

Enumerative and Analytic Studies

Description versus prediction

Donald J. Wheeler

The ultimate purpose for collecting data is to take action. In some cases the action taken will depend upon a description of what is at hand. In others the action taken will depend upon a prediction of what will be. The use of data in support of these two types of action will require different types of analyses. These differences and their consequences are the topic of this article.

DESCRIPTIONS OF WHAT IS AT HAND

When the action to be taken will only affect the current situation, the lot of material on hand, or the product already made, then the problem of data analysis is essentially the problem of using measurements for description. The interest centers in what is, not how it got that way, or what it ought to be, or what it might have been. This is what Dr. W. Edwards Deming called an enumerative study.

In the case of 100 percent testing there is no statistical inference to be made in an enumerative study. When the testing is non-destructive each item can be sorted according to the test result and will be either shipped, reworked, downgraded, or scrapped. Here the daily summary is descriptive and complete. The only uncertainty in the result is the possibility of some error in the measurement or the test or the verification procedure.

When the uncertainty in the measurement procedure is insignificant we consider these daily summaries to be exact. When this uncertainty becomes substantial we may need to evaluate the measurement procedure to assess the degree of uncertainty in our 100 percent test and verification procedures.

DESCRIPTIONS OF WHAT IS AT HAND BASED ON SAMPLES

While 100 percent inspection answers the enumerative question, it can be expensive. Moreover, 100 percent inspection does not work when measurements are destructive. So the enumerative problem quickly became one of how to work with a sample rather than a census. When only some of the material at hand is measured, or only some transactions are audited, our description of the current situation will depend upon an assumption that the measured items in the sample are essentially the same as the unmeasured items in the lot. This assumption of a representative sample is the key to taking the right action on the lot.

Clearly, the assumption of a representative sample will rest upon how the sample is obtained from the lot. Here the specifics of the situation and the consequences of the actions contemplated will need to be considered.

Logic tells us that in order to have a reasonably representative sample we will need to use a sampling method that provides an "equal and complete coverage" for the lot. Complete

coverage means that the list from which the sample items are selected must effectively include the whole lot, or very nearly the whole lot. Equal coverage means that each item on the list has the same chance of being included in the sample. This equal and complete coverage is what makes random sampling plans the preferred approach for enumerative studies.

When we have a random sample, the uncertainty in the extrapolation from the sample to the lot may be characterized by the formulas of statistical inference. Descriptive statistics are taken as point estimates for the lot as a whole. Confidence intervals are used to define ranges of descriptive values that are logically consistent with the observed data. And tests can be used to determine if the lot meets some specification value. (This is the context in which Student's *t*-test was developed—as a master brewer for Guinness, W. S. Gossett wanted to know how much to pay for shipments of hops, etc. As a shipment was unloaded periodic samples were obtained from throughout the shipment and these were tested to determine the average value for some characteristic of the ingredient.)

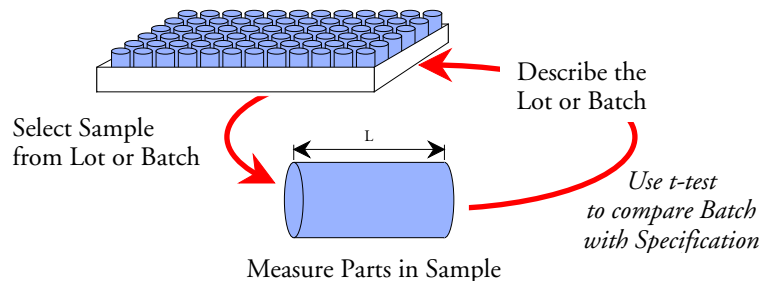


Figure 1: An Enumerative Study

Thus, in an enumerative study, if the sample has been obtained in a manner that is likely to make it representative of the lot, then the methods of statistical inference may be used to quantify the uncertainty in the extrapolation from the sample to the lot as a whole. And the action taken on the lot will depend upon two things: a judgment about the representativeness of the sample, and the outcome of the statistical inference.

Of course, when a sample does not properly represent the lot, both the inferences and the actions taken can be incorrect. This is why the consequences of an incorrect action determine how much effort should be applied to obtaining a sample that will be reasonably representative. While random sampling plans are preferred, they can be complex and hard to implement in practice. What is reasonable for research may not be feasible in production. This is where we get into using systematic or judgment samples (like Gossett did) in place of random sampling. Regardless of how the sample is obtained, before we are ready to take action on the lot we have to make a judgment about the assumption of representativeness for the sample.

THE TWO SAMPLE T-TEST

Gossett solved the simple enumerative problem when he gave us Student's *t*-test. One of the next questions was "How can we compare two or more batches?" The two-sample *t*-test, and Sir Ronald Fisher's generalization thereof, the *F*-test, were developed over the next seventeen years.

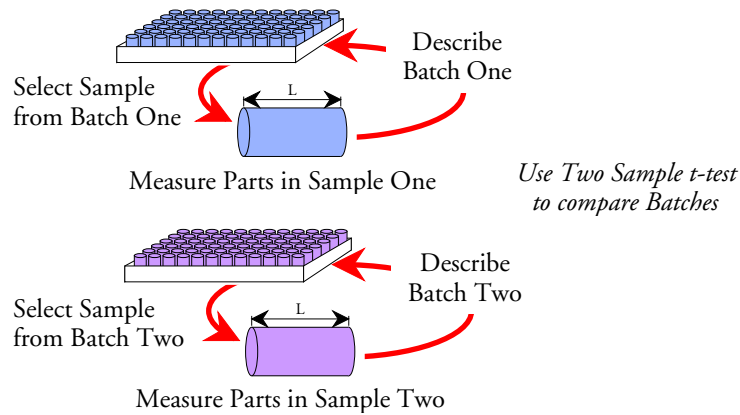


Figure 2: Comparing Two Batches

However, with this generalization of Gossett's result there was also a change in the nature of the question asked of the data. Fisher was working at Rothamstead Experimental Station analyzing the results of agricultural experiments. By changing from the comparison of two batches of material on hand to the comparison of two conditions, such as two varieties of wheat, the question became one of predicting what might be. Not only do we want to know if variety A had a higher yield than variety B during the past season, but we also want to know if this difference will persist in future growing seasons. So a description of what had happened in the past would no longer suffice—we needed to extrapolate from historical data into the future.

In general, the agricultural approach to this question of prediction has been to use a large number of experimental plots, representing many different environmental conditions, and to assign these plots in some structured, but randomized, manner to each of the two varieties being tested. If a detectable difference between the two varieties could be found across all these different conditions, then it would be reasonable to expect that this difference might persist from season to season.

The basis of the scientific method is the replication of results, and the agricultural approach builds this replication into a single growing season by using many different experimental units representing multiple environmental conditions. In this context the statistical analysis consists of filtering out the probable noise due to the experimental units to see if any potential signals can be found that correspond to the treatments being studied. The alpha-level for the statistical inference does not signify if the potential signals will be useful, nor does it tell us under what conditions we can make a prediction based on these results. The alpha level simply defines how the experimenter separated the potential signals from the probable noise. When a potential signal is found (that is, when a p-value smaller than the alpha-level is observed) a judgment will still have to be made by those with the appropriate subject matter knowledge regarding both the usefulness of the result and under what conditions the result might be predictive.

But what happens when we cannot use the agricultural model of many different experimental units within a single iteration? In business and industry we commonly do not have the luxury of multiple experimental units. Neither can we wait around for a whole year to collect useful data. Data have to be collected, analyzed, and actions taken on a much tighter schedule and with a much smaller budget. Nevertheless, all of the questions of interest will still pertain to

prediction. In what follows I will look at the problem of prediction outside of the agricultural model.

PREDICTIONS OF WHAT WILL BE

When action is to be taken on the process that produces our measured items we do not care about the present conditions except as they pertain to what will be. Our extrapolation is from the measured items to items not yet made. In consequence our interest centers in the process—the forces that give rise to yesterday's, today's, and tomorrow's items. This use of data to make predictions so we can take action on the process is what Dr. Deming called an analytic study. Many of the most important questions in business, industry, and science require predictions of what will be.

Since we cannot measure that which has not yet been made, the observations used in an analytic study are often the same as those used in an enumerative study. Nevertheless, the issues of prediction are different than those of description, and so our analysis will also need to be different.

The first difference involves how we interpret our data. For the purpose of prediction it matters not whether the data represent 100 percent of the product made or represent some fraction thereof. Moreover, *since it is impossible to take a random sample from the future, in an analytic study all of our observations become a judgment sample regardless of how they were obtained.* Here our extrapolation is over time, and so it is the temporal spread of our data that becomes our primary consideration. This places a premium on systematic methods of selecting items from the product stream (as opposed to the random sampling methods used in enumerative studies). Of course the data still have to be organized rationally. That is, we have to respect the structure and context of our data when we apply them to the problem of prediction. But it is the *judgment* about the *rationality* of the application of the data that is the basis for our degree of belief in the predictions made. This why analytic studies place an emphasis on rational sampling and rational subgrouping as opposed to the many different random sampling procedures developed for enumerative studies.

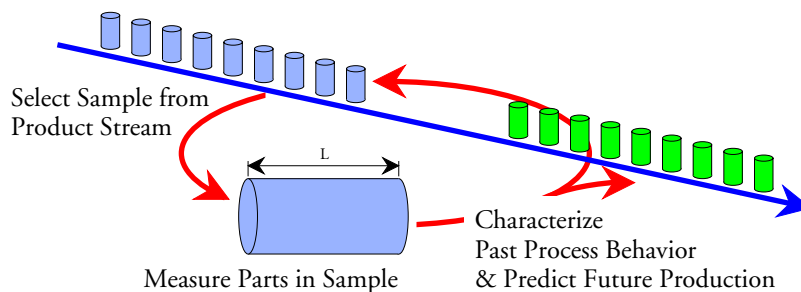


Figure 3: An Analytic Study

A second difference between enumerative studies and analytic studies is the type of statistics used. Enumerative studies tend to be built upon symmetric measures of dispersion (global measures) while analytic studies use measures of dispersion that depend upon the time-order sequence of the data or the structure of the experiment (within-subgroup measures). Since

prediction involves a characterization of the process behavior over time, this distinction is crucial.

DATA ANALYSIS FOR ANALYTIC STUDIES

The purpose of an analytic study is to take action on the process that produced the measured items. Since these actions must involve the process inputs, it will be helpful to think of these process inputs as belonging to one of two groups. The “control factors” will be those process inputs used to control the process in production. The remainder of the process inputs, both known and unknown, will be part of the group of “uncontrolled factors.”

By choosing the levels for our control factors we try to operate our process average near the target value. By holding our control factors constant we prevent them from contributing to the variation in the product stream. Thus, control factors provide a way of adjusting the process average, but have little to no impact upon the process variation.

Virtually all of the process variation will come from the group of uncontrolled factors. To reduce the variation in the product stream, and thereby to reduce the excess costs of production and use, we shall have to move some process inputs from the set of uncontrolled factors to the set of control factors.

This means that if we wish to adjust the process average we will need to study the levels for the group of control factors, but if we wish to reduce the process variation we will need to study the inputs in the group of uncontrolled factors.

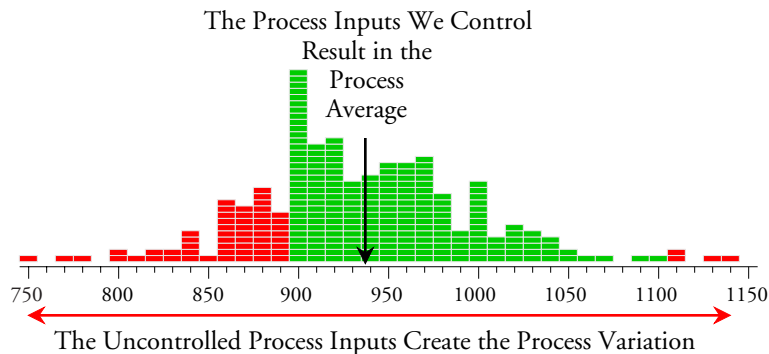


Figure 4: Process Inputs Belong to One of Two Groups

This is where the tyranny of economics raises its head. We cannot afford to make every possible process input a control factor, even if we had sufficient knowledge to do so. Moreover, the very large number of inputs in the group of uncontrolled factors makes it exceedingly difficult to identify which inputs to study. So, instead of seeking to study the effects of selected process inputs, we need an approach that will not only allow us to learn about the large number of process inputs in the group of uncontrolled factors, but will also allow us to do so without having to identify specific inputs in advance. And this is what the process behavior chart allows us to do.

A process behavior chart makes no assumptions about your process or the data that characterize the product stream. It simply lets the data define both the generic process potential and the actual process performance. It then combines these two in a single graph to allow the

user to characterize the process as behaving either predictably or unpredictably.

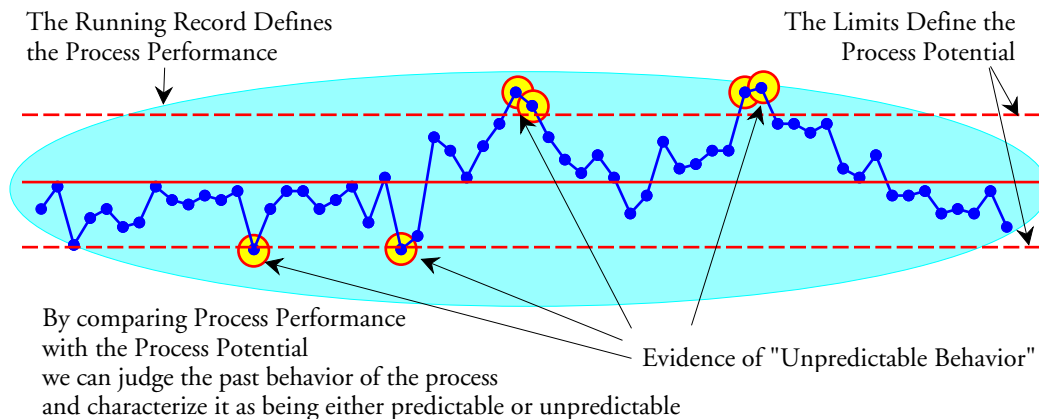


Figure 5: We Characterize a Process by Comparing Performance with Potential

Since the group of uncontrolled factors are the source of the variation in the product stream, this characterization of the process behavior tells us something about the group of uncontrolled factors.

PREDICTABLE OPERATION

If the process has been operated predictably in the past, we think of the group of uncontrolled factors as consisting of a large number of cause-and-effect relationships *where no one cause has a dominant effect*. The routine variation of a predictable process is said to be the result of common causes.

Here it would be a mistake to single out an uncontrolled input and add it to the set of control factors. The large number of inputs, plus the lack of any single input having a dominant effect, conspire to make process changes uneconomical. When operated predictably a process is operated up to its full potential. It will have the minimum variance that is consistent with economic operation. This is why seeking to control a common cause of routine variation is a low payback strategy.

When the process has been operated predictably in the past, then it is logical to expect that it will continue to be operated predictably, and the past behavior is the basis for our predictions of what will be.

UNPREDICTABLE OPERATION

If the process shows evidence of unpredictable operation we think of the group of uncontrolled factors as containing one or more cause-and-effect relationships *whose dominant effects show up above and beyond the background of the routine variation*. The causes with these dominant effects are called assignable causes of exceptional variation.

Here the process is offering strong evidence that uncontrolled inputs with dominant effects exist. When a process is operated unpredictably it is being operated at less than its full potential. It will not have minimum variance, and it will not be operated economically. When assignable

causes of exceptional variation are present it will almost always be economical to identify these process inputs and make them part of the set of control factors.

This is a high payback strategy for two reasons. When we make an assignable cause part of the group of control factors we not only gain an additional lever to use in adjusting the process average, but we also remove a large chunk of variation from the product stream. In this way, even though the process may have been unpredictable in the past, we learn how to improve the process and come closer to operating it predictably and on-target in the future.

It is not logical to assume that a process that has been operated unpredictably in the past will spontaneously begin to be operated predictably in the future. Unless we intervene, the uncontrolled process inputs known as assignable causes, with their dominant effects, will continue to take our process on walkabout. So, even though our natural process limits may approximate the hypothetical process potential, no computation can provide a reliable prediction of what an unpredictable process will actually produce.

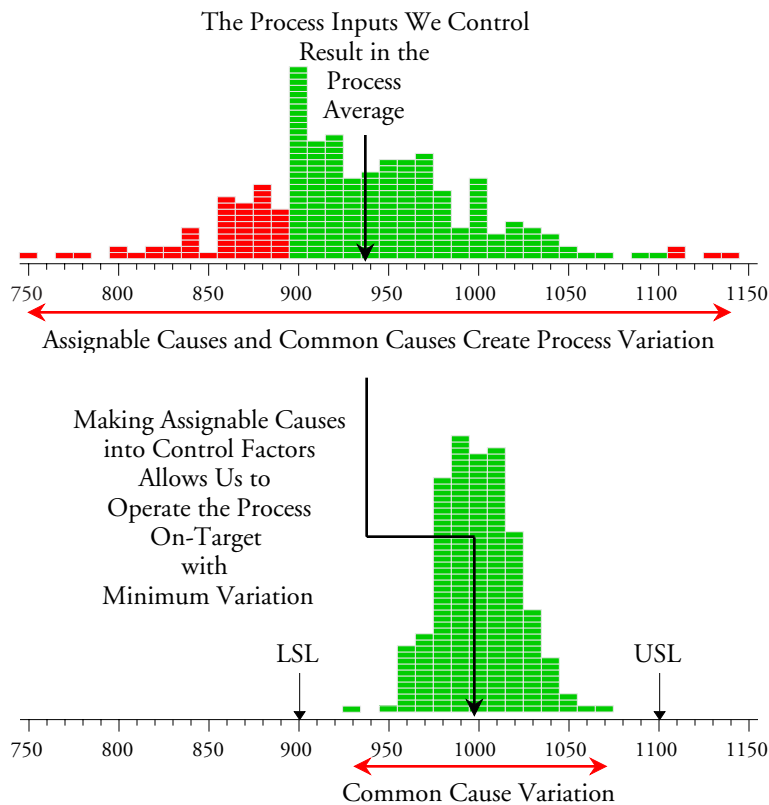


Figure 6: Assignable Causes Belong in the Set of Control Factors

SUMMARY

Enumerative studies seek to describe what is. They enumerate or estimate what is at hand in order to take action thereon. Statistical inferences can cover the uncertainty involved in the extrapolation from a representative sample to the lot, but they cannot begin to cover the uncertainties due to having a non-representative sample. In the end, some judgment has to be

used in deciding if the sample is logically representative of the lot before action can be justified.

Imposing the techniques and requirements of enumerative studies upon an analytic study is a sign of confusion.

Analytic studies seek to make predictions so that appropriate actions can be taken on the production process. Here the extrapolation is over time, and all samples become judgment samples. The temporal spread of the data and the systematic selection of the items to measure are the important issues.

In an analytic study, experimenting with the levels of the control factors may let you tweak the process average, but it will seldom lead to a reduction in the process variation. Process variation comes from the group of uncontrolled factors, and the process behavior chart lets us consider all of the uncontrolled factors together.

A predictable process will exist as a well-defined entity. It will have a consistent process average and it will be operating with minimum variance. We may use our data to make predictions, and no action may be needed.

An unpredictable process will not be operating with minimum variance. Estimates of process characteristics will be premature. Computations are useless for prediction—action is required. Until the assignable causes are found and made part of the set of control factors all prediction is futile. Working to make assignable causes part of the set of control factors is a high payback strategy.

Looking for assignable causes of exceptional variation when the process is being operated predictably will be a waste of time and effort. Seeking to control common causes of routine variation is a low payback strategy.

Assuming that the past will predict the future when assignable causes are present is merely wishful thinking. Failure to control assignable causes of exceptional variation will increase process variation and result in excess costs downstream.

Finally, the only reason to collect data is to take action. Probabilities associated with the data analysis cannot justify any action. Whether the analysis is enumerative or analytic, the results will have to be judged in the light of subject matter knowledge before actions can be taken. So while the elaborate constructions of mathematical statistics and the resulting probabilities may make the decision problem appear to be rigorous and exact, in the end actions always involve some element of judgment.