MINI PAPER

STATISTICS AND REALITY – PART I

Davis Balestracci
Harmony Consulting
(651)-269-0954
davis@dbharmony.com
www.dbharmony.com

Introduction

Required "Statistics from Hell 101" courses are virtually *worthless*…and perpetuate myths that statistics can be used to "massage" data and prove anything. Common uses on organizational data—variance tables, trends analysis, rankings, stretch goals, tougher standards—can actually *sabotage* improvement!

Whether or not you understand statistics, you are already using statistics! The key skill needed is the ability to respond to variation appropriately so as to *ask better questions* – mathematical skills aren't necessarily the focus!

Data "Sanity": Statistical Thinking Applied to Everyday Data

People generally do not perceive that they need statistics—the need is to solve their problems.

Given the current rapid pace of change in the economic environment along with the "benchmarking," "re-engineering," "total customer satisfaction," and, most recently, "Six Sigma" crazes, there seems to be a new tendency for performance goals to be imposed from external sources, making improvement efforts flounder when:

- Results are presented in aggregated row and column formats complete with variances and rankings,
- Perceived trends are acted upon to reward and punish,
- Labels such as "above average" and "below average" get attached to individuals/institutions,
- People are "outraged" by certain results and impose even "tougher" standards.

These are very well meaning strategies that are simple, obvious...and wrong! They will mislead analysis and interpretation…and insidiously cloud decisions everyday in virtually every work environment.

Realities:

- Taking action to improve a situation is tantamount to using statistics,
- "Traditional" statistics have severely limited value in real world settings,
- Understanding of variation is more important than using techniques,
- Statistical *thinking* gives a knowledge base from which *to ask the right questions*,
- Unforeseen problems are caused by the exclusive use of arbitrary numerical goals, "stretch" goals, and "tougher" standards for driving improvement,
- Using heavily aggregated tables of numbers, variances from budgets, or bar graph formats as vehicles for taking meaningful management action are many times futile and inappropriate,
- There is poor awareness of the true meaning of "trends," "above average," and "below average."

Key Concept: Process - Oriented Thinking

- The statistics needed for quality improvement are based in the context of **process**. What is a process? *All work* is a process! Processes are sequences of tasks aimed at accomplishing a particular outcome by processing inputs to produce some type of output. Everyone involved in the process has a role of supplier, processor or customer. A group of related processes is called a system.

Process-oriented thinking is built on the premises of:

- Understanding that *all* work is accomplished through a series of one or more processes, each of which is potentially measurable,
- Using data collection and analysis to establish consistent, predictable work processes yielding desired outcomes, and these outputs go to both internal and external customers,

STATISTICS AND REALITY – PART I

- Reducing *inappropriate* and *unintended* variation by eliminating work procedures that do not add value (i.e., only add cost with no payback to customers),
- Process inputs falling into the six general categories of "people," "methods," "machines," "materials," "measurements (data)," and "environment," each of which is a potential source of variation,
- Improving quality = improving processes,

Aside from the new perspective of looking at your jobs and workplaces as processes and systems, process-oriented thinking must also be applied to a quality professional's *data collection process.*

How a Quality Professional "Adds Value": Recognizing the Use of Data as a Process

People, Methods, Machines, Materials, Environment and Measurements inputs can be a source of variation for any one of the measurement, collection, analysis or interpretation processes!

The use of data is really made up of four processes – Measurement, Collection, Analysis, and Interpretation – each having "people, methods, machines, materials, measurements (data inputs via raw numbers), and environment" as inputs. (See Figure below) *Any one of these six inputs can be a source of variation for* **any** *one of these four processes* – They lurk to contaminate your data process and mislead you as to what is going on in the actual process you are trying to improve!

People, Methods, Machines, Materials, Environment and Measurements inputs can be a source of variation for any one of the measurement, collection, analysis or interpretation processes!
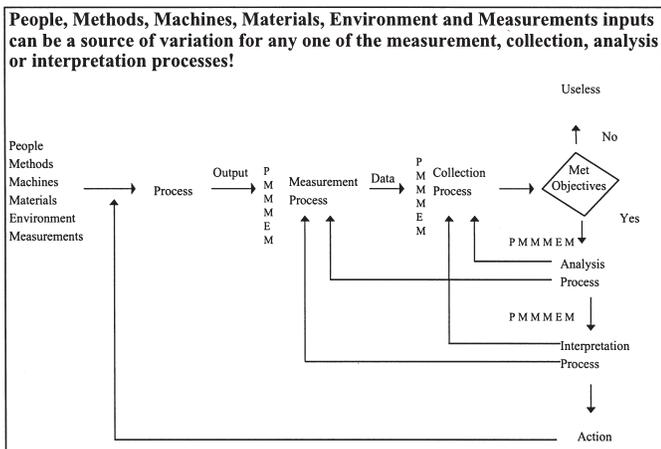
So, any process produces outputs that are potentially measurable. If one chooses, one can obtain a number (a piece of data) characterizing the situation via a **measurement** process. If the objectives are not understood or people have varying perceptions of what is being measured, the six sources of variation will compromise the quality of this measurement process.

These individual measurements must then be accumulated into some type of data set, so they next pass to a **collection** process. If the objectives are clear, the designed collection process should be relatively well defined – the appropriateness of an analysis depends on how the data were collected. If the objectives are not clear, the six sources of variation will once again act to compromise the process (Actually, from the author's experience, it is virtually guaranteed that the six sources will compromise the collection process anyway!).

If the objectives are passive and reactive, eventually someone will extract the data and use a computer to "get the stats." This, of course, is an **analysis** process (albeit not necessarily a good one) that also has the six sources of inputs as potential sources of variation. Or, maybe more commonly, someone extracts the data and hands out tables of raw data and cursory, computer-generated summary analyses at a meeting that becomes the analysis process, which is affected by the variation in perceptions and abilities of people at the meeting.

Ultimately, as you now are starting to realize, it all boils down to interpreting the variation in the measurements. So, the **interpretation** process (with the same six sources of inputs) results in an *action* that is then fed back in to the original process.

Now, think back for a minute to the many meetings you attend. How do unclear objectives, inappropriate or misunderstood data definitions, unclear or inappropriate data collections, passive statistical "analyses," and shoot-from-the-hip interpretations of variation influence the agendas and action? In fact, *how many times are people merely reacting to the variation in these elements of the DATA process* – and making decisions that have NOTHING to do with the process being studied?

STATISTICS AND REALITY – PART I

Another danger inherent in this data process is: *Data not collected **specifically** for the current objective can generally be "tortured" to "confess" to someone's hidden agenda!*

Research vs. Improvement

So, if the data process itself is flawed, many hours are spent "spinning wheels" due to the contamination from the "human" variation factors inherent in the aforementioned processes—People make decisions and react to their perceptions of the *data* process and *not* the process allegedly being improved!

Many physicians will argue that the process-oriented approach is invalid because it doesn't follow "established" procedures of clinical research. However, let's look at *research as a process.*

In research, all input variations are *tightly controlled* such that observed differences in the "control" and "treatment" groups can be attributed to the "methods" input and no other. The protocol is also excruciatingly detailed as to how measurements are defined, collected, and analyzed so as to reduce variation in the data process. Controlling variation is expensive, which is why research is expensive.

Hence, research statistics is actually a very specialized subset of process-oriented statistics! It makes the assumption and has the luxury of "ignoring" the everyday factors lurking to compromise results in busy uncontrolled practice environments.

However, after a significant result is published, do the researchers have any control over how clinicians use it? The "rigor" is gone, and human variation in interpretation and use of the protocol virtually guarantees that they won't achieve the same results as reported. This variation, usually either inappropriate or unintended, could be present in any or all of the six inputs…of five processes (actual & data measurement, collection, analysis, and interpretation)!

So, in understanding any variation between the research results and actual results, it becomes necessary *to expose the variation* between individual use and the research use of the protocol and reduce any inappropriate and unintended variation. As will be discussed in the next sections, traditional statistical methods are, for the most part, invalidated.

Do not underestimate the factors lurking in the data process that will contaminate and invalidate statistical analyses. *Objectives* are crucial for properly *defining* a situation and determining how to *collect* the data for the *appropriate analysis*. Statistical theory can interpret the *variation* exposed by the analysis to take *appropriate action.*

Further Clarification: "Enumerative" Statistics vs. "Analytic" Statistics

(I am indebted to the work of David Kerridge for much of the following explanation as well as those in the subsequent sections **Prediction** and **Unknown & Unknowable**)

There are actually three kinds of statistics, and they can be summarized as follows:

Descriptive – "What can I say about this *individual* patient?"

Enumerative – "What can I say about this *specific group* of patients?"

Analytic – "What can I say about the *process* that produced the result in this group of patients?"

An *enumerative study* always focuses on the *actual state* of something *at one point in the past* – no more, no less. For example, one can literally summarize the results of all the participants in any clinical trial once it is completed. Now what?

An *analytic study* usually focuses on *predicting the results* of action *in the future* – in circumstances we cannot fully know. *It is this predictive way of thinking that is fundamental to quality improvement.*

Both kinds of statistics count or measure samples. However, suppose it is desired to know which of two antibiotics is better in treating a certain disease? It is impossible to take a random sample of *all the people who will be treated in the future:* It isn't even known who specifically will get this disease in the future!

This can be described as sampling from an **imaginary** population. The practical difference is that we must not rely on what happens *from the results of any one experiment:* We must repeat the experiment under as many different circumstances as we can to

STATISTICS AND REALITY – PART I

establish an increasing degree of belief in the result. This is in very strong contrast with what is normally taught in most statistics textbooks, where it describes the problem as one of "accepting" or "rejecting" hypotheses.

The famous Walter Shewhart stated the difference by means of an example:

> **"You go to your tailor for a suit of clothes and the first thing that he does is to make some measurements; you go to your physician because you are ill and the first thing he does is to make some measurements. The objects of making measurements in these two cases are different. They typify the two general objects of making measurements. They are:**
> **(a) To obtain quantitative information**
> **(b) To obtain a causal explanation of observed phenomena"**

Prediction

The distinction between enumerative and analytic studies means we must look for repeatability over many different populations consistently over time. Analytic thinking relates to sampling from a *process*, rather than a well-defined, finite *population*. Furthermore, most mathematical statisticians state statistical problems in terms of repeated sampling from the *same* population under circumstances where nothing changes over time! This leads to a very simple mathematical theory, but does not relate to the real needs of the statistical user. Especially in medicine, one cannot take repeated samples from the exactly the same population, except in the rare cases.

Getting back to comparing two antibiotics in the treatment of some infection, suppose a conclusion is made that one did better in tests. How does that help?

Suppose that all testing was done in one hospital in New York in 1997; however, someone may want to use the antibiotic in Africa in 2002. It is quite possible that the best antibiotic in New York is not the same as the best in a refugee camp in Zaire. In New York the strains of bacteria may be different: and the problems of transport and storage really are different. If the antibiotic is freshly made and stored in efficient refrigerators, it may be excellent. It may not work at all if transported to a camp with poor storage facilities.

And even if the same antibiotic works in both places, how long will it go on working? This will depend on how carefully it is used and how quickly resistant strains of bacteria build up. The effectiveness of a drug may also depend on the age of the patient, or previous treatment, or the stage of the disease. Ideally it is desirable to have one treatment that works well in all foreseeable circumstances, but this may not be possible.

Unknown & Unknowable

There are usually no difficulties with carrying out the objectives of an enumerative study, which usually involves estimation, other than how to choose the sample; however, in process improvement, an analytic process, a number of subsequent practical problems still remain.

Random sampling is often used in analytic studies, but this is not the same as sampling in an enumerative study. For example, consider a chosen group of patients who attend a particular clinic and suffer from hypertension. Either a random method or some complicated method involving random numbers is used to determine who is to get which treatment. But the resulting sample is not necessarily a random sample of the patients *who will be treated in the future at that same clinic*.

Still less are they a random sample of the patients who will be treated in any other clinic. In fact the patients who will be treated in the future will depend on choices that you and others have not yet made! And those choices will depend on the results of the study currently being done and on studies by other people that may be carried out in the future.

So with an analytic study, there are two distinct sources of uncertainty. The first is similar to an enumerative study, that due to sampling.

The second is due to the fact that one is predicting what will happen at some time in the future – to some group that is different from the original sample. This uncertainty is "*unknown and unknowable.*" It is rarely

STATISTICS AND REALITY – PART I

known how any produced results will be used, so all one can do is to warn the potential user of the range of uncertainties that might affect different actions.

This is rarely done. Furthermore, how does one even express it? But the uncertainties of this kind will in most circumstances be an order of magnitude greater than the uncertainty due merely to sampling – making it very dangerous to pretend to be more certain than warranted. Such false certainty many times leads to wrong choices, but the result, in most statistics courses, has been a theory in which the unmeasured uncertainty has just been ignored.

"Statistics" in a Quality Improvement Perspective

So, it can be seen that statistics is not merely the science of analyzing data, but the **art** and science of **collecting** and analyzing data. Given any improvement situation (including daily work), one must be able to:

1) Choose and define the problem in a **process/systems** context,
2) Design and manage a **series of** simple, efficient data collections to expose inappropriate and unintended variation,
3) Use comprehensible methods presentable and understandable across **all** layers of the organization, virtually all **graphical** and NO raw data or bar graphs (with the specific exception of a Pareto analysis), and
4) Numerically **assess** the current state of an undesirable situation, **assess** the effects of interventions, and **hold the gains** of any improvements made.

Summary

As quality professionals, it is important to realize that data analysis goes far beyond the routine statistical "crunching" of numbers. The greatest contribution to an organization is getting people to understand and use a *process-oriented* context in analyzing situations as well as principles of good, *simple, efficient* data collection, analysis, and display. This cannot help but enhance the health care quality professionals' credibility. It will also help gain the confidence and

cooperation of organizational culture during stressful transitions and investigations.

It will be vital to put a stop to many of the current well meaning but ultimately damaging ad hoc uses of statistics – *Whether or not people understand statistics, they are already using statistics*...and with the *best* of intentions.

Useful References

Balestracci D. "Data 'Sanity': Statistical Thinking Applied to Everyday Data," Special Publication of the American Society for Quality Statistics Division, Summer 1998. Available through American Society for Quality Publications, Item S0808.

Balestracci D & Barlow J. *Quality Improvement: Practical Applications for Medical Group Practice, 2nd Edition.* Englewood, CO. Center for Research in Ambulatory Health Care Administration (CRAHCA), 1996. Phone: (303)-397-7888 to order ($70). [ISBN 1-56829-015-2]

Berwick DM.
"Controlling Variation in Health Care: A Consultation from Walter Shewhart." *Medical Care.* 1991; 23(12). [Shows "statistical thinking" applied to medicine] "Run to Space" and "Why the VASA Sank: Managing for World Class Improvement" (video of keynote addresses to the 7th and 9th, respectively, Annual National Forums on Quality Improvement in Healthcare). Boston, MA: Institute for Healthcare Improvement (call (617)-754-4800 or visit www.ihi.org for purchase information). [Brilliant videos!]

Carey R & Lloyd R. New York: Quality Resources, *Measuring Quality Improvement in Healthcare: A Guide to Statistical Process Control Applications,* 1995. [ISBN 0-527-76293-8]

STATISTICS AND REALITY – PART I

Executive Learning Inc. Brentwood, TN: Executive Learning Inc. (call 1-800-929-7890 for preview & purchasing information).
[Materials are *very* well designed and use a context of health care. The "Variation" video and the series "Data planning," "Data Collection," and "Pareto Analysis" videos are particularly outstanding—brief and to-the-point]

Joiner B. *Fourth Generation Management: The New Business Consciousness.* New York, NY: McGraw-Hill, Inc., 1994. [ISBN 0-07-032715-7]

Langley GJ et al. *The Improvement Guide.* San Francisco: Jossey-Bass, 1996. [ISBN 0-7879-0257-8]

Wheeler DJ.
   Understanding Variation: The Key to Managing Chaos. Knoxville, TN: SPC Press, Inc., 1986.
   Building Continual Improvement: A Guide for Business. Knoxville, TN: SPC Press, Inc., 1998.
   Classics from Donald J. Wheeler. Knoxville, TN: SPC Press, Inc., 1999.
   [A composite of 25 one-page absolute "gems" from his columns in Quality Digest]

DAC EDUCATION & TRAINING COURSE REVIEWERS

ASQ's Education and Training Board is responsible for reviewing and approving all new and current ASQ course offerings. The Board reviews the business aspect of course proposals, but we rely on the divisions for review of the technical content of the courses. We anticipate an increased number of new course proposals this year, particularly courses that will be offered on-line.

We are asking each division to identify three to five people who would make good technical reviewers for proposed courses. Please send those names to John Dew, ETB-DAC Liaison, at jdew@aalan.ua.edu by early August.

   Thanks,

John Dew
Dr. John R. Dew
Director for Continuous Quality Improvement
The University of Alabama
205-348-9831