

1. [What are Adaptive Sampling Designs?](#)
  2. [Why Aren't They Always Used?](#)
  3. [Our Attempts To Help Make Them More Useful](#)
  4. [Questions? Need Consultant Help?](#)
- 

## What are Adaptive Sampling Designs?

Adaptive sampling designs for statistical experiments, also known as response-adaptive designs, are ones where the accruing data (i.e., the observations) are used to adjust the experiment as it is being run. In a typical non-adaptive experiment, decisions such as how to sample during an experiment are made and fixed in advance. For example, in a classical clinical trial comparing two different treatments, patients are assigned to the two treatments with half being assigned to each therapy. At the end of the experiment an evaluation is made as to which treatment is more effective.

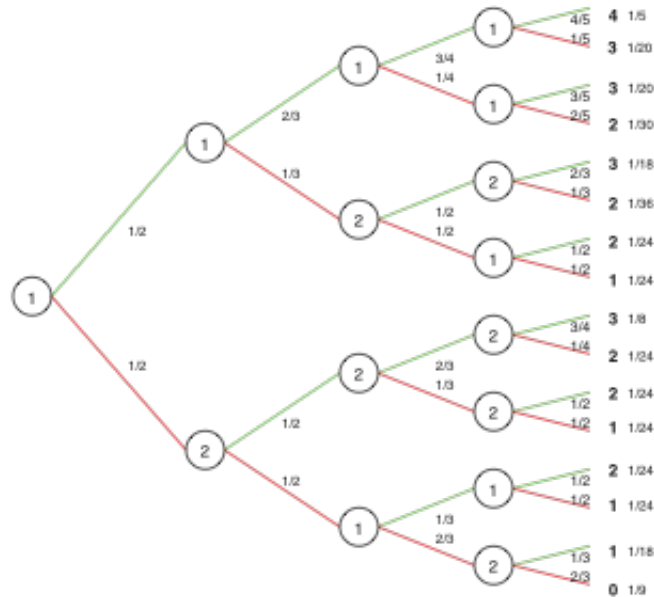
In contrast, in an adaptive clinical trial, patient outcomes can be used as they become available to adjust the assignment of future patients, assigning more of them to the better treatment. This allows one to improve expected patient outcomes during the experiment, while still being able to reach good statistical decisions in a timely fashion. Thus, adaptive procedures can offer significant ethical and cost advantages over standard fixed procedures. A situation in which adaptive procedures have become particularly relevant are AIDS clinical trials, since many new drugs come to market, yet classical randomized studies may pose ethical dilemmas and time delays. They are also quite important in pre-clinical animal trials, phase I trials, and phase I/II trials. While there had been some opposition to their use quite some time ago, now NIH, FDA, EPA, etc. are encouraging their use.

Adaptive designs also arise naturally in computer-related areas such as computer systems resource allocation or parameter selection for repeated computer simulations, since the computer can both collect data and run a program to decide what to do next. They can be extremely useful in industrial settings as well, such as acceptance testing. Generally speaking, adaptive sampling procedures are of interest to investigators using statistical experimental designs, to those concerned with medical ethics, to people working in stochastic optimization, control theory, decision theory, machine learning (where this subject is sometimes known as "active learning"), as well as to computer scientists.

## Bandits

Some aspects of this work fall within the area known as *bandit theory*, since one useful model is to think of the sampling options as the arms of slot machines ("bandits") with unknown payout

distributions. Many problems can be viewed as trying to maximize profits by observing the payouts (the outcome of a "pull" on the arm), and using this information to adjust the decision as to which arm to pull next. There may be a cost each time an arm is pulled (as occurs with the real 1-armed bandits we encounter in gambling establishments), or there may be a limit on the total number of pulls allowed. In some settings the goal may not be to maximize profits, but instead to estimate the difference in payout rates among the arms. Of course, we hope to have an expected gain from playing our experiments, not the expected loss that occurs in Las Vegas. Here is an example of a simple 2-armed bandit design, where the outcome of each arm is either a success or a failure. There will be exactly 4 pulls, and the goal is to maximize the expected number of successes.



The blue numbers indicate the arm pulled. The green lines slanted upward indicate a success, and black lines slanted downward indicate failure. The fractions adjacent to a line indicate the probability of obtaining that outcome. We assume all pulls are independent and each arm has some unknown probability of providing a success (i.e., it has some unknown bias). We then use a Bayesian approach where we assume that these probabilities are independent and uniformly distributed on  $[0, 1]$ . That is, we assume we know something in advance, such as assuming the arms aren't highly likely to be successful 99% of the time, but we don't want to make strong assumptions, so we assume that all success rates are equally likely (if you know anything about continuous probability distributions then mathematically this statement isn't correct, but this is an intuitive explanation, not a formal one!).

The experiment starts on the left with a pull on arm 1. (The choice of arm 1 was arbitrary - we could have just as easily started with arm 2). We initially assumed that arm 1 had  $1/2$  probability of success, and if we observe a success, since we assume that there is some underlying bias, now our estimate of success probability increases  $2/3$ . Similarly, had the first pull been a failure, then our estimate would drop to  $1/3$ . If we have a success then the diagram shows that we try arm 1 again. If we observe failure on the second pull, then we try arm 2, and so forth. The black numbers at the right indicate the number of successes observed if that is the path followed, and the red fractions at the right indicate the probability of that path occurring. This probability is the product of the probabilities along the edges of the path.

There is an important difference between this experiment, in which each arm has some unknown bias, and pulling on arms where it is known what the success rate of the arms are. Suppose, for example, that we know that the success rate for each arm is  $1/2$  (as opposed to its *expected value* being  $1/2$  as it is with the uniform prior). In this case, pulling the arms provides us with no new information, and thus, in 4 pulls the expected number of successes is 2, regardless of which arms are pulled. In our adaptive experiment, however, the outcomes teach us something about the different arms, and by adapting to use the more successful arm, we increase the expected number of successes to more than 2.27, even though initially we only assumed that each arm was equally likely to produce a success or a failure. The more pulls we use, or the more arms there are to choose from, the more the adaptation can help.

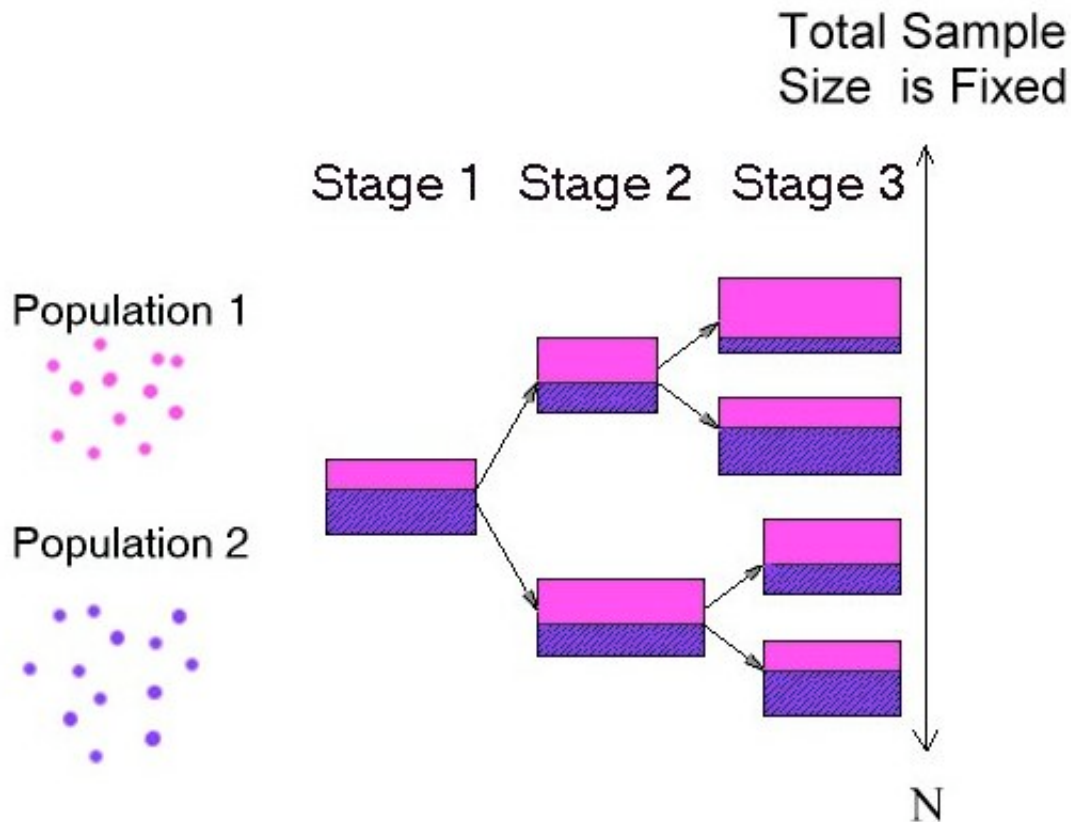
## Generalized Bandits

The bandit model is extremely general. For example, suppose you have several restaurants in your area, and you want to eat at the best. You can keep going to the same restaurant, or try different ones in the hope that they might turn out to be even better. Going to the same restaurant is sometimes called "exploitation", while trying new ones is "exploration". In clinical trials these concepts are sometimes called "learn and confirm", with the learning coming in early stages of the clinical trial process (such as stage II), and the confirming coming in subsequent stages (such as stage III). Optimal solutions to such bandit problems make tradeoffs between these aspects. Note that if you are soon going to leave the area then you'll emphasize the exploitation part, while if you are new to the area then you'll explore.

The restaurant example provides another possible goal and set of actions: rather than wanting to just identify the best and keep going there, you might instead want to keep a small group that you go to. This is related to what is known as "multiple plays", and is useful because it allows one to incorporate other considerations such as variety, or ability to go to a different restaurant if the one you want is full. This is an especially useful strategy if you live near the University of Michigan and you are trying to find a restaurant on a weekend when there is a home football game.

## Staged Allocation

Sometimes you want to finish the learning more quickly, perhaps because each outcome requires quite a while to observe or perhaps it is expensive to set up. In this case people often perform sampling in stages, making many observations at the same time. Typically this was done by evening dividing the patients among the arms at each stage. Further, just as before, one can do better by adapting as the observations are obtained, adjusting the size of the next stage and the number of observations on each arm. Here is an example, where the total number of observations is fixed. The width of each box indicates the number of observations, and the colors indicate the relative proportions of the two options.



Staged allocation is much more difficult to optimize, and we designed the only algorithm capable of optimizing it in a feasible amount of computing time.

### Some Other Forms of Adaptive Sampling

One form that is quite important is when you are trying to find a new streak of gold. You may take rock samples at many different places, and then when you've found some that look promising you'll go back and take more detailed samples in the area. Then you might select a single spot where you'll spend a great deal of time trying to locate the streak.

In this example, geography and location in space are important. You're determining your second round, and perhaps third round, of detailed sampling based on the results of sampling on a much coarser grid. You are assuming that nearby areas are correlated, so that if there is a gold streak then some of the rocks nearby will have at least a little gold in them. It's not exactly arms as in the bandit model, but you'll still balance exploration and exploitation aspects if you want to do well. You should also be prepared for a negative expected outcome.

Another example is when you are trying to estimate the difference between the average income of workers in fast food restaurants versus the average income of waiters in regular restaurants. You survey workers in each population and quickly discover that the fast food workers all get quite similar, low, pay, while the waiters have widely varying income, based on factors such as the quality of the restaurant, the skill of the waiter, etc. Thus if you want to accurately estimate the difference in

the average pay, you'll need to sample many waiters to improve your estimate of their average. If you'd known in advance the variance of the waiter's and fast food worker's salaries, then you could have figured out in advance how many to sample from each group, but since you don't know this you first do some sampling to estimate it, and then refine the number of samples you'll take from each group.

## Why Aren't Adaptive Procedures Always Used?

Unfortunately, adaptive procedures are more complicated to design and to analyze, and in some settings are more difficult to implement. Most people using formal experimental designs use fixed sampling procedures where they decide in advance how many samples of each alternative they are going to collect. These fixed procedures are easier to design and analyze, because they don't have as many alternatives, but they are less efficient, sometimes significantly so. In many cases, however, users are unaware of the option to use adaptive designs because standard statistical courses and packages do not include them. In medical problems, people may be wary of adaptive designs because of the confusion between a controlled clinical trial using an adaptive design versus allowing the physicians to decide each patient's treatment based on professional experience. The latter can introduce many biases into the learning process, especially because physicians tend to use the exploitation approach for each patient, as opposed to including some exploration.

There are some complicated ethical issues here - which is better the better option?

- For each patient, do what you currently think is best.
- Occasionally run some clinical trials which might find that something new is better, helping future patients, though with the chance that the new option isn't as good as the old one and some of the patients in the trial will get inferior treatment.

One of the important aspects of adaptive trials is that they can decide which option is best, and assign relatively more patients to it, much faster than a fixed trial. Note how similar this is to the bandit problem.

## Our Research to Help Make Them Useful

We have developed a variety of new approaches and computational algorithms that make it much easier to exploit such designs. Here is a [more detailed explanation](#).

## Questions? Consultant Help?

If you have questions or would like consultation help in this area, please contact us:

Janis Hardwick

jphard @ umich · edu     [www.eecs.umich.edu/~jphard/](http://www.eecs.umich.edu/~jphard/)

Quentin F. Stout

qstout @ umich · edu     [www.eecs.umich.edu/~qstout/](http://www.eecs.umich.edu/~qstout/)



Copyright © 1997-2017 Quentin F. Stout.