

Quality Procedures in Statistical Sampling

Mary Batcher and Wendy Rotz
Ernst & Young LLP

Abstract

Ernst & Young's Quantitative Economics and Statistics Group (QUEST) conducts over 150 samples annually to estimate values in regulatory, audit, and litigation settings. There is constant pressure in private industry to deliver complex analyses with short notice – within a week, or even within a day. Often, after the sample was drawn and the study is in progress, newly discovered information will alter the sampling frame requiring design adaptations and more detailed computations later during estimation. When audit and survey results are returned for estimation months later, staff who performed the original work may be juggling multiple projects and will work late hours to accommodate the deadline, or may even be unavailable due to turnover or their current assignments. In this setting there are many factors that could lead to errors, inefficiencies, and difficulties in reproducing the work. Yet trust in our quality is critical to our business, efficiency is essential for profitability, and a clear audit trail is mandatory. With support from top management, applying lessons from Juran, Deming, and basic quality tools, QUEST developed standard procedures, self-documenting processes, automated quality steps, and check sheets to assure high quality and accurate answers are delivered timely, within budget, and with an easily reproduced audit trail.

Key Words: Quality, Process, Control, Sampling, Statistical, Consulting

1. Introduction

Historically, statisticians were consulted to facilitate quality improvement efforts in assembly lines that manufacture consumer products. In recent years quality improvement tools have been applied to more traditional white-collar processes such as engineering and architectural work, software design, even processing paperwork. Today, businesses and government agencies alike consult statisticians to aid them in using statistical quality tools for these tasks.

Yet, how often do we as statisticians turn our quality tools and experiences to examine and improve our own statistical consulting process? We have customers who employ us; we follow processes to do our statistical work; and in the end, we deliver products such as our calculations, tables, estimates, or statistical designs.

How can we as statisticians practice what we teach by applying these same quality tools to our own work. Process flow charts, fishbone diagrams, the quality improvement cycle and real statistical measures can be applied to our statistical consulting.

This paper is a case study from Ernst & Young LLP.

2. Background

Ernst and Young is a big-four accounting firm with over 30 statisticians, economists and analysts in the Quantitative Economics and Statistics (QUEST) group in Washington D.C. QUEST conducts 100 to 200 samples annually mostly for the purpose of estimating values for corporate tax returns.

There are occasional late hours to meet a demanding schedule. Sometimes QUEST is asked to design a sample or produce estimates in less than a day. QUEST statisticians juggle multiple clients so, especially when there are tight deadlines and/or during peak periods, there may be high turnover on individual projects as team members are temporarily unavailable and new team members are pulled in to meet immediate client needs.

One such peak period occurred at a time when there were many new staff and fewer managers. We needed a method of doing complex sample designs faster to respond to tight deadlines and handle a higher workload, without compromising quality.

With the pressure of tight deadlines, high workload, newer staff, and long hours, the age-old quality problems began to occur: errors and inefficiencies. Errors caused rework which in turn caused delays and additional project costs. Sound familiar? These are the common complaints statisticians hear in quality improvement consulting.

Worse than the rework, errors, if detected too late, could erode client trust or deter acceptance of statistical methodologies. Errors are just plain bad for business – any business.

This was our impetus for change.

To begin, a high level of management asked for a stop-gap measure: a check sheet for typical items we deliver that would ensure key quality checks are conducted every time.

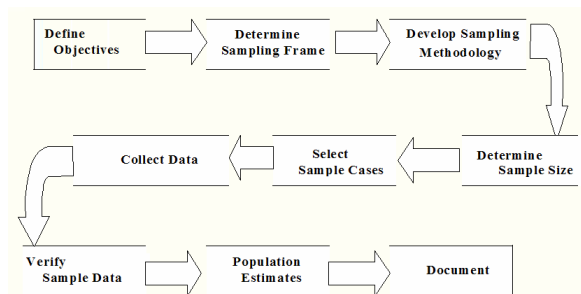
However, this was a time to recall a fundamental principle: quality is built into the process, not at the end. Indeed Juran applied the Pareto principal to quality: 20% of all defects cause 80% of the problems. We know that quality improvement efforts focusing on people, such as training, reminders, job aids, and check sheets, will only marginally improve quality. The majority of quality gains come from process improvements.

Job aids in the form of check sheets, while helpful, would not be a key solution by themselves. When developing these check sheets, the process needed to be assessed, and adjusted to prevent errors if possible, or catch them at a much earlier stage – ideally by the staff performing the tasks.

3. Process Flow Chart

Our process had been long established and we already had a process flow chart. See Figure 1.

Figure 1. Sampling Process Flow Chart



At the time we were following typical steps in a statistical project: 1) work closely with the firm’s tax professionals, who are QUEST’s internal clients, to define the objectives and necessary estimates, 2) determine how to construct the sampling frame and what is included in the population with input from internal and external clients, 3) develop a sample design with sample size options, 4) discuss the sample sizes with the clients to determine the right sample size that balances costs with accuracy needs

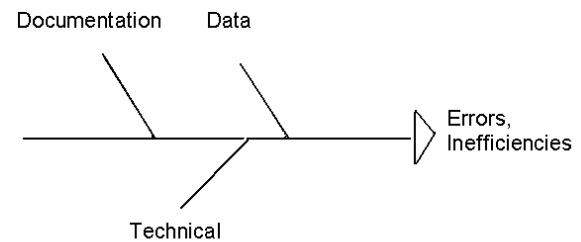
and sampling risks, 5) draw the sample, 6) collect the necessary information (done by tax professionals), 7) check the data returned to the statisticians, 8) produce estimates, and finally, document the steps and the statistical methodology.

As it turned out, a major flaw in the process was that this last step was – well, last. This is discussed more below.

4. Error Analyses and Fishbone Diagrams

It appeared that along this process there were many types of errors that could occur and at they seemed to have several unrelated causes. But after consideration, they fell into three major groups: documentation, data handling, and technical application, as illustrated in the fishbone diagram below. See Figure 2.

Figure 2. Main Fishbone Diagram



Note that time pressure was not included as a cause in these diagrams. Rushing through technical work does indeed contribute to errors, and managers, when possible, will negotiate more reasonable deadlines. However, the nature of our business requires us to respond to tight deadlines, so our process must be able to deliver high quality work in time-sensitive settings.

Each cause, or “fin” in Figure 2, can be examined as a fishbone by itself. See Figures 3 through 5.

4.1 Documentation

The root cause of most of the problems was documentation. Some errors and rework occurred because important nuances had not been conveyed when a project was transferred or had been forgotten until managerial review.

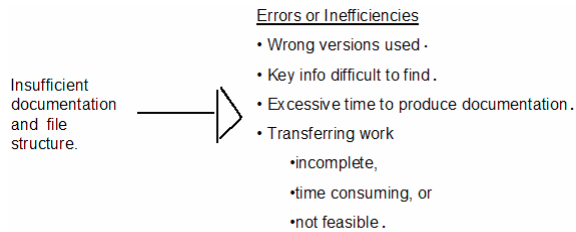
In addition to rework from errors, there were inefficiencies in transferring projects from one staff to another. Incomplete verbal reporting combined with a sea of extensive, but not well organized documentation, led to delays in handing over projects, writing reports, and responding to requests

for work papers. There were instances when this difficulty could preclude bringing on available staff, as the time spent explaining the project's history might outweigh the time necessary to complete the next required tasks.

In some instances identifying the most current data sets and figures was challenging and some rework occurred because obsolete data were used erroneously.

Initially, it appeared as though there were several root causes associated with documentation such as wrong versions used or the difficulties in transferring work. However, when analyzed, there was really only one underlying cause: the documentation and filing structure were inadequate. See Figure 3.

Figure 3. Documentation Fishbone



A self-documenting process was needed. The project stage, current versions, and completed work needed to be self-apparent. An ongoing history of the project was needed without requiring the review of dozens and dozens of files and other documentation. A self-documenting process would ease the transfer of projects among staff, reduce errors due to wrong versions, expedite writing final reports, and facilitate responding to documentation requests.

To meet these needs, standardized paper and electronic file structures were developed. Electronic files were kept in a centralized location, not on individual personal computers. The files were organized by stage in the process, not by the individual performing the work.

Naming conventions were used for electronic files including the revision date in the file name. Programs were numbered sequentially. More automated tables required for checks and reports were incorporated into our SAS code templates.

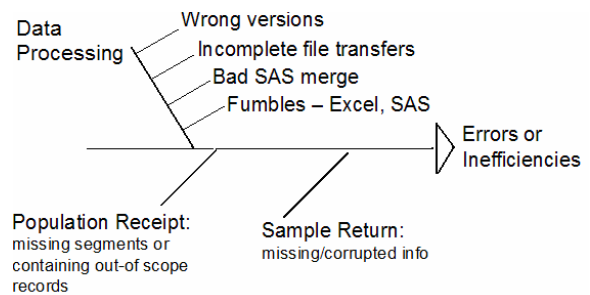
Check sheets for the key statistical steps were developed and the final check on each one was the "Could I be hit by a truck?" check. This required

that the files left in the directories would be kept current and self-apparent so that anyone could easily pick up where the project was left off. This included the removal or at least separation of old versions as soon as new versions were created, as-you-go documentation, and clean SAS runs and logs at the end of each task.

4.2 Data

The next major problem area was data, including population data, sample data and data processing. See Figure 4 below.

Figure 4. Data Fishbone



In statistical calculations there are numerous opportunities for data processing snafus. One of our problems was identifying the correct version of the data, but this was addressed through the documentation solutions discussed above. There are many other common data processing problems in statistical work.

Data can be lost or corrupted when electronically transferred or when imported from one medium to another. Merging of files can have unexpected outcomes and there is always a possibility for odd coding or unknown SAS idiosyncrasies.

Excel formulas are especially error prone when cells are copied and pasted for repeated calculations. Excel calculations are also prone to incomplete updating when new data is used in an old spreadsheet containing a mixture of automated and manual data and formulas.

Our process needed to prevent or detect these problems early. Our staff now track control totals through all data steps from the receipt of system files for the data frame to the production of the final report. Automated steps in SAS templates provide most of the totals required. A policy was developed to subset data rather than delete records. This allowed the removed records to be examined and accounted for.

Then, there is the data itself. A prevalent and expensive data problem is correcting the sampling frame after a study is underway. We typically construct a sampling frame from administrative records that were collected for an entirely different business purpose. Sometimes, inadvertently, we do not receive a portion of the population. More often, there are many out-of scope records. We rely on the clients, lawyers, accountants, and tax professionals to provide direction on the assembly and trimming of system files to create the sampling frame. However, even the most knowledgeable clients are often surprised by the motley data found in their system once a sample is drawn and some of the records receive closer scrutiny.

Checking the adequacy of the old design for the new frame, altering the design if necessary, and adapting the sample is expensive rework. This introduces opportunities for errors, design inefficiencies, and causes new versions. It would clearly be best avoid changes to the frame at a late stage in the process.

One measure we found helpful is to echo back summaries of the data files we received including counts and dollar totals by key categorical variables. This enables those who are familiar with the scope issues and the company's data to more readily recognize out-of-scope categories and potentially identify key missing areas.

Another source of data problems is the sample file returned for estimation. Non-statisticians, though well intended, may unwittingly corrupt a sample by adding new records or making substitutions. Unaware of the importance of statistical data fields, such as stratum indicators, sample files may be returned with these fields either missing altogether or scrambled in a bad Excel sort.

We check control totals by stratum in the returned samples and perform other checks. To avert these kinds of problems in the first place, when we deliver the sample, we also include communication concerning data handling. We explain the importance of the statistical fields and warn not to make substitutions or additions to the sample selections. Sometimes we even add a column for the data to be entered and include basic built-in checks in the sample spreadsheet we deliver.

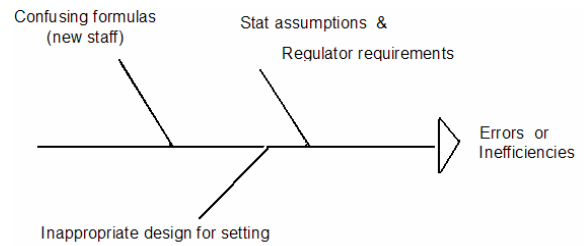
4.3 Technical

Interestingly, this area had been given the most managerial attention until this point. However, after

analysis, it was determined to be the least significant source of errors and inefficiencies. See Figure 5.

New staff would occasionally confuse formulas, managers needed to verify designs were appropriate for the setting, and occasionally work was turned in for review without verifying that statistical assumptions or regulatory requirements were met. These were less frequent, relatively easy to spot, quick to correct, and required less rework.

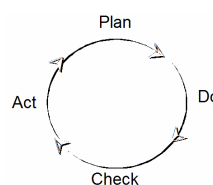
Figure 5. Technical Fishbone



Nonetheless, to assure technical quality, managers are involved at all stages coaching judgment calls, staff receive on the job training, and we have standardized SAS code covering 90% of our calculations. Also, in our new process, the verification of key statistical assumptions and regulatory requirements does not begin at the estimation stage. Instead, we ensure the sample design is likely to yield results that meet assumptions and requirements at the design stage.

While these technical improvements were helpful, we found emphasizing the practical implementation, such as data handling and documentation, had more impact on our quality success.

5. Quality Cycle



When developing check sheets, we took the opportunity to adjust the process. We added critical steps, standardized the file structure, and began a self-documenting process. We

required staff to do basic checks prior to managerial review. Was it helping?

Recall the Deming/Shewart PDCA quality cycle is: Plan, Do, Check, Act, then it is back to Plan again. Outside manufacturing, businesses often have difficulty determining appropriate measures for non-tangible products. What was appropriate in this setting?

A summary of our measures is below. We found that the additions to our SAS code, standardized and current documentation together with key checks at early stages did facilitate our efforts. Documentation requests, transferring a project, designing samples, and producing reports all take markedly less time.

Accuracy was not measured prior to implementing the quality steps. However, when the new check sheets were first implemented, staff identified two or three problems weekly. Now that number is down to 1 or two per month.

Figure 6. Measures

	Before	After
Documentation Requests	2-3 weeks	2-3 days
Transfer Projects	2-4 hours	15 minutes
Design Sample	8-16 hours	< 4 hours
Report	10-20 hours	4-6 hours
Early Error Detection	not measured	First month: 2-3 per week. Now: 1-2 per month

Quality improvement is indeed a continual process. When a new kind of error is found, we reassess the process, assess the need for extra steps or checks, and make the necessary changes.

We also have ongoing meetings to discuss changes to the process, SAS templates, and checks. In addition, there are ongoing training sessions for each stage in the process to discuss statistical theory, raise awareness of developing issues, and share the kinds of errors that have been found at that step in the process.

6. Conclusions

Upper management has been wonderfully supportive of our efforts and has encouraged the commitment of additional resources to quality improvement. In a wider quality improvement effort, many of the practices developed for sampling were adapted for QUEST policy on all projects.

We found that our preventative measures reduced errors, but we still need to check for them. Our check sheets have undergone several revisions since their initiation. The process, code, and checks are continually adapted as necessary.

Increasingly more steps have been automated. These have streamlined efforts, improved our response times, and have reduced opportunities to introduce errors into our calculations.

However, there was an unforeseen negative outcome. Before the extensive automation, we had a less efficient process and certainly more errors were made. However, more skills were needed for basic designs and staff learned from simple designs and learned from their errors. They were able to work independently and advance to more difficult projects sooner.

Now we have a highly efficient process with fewer errors. However, there is less learning from basic assignments. While new staff may be able to function quickly for our common projects, it now takes staff longer to work independently and to be able to tackle more advanced statistical problems. Stretch assignments are more of a stretch for newer staff.

Therefore, our managers now need to be more active to ensure newer employees are gaining the technical experience they need from their assignments and are given sufficient challenges. Incorporating staff into ongoing process improvement activities is a means of providing these growth opportunities.

In summary, we now more actively develop staff and spend less time doing rework. We reduced the time and therefore costs to complete typical tasks. Our built-in documentation is more efficient and there are fewer errors made in calculations.

In conclusion, we learned that as statisticians, we can indeed improve our own statistical work using our familiar quality tools.

References

[1] Brassard, Michael, *The Memory Jogger Plus + Featuring the Seven Management and Planning Tools*, Goal/QPC, 1989

[2] Bluvband, Zigmund, *Quality's Greatest Hits: Classic Wisdom from the Leaders of Quality*, Quality Press, 2002

[3] Brecker Associates Inc., *Quality-Based Problem-Solving / Process Improvement*, <http://www.brecker.com/quality.htm>

[4] Juran, Joseph M., *Juran on Quality by Design : The New Steps for Planning Quality into Goods and Services*, The Free Press A Division of Simon and Schuster Inc 1992

[5] Juran, Joseph M., Godfrey, A Blanton, *Juran's Quality Handbook 5th ed.*, McGraw-Hill, 1999

[6] Reh, F John, *How the 80/20 rule can help you be more effective*,
<http://management.about.com/cs/generalmanagement/a/Pareto081202.htm>

[7] SkyMart, *Management Glossary*,
<http://www.skymark.com/resources/qualglos.asp>

[8] Syrett, Matthew, *Do You Really Understand the 80/20 Rule?*
<http://www.marketingprofs.com/5/syrett10.asp>,
October 18, 2005