Practical Hypothesis Testing to Guide Decision Making in a Complex Adaptive System

By Tom Coyne

One of the most frequently heard complaints among corporate executives is the apparent practical irrelevance of much academic research on business and investment issues. In recent years, I have come to realize that one of the root causes of this gap are the different statistical methods and standards of proof that academics and business leaders use in their day-to-day work.

Whether or not we do it consciously, our problem solving process usually proceeds through up to seven stages:

1. We allocate and direct our scarce attention, and observe a problem – i.e., an actual or potential situation wherein results would fall short of our goals.
2. We explain why the problem exists, by recalling or constructing a causal model of the situation.
3. On the basis of our explanation, we generate options/plans for improving the situation
4. We use our causal model to predict the likely impact of each option, or combination thereof.
5. We decide which option or options to initially pursue.
6. We implement our plan.
7. We observe the results, compare them to our expectations, and, if necessary, adapt our causal model and/or plan.

Perhaps the most challenging step in this process is the second one – causal reasoning. I believe that an important source of difficulty we experience with causal reasoning in business lies in the way we usually learn about this critical process in school. For most of us, our first formal experience with causal reasoning is in our earliest science classes, and our exposure to the scientific method. More specifically, we first learn about causation in the context of fixed, deterministic systems, which are usually characterized by relatively simple cause and effect relationships that are consistent and unchanging over time. This allows us to test hypotheses using experiments that are both carefully controlled and easily repeated.

Unfortunately, social systems, like organizations, industries, markets, societies, economies, and governments aren't the deterministic systems we first encountered in our early science classes. Rather, they are <u>complex</u> systems, in which effects have multiple interacting causes, which are often characterized by time delays and non-linearity. Moreover, they are also <u>adaptive</u> systems, populated by intelligent actors that modify their behavior over time in light of the effects it produces. One signature characteristic of complex adaptive systems (CAS) is that their effects are impossible to predict to anything close to the same degree of accuracy that is possible in deterministic systems (just to clarify, some physical systems like weather are complex, but not adaptive; even with the best models and most powerful computers now in use, the accuracy of weather forecasts currently degrades sharply after seven days). Another CAS characteristic is that many of the results they produce tend to have an exponential (i.e., power law) rather than Gaussian (i.e., normal or bell curve) distribution.

Having clarified the nature of the complex adaptive systems within which business leaders must make decisions, let us move on to a short discussion of the statistical tools that are typically used to assess causal hypotheses about the way they operate.

Traditional (frequentist) statistical hypothesis tests are based on the (usually implicit) assumption that the underlying system/data generating process doesn't change, and our task is to discover the true nature of its operation. Given this assumption, the traditional approach to testing a hypothesis about the system asks how likely it is that a given result could arise solely by chance. For example, let's assume I know the distribution of product quality when I use a certain mix of chemicals in the process used to produce it. My hypothesis is that a different mix will produce an improved result. After implementing the new mix, I subsequently observe a five percent increase in product quality. Traditional statistical methods ask the question, "how likely is it that this result would have occurred by chance?" The "p-values" one typically encounters in traditional hypothesis tests measures this probability, with lower values signifying a lower probability that the observed result would have occurred by chance. If the p-value is less than a set threshold value, the hypothesis is accepted as true. To further reinforce our confidence in the truth of the hypothesis, we can repeat the experiment (perhaps at a different plant) to see if we obtain a similar improvement in performance and associated p-value.

Now consider the challenge of applying this approach to causal reasoning and hypothesis testing in a situation where the underlying data generating process is constantly evolving. As you would expect, this situation creates a very substantial challenge for the application of traditional hypothesis testing procedures. To their credit, academic researchers have sometimes found ways to meet this challenge, through creative experiment design and/or the use of complex statistical methods. Unfortunately, in doing so they have too often made their results more difficult for business leaders to understand and apply. The result has been the aforementioned gulf between academic business research and the needs of real world business leaders.

In my experience, Bayesian statistics offer hope for bridging this gulf.

While traditional frequentist statistics seek to infer the probability of observing a series of data assuming that a hypothesis about the unchanging data generating process is true, the Bayesian approach attempts to infers the probability that a hypothesis is true, given a set of newly observed data, <u>combined with</u> the decision maker's prior belief in the truth of the hypothesis (which can be based on experience, intuition, and/or previous analysis results). Put differently, the frequentist approach focuses on the properties of a sequence of data, while the Bayesian approach focuses on the properties of a decision maker's beliefs about the current (not unchanging) state of the underlying data generating process. To simplify, think of it as the difference between trying to hit a stationary versus a moving target. While these differences between traditional and Bayesian statistics may seem like arcane distinctions, they represent very different underlyng philosophies of causal reasoning.

For the Bayesian decision maker, the critical question is "what is the probability I would observe this new evidence if my hypothesis is true, compared to the probability I would observe it if the hypothesis is false?" The greater this "likelihood ratio", the greater the increase in the decision maker's prior probability/belief that the hypothesis about the current state of a complex adaptive system is true. This Bayesian approach also helps a decision maker avoid information overload, by focusing his or her information search on those pieces of evidence with the highest likelihood ratios. Even if they've never taken a stats course, most organization leaders, who have to make daily decisions in the face of irreducible uncertainty about the true nature of a complex adaptive system, are practical Bayesians, whether they know it or not. For more on Bayesian reasoning, see The Theory That Would Not Die by Sharon McGrayne, which is an excellent and highly readable introduction to this subject.

Tom Coyne consults on enterprise and investment risk management. He began his career as a credit officer at Chase Manhattan Bank. He subsequently spent almost twenty years as a management consultant, specializing in organizational performance improvement, turnarounds and growth. He has also served as the CFO, CRO, and CEO of public and private companies. Tom is also a frequent contributor to investment research publications, and a member of the top ranked team in the Intelligence Advanced Research Projects Agency's multiyear forecasting tournament. Tom can be reached at tcoyne@sachuestadvisors.com.