

Understanding Simulation Error

Teddy Weaver

Contents

It can be hard enough to predict outcomes of a single event, like rolling a pair of dice or hiring process for a job applicant. Now what about a series of events, like 100 or even 1,000 job applicants? Understanding series of related events, like these two examples is a **stochastic process**, where we are able to build models in an attempt to predict future outcomes.

The models for each process will vary significantly. For rolling two dice, the model is probably quite simple, whereas job applicants could have a wide range of factors (location, position, interviewers, years of experience, etc.) that need to be taken into consideration. Applications of stochastic models are prevalent in a number of industries including finance, computer science, health care, and many others.

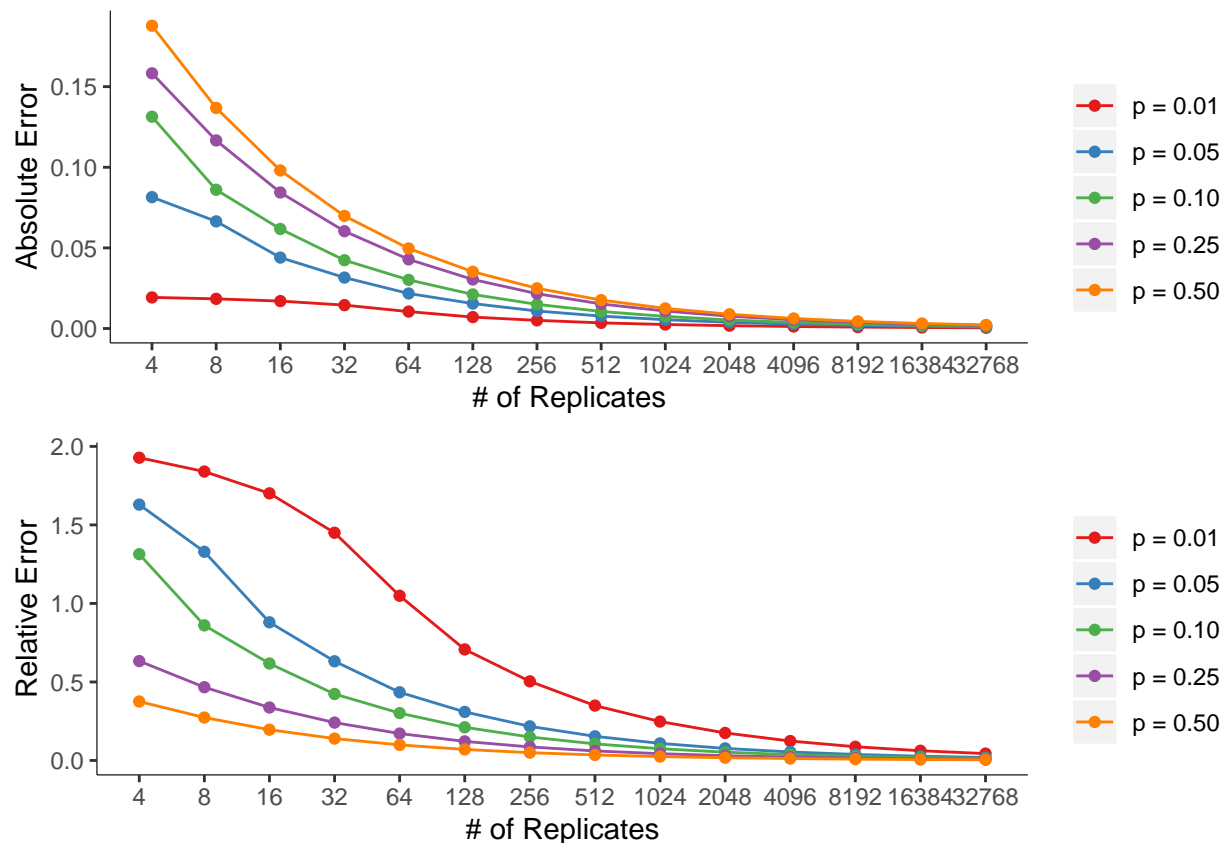
Monte Carlo method is a common simulation technique to evaluate the uncertainty and accuracy of stochastic models. It involves repeatedly running the simulation of the model to minimize error.

As another example, let's think about flipping a coin - another stochastic process. How many times have you had an even number of heads and tails? Probably not very often! Let's say we flipped 10 times with 6 heads and 4 tails - a pretty even result, but we would have expected 5 heads and 5 tails. Our **absolute error** in this case is 1 (6-5), and our **relative error** is 10% (1/10 flips).

One way to reduce this error is to repeat the Monte Carlo simulation many, many more times.

But how many times? What a great question!

To better visualize this, below are two graphs absolute and relative error. Each line represents a different probability value (p) and the y-axis is the number of times we repeated each simulation.



What does this mean?

- In both we see the greater the probability, the fewer simulations you need to run with diminishing returns around 2048 replicates.
- **Absolute error** has relatively uniform behavior for each probability, converging around 4,096 replicates.
- **Relative Error** the lower the probability the more replicates are needed to reduce the observed error.

As with the expected value of the Martingale strategy from the last post, it is important that the interpretation of error, both absolute and relative, remain contextual. For example, if we had \$1M to blow at the roulette table, we wouldn't really care about the absolute error ($p = .5$). Our expected return would be -\$50 but if we're off by 50%, we're not really hurt by the additional \$25 lost because it's such a small percentage of our initial amount (.0025%). If we only had \$100, this would be a very different story.