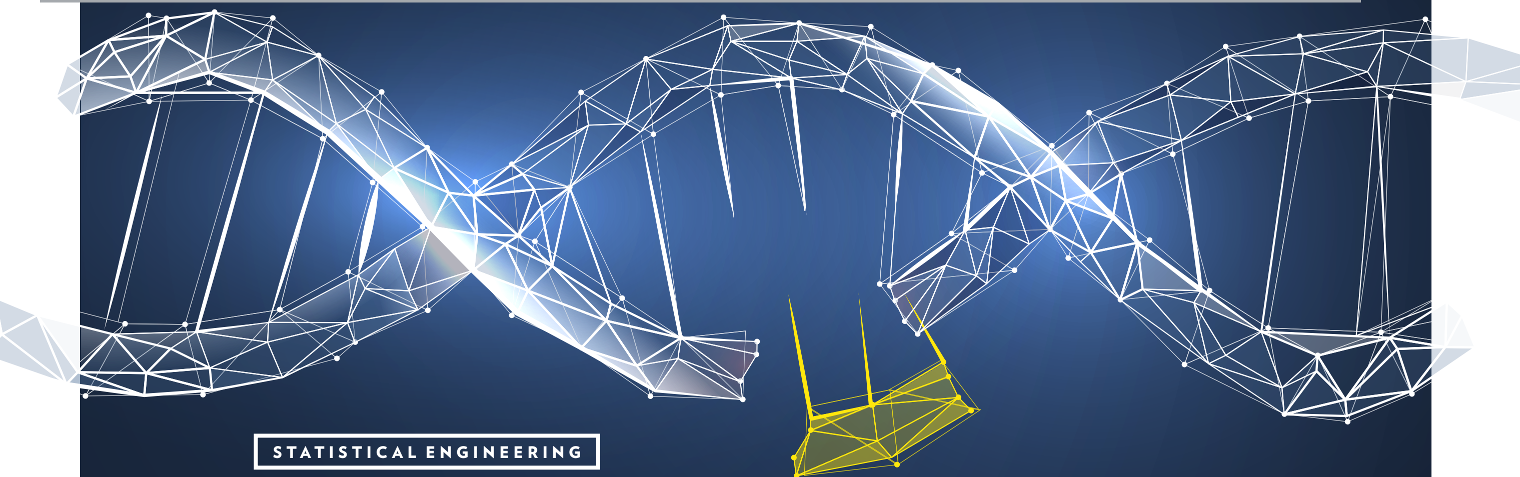
STATISTICS SPOTLIGHT

Solving quality quandaries through statistics



STATISTICAL ENGINEERING

New Day Dawning

A fresh discipline mutates from a widening statistical gene pool by Lynne B. Hare

Over time, biologists say, some species spread widely around the globe. As they spread, they adapt to new surroundings. Their gene chain lengthens to the extent that opposite ends might be unrecognizable to each other. They may not even be able to mate with each other. That's how new species are formed.

So, at some early point in time, was there only one pair of birds? Did their offspring and theirs and so on through millions of years become strangers? Not a clue. It must be more complicated than that, but basically it seems a logical conclusion, consistent with theories of evolution. We should probably believe it, if only in broad general terms.

There are parallels in our present-day behavior. In its early days, statistics evolved from curiosity bolstered by the scientific method. It was not a discipline unto itself. Chemists, biologists, physicists and other scientists contributed to its development, inspired by their immediate applications. Mathematicians added the formal touches. But in a process not unlike the evolutionary one, the statistical discipline was formed, as were academic departments for its care and feeding. Proponents still recognized and communicated with each other. (We will leave mating prospects out of the present discussion.) Suffice it to say that statistics as a discipline became a reality.

Major issues

To get a degree in statistics, a student must be proficient in math. That's because statistics' underpinnings are decipherable only to those knowledgeable of calculus and probability theory. Without sufficient knowledge of theory, practitioners misapply, causing errors. So, the study of statistical theory is a just and right thing to do; there can be no doubt. Without it, the field cannot grow, and without growth, the field will atrophy.

Singular concentration on theory, however, can be the road to perdition, especially if a pure theoretician ventures into domains of statistical application. This is because appropriate use of statistical methods depends highly on the field of application and skills beyond theory. Pure statistical theoreticians don't typically spend much time on applications—not that it's beneath them. Their time is better spent on theory development.

At one university with a statistics department staffed with theoretical and applied statisticians, some of the applied statisticians had moved on or were otherwise unavailable. There was high demand from client departments

iStock.com/Yevhen Lahunov

FIGURE 1

The statistical engineering approach to opportunity identification and problem solving

for scientific collaboration, so the statistics department chair (a theoretician) assigned some other theoreticians to stand in. Mayhem is too weak a descriptor to explain followed.

This is a two-way street, of course. To ask an applied statistician to teach a course on advanced theory is to invite disaster, too.

> *A musical gene chain: Most likely the first person to stretch strings across a hollow log used the result to make music. Now, violin makers are not usually concert violinists, and certainly not without practice. Not many concert violinists dabble in woodwork.*

But the bidirectional path is manifest in other forms. Liberal arts colleges encourage students to broaden their horizons by crossing disciplines beyond their majors. The best among them knit disciplines together by encouraging enrollment in pan-disciplinary courses.[1]

Still, the declaration of majors forces the gene chain to lengthen, and the length increases as students move into graduate school or other forms of specialization. To be sure, many students of science and engineering are encouraged (or perhaps forced) to study some "statistics for research."

These courses, however well executed, may be considered shots to the foot unless they put statistics and statistical methods in the framework of opportunity revealing, problem solving strategies. Statistical internships must be encouraged and toy problems discouraged with a process view taking their place. Failing that, statistics becomes the enemy and "my worst course in college." That's because they usually lack relevance, and any real value they might have had fades before the arrival of opportunities for their use in decision making. What a waste!

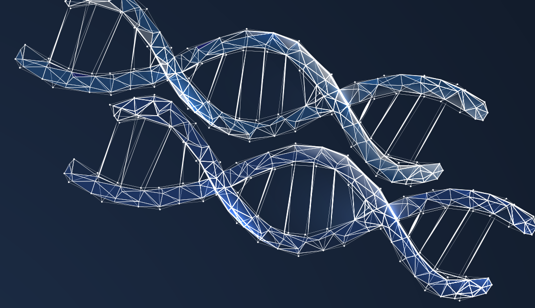
More than a new name

The redemptive key is a framework for the realization of opportunities through the application of problem solving strategies. Where might we find that? At the far, applied end of the statistical gene chain lays statistical engineering, a relatively new manifestation of genetic drift. Make no mistake: Statistical engineering is not simply a renaming of applied statistics. It is much larger than that. We might even say a mutation because it sets the stage, scenery, lighting and all for real multidisciplinary breakthrough.

Formally, statistical engineering is "the study of the systematic integration of statistical concepts, methods, and tools, often with other relevant disciplines, to solve important problems sustainably."[2] Much is crammed into that definition.

If you've spent any time attempting to solve complex problems, you understand the overwhelming temptation to get in there and fix things. Apply your own brand of wizardry, don't ask questions, take no prisoners and "get 'er done." Down deep, you also know that approach almost never works.

That's because the real solution usually requires a multidisciplinary view. It starts with identifying the opportunity. Do that formally with all hands on deck so there can be no variability of understanding. (After all, managing variability is a very statistical thing to do.) The identification should cross disciplines and organizational departments, and it is the real cornerstone of success.[3]



Next, structure the problem by determining associated processes, subprocesses and their related metrics. If we all agree that line seven efficiency is brutally low, that's not enough. We must know how that efficiency is measured, what feeds the line, how the feed's quantity and quality are measured, and who controls it and how. Now, we must know how accurate and precise these measurement processes are. Not done yet, we must know similar features of each unit operation down the line from source to sink, all the way out to the eventual consumer (Figure 1).

It is good politic to know the organizational context of the problem. Who wants it solved? Who might be embarrassed if it is solved? What was attempted and had failed? Those who find workable solutions are not always immune from attack.

It is good politic to know the organizational context of the problem. Who wants it solved? Who might be embarrassed if it is solved? What was attempted and had failed?

At one organization, a problem had persisted for at least seven years. A new researcher collaborated with R&D teammates and a statistician, as encouraged by the director. A designed experiment was applied, data analyzed and interpreted, and a solution emerged within a few weeks. Some outside R&D sniped, saying the solution would cost the company money. A cost-benefit analysis later showed great gains. Still, internal jealousies could have been avoided had the full problem context been understood. The sociological context is important for organizational stability and just plain getting along.

Following the understanding of context comes strategy development. This is absolutely essential—creating a cohesive plan developed and embraced by all team members. The strategy defines steps of the broad approach to be taken toward solution. Emerging from the strategy are individual tactics to be shouldered by individuals or subgroups of the overall team. These may include the use of designed experiments, analyses of retrospective data, exploration of tangential issues and others.

Finding a sustainable solution

Care must be taken during this phase to avoid nonconforming splinter strategies running counter to the main overall strategy. They can cause great damage as can be seen, for example, when individual opposing members of sports teams begin to fight. Penalty flags fly and whistles blow because temper-induced strategies interfere with the overall strategy (hockey might be an exception). We don't usually think of it in those terms, but that is actually what's going on.

All the efforts must culminate in a sustainable solution. The work is not complete without an instantiation of means to ensure permanency. This necessitates installation of monitoring and controlling safeguards to prevent process drift.

Although there are some lights on the horizon, little if any of this is taught in academia, where data are too often accepted at face value, and real, complex problems are simplified into well-defined toy problems. Statistics graduates are not yet prepared to take this subject on. It will take the establishment of new course—and perhaps entirely new departments—to teach statistical engineering.

Make no mistake: Statistical engineering is not your parents' applied statistics. Its new day is dawning. **QP**

REFERENCES AND NOTE

1. Colorado College, "Catalog of Course—General Studies" GS101 "Freedom and Authority," https://tinyurl.com/2pfjk7fa.
2. International Statistical Engineering Association, https://isea-change.org.
3. This and what follows borrows heavily from the first chapter of the body of knowledge in reference 1. Its editor is Roger W. Hoerl.

Lynne B. Hare is a statistical consultant. He holds a doctorate in statistics from Rutgers University in New Brunswick, NJ. He is past chair of the ASQ Statistics Division and a fellow of ASQ and the American Statistical Association.