COGENT
ECONOMICS

Statistical Sampling in Mortgage Quality Control

By Hakki Etem

Introduction

Statistical methods are well-established tools for efficiently measuring and improving product quality in a variety of industries. Unfortunately, statistical analysis has been slow to gain acceptance in the mortgage industry, although the ability to originate the best quality product at the lowest possible cost is just as valuable to mortgage originators as it is to automobile manufacturers. There are many reasons why the mortgage industry has avoided statistical methods, but surely one reason is the subject itself: few disciplines can be as mind-numbing as statistical theory.

Nevertheless the most effective way for QC managers to measure and improve loan quality -- at the lowest possible cost -- is to employ statistical methods. Although this means that QC managers must necessarily become familiar with basic statistical concepts, with the right tools and professional support the process can be greatly simplified. The COGENT Statistical Sampling Methodology is such a tool.

Overview

In the mortgage industry context, a sampling methodology is a *strategy* that begins with a question: What information does the QC manager need to learn from reviewing loans? Peter Drucker would ask two more basic questions: What is the task? Why do it at all?

Following the conventional wisdom, or continuing to do mortgage QC "the way it has always been done," is not a particularly reliable starting point. Mortgage QC has often been viewed as a necessary evil, required by agency regulations and performed with apathetic management support. Another view, one adopted by hundreds of vibrant companies, holds that QC can reduce the cost of poor quality products, strengthen the company's competitive position, and ultimately return far more to the company than its direct expense. This view requires a thoughtful response to the deceptively simple questions posed above. That response becomes, in turn, the foundation of a sound and efficient sampling methodology.

Efficient sampling involves a central question: How many loans must be sampled, and how, to gain the necessary information? Statistical analysis is the tool that, combined with a sound sampling methodology, allows the QC manager to determine the minimum number of random loan reviews required. Without statistics, a QC department will almost certainly either sample too many loans, an inefficiency with a measurable cost, or sample too few loans, which means no useful information is gained.

Production QC in the Mortgage Industry

The basic role of the QC department is to monitor loan quality trends and respond alertly to unusual spikes of poor loan quality. That response occurs within a feedback loop to production personnel, who are in a position to effect changes in the origination process and bring loan quality back within the lender's normal or acceptable range. Without statistics, it is impossible to determine reliably what the boundaries of the normal quality range are. The best QC managers in the most enlightened companies will, together with production staff, participate in what Joseph Juran calls *quality improvement* projects: short-term efforts designed to permanently improve the normal quality range. Here again statistical analysis is required to measure the effectiveness of

the quality improvement project: Has the normal quality range truly shifted towards higher quality?

The QC process - monitoring loan quality and providing feedback - implies that reviewed loans must be divided into two categories, acceptable and defective. These qualitative assessments must be uniform, valid, and accepted by the lender's investors and production departments. If these conditions are not met, the effort to stabilize loan quality is compromised; unfortunately, the vast majority of mortgage lenders have yet to establish this elementary quality foundation. Many mortgage QC departments, for example, do not make this summary, qualitative assessment of reviewed loans. Those that do rate loans as acceptable or defective, moreover, do not use uniform criteria, since the industry lacks a common definition of a defective loan.

For lenders who sell loans into the secondary market, the most common definition of a defective loan is one that is technically "unsalable", or does not meet investor standards. Such a loan may require indemnification or repurchase, which has a measurable economic cost. For portfolio lenders, the capacity of a loan to make timely payments of interest and principal is surely central to the definition of mortgage quality. But few QC managers have validated their definition of a defective loan by comparing the actual performance of reviewed loans (i.e., default rates) with an appropriate benchmark. And if this validation is lacking, it is unlikely that the production departments will respond in a meaningful way to QC feedback.

That feedback, finally, must be presented clearly and succinctly if it is to be effective. All QC departments prepare monthly reports to senior management; unfortunately, these reports often amount to a list of every exception noted in every file reviewed. They are too long and contain little information (as distinct from data), which means they are rarely read. An unread report is of no use whatsoever.

Statistical Sampling Methodology

Earlier we observed that the design of a QC sampling strategy must begin with a question: What information do we seek to learn from the sample? Our information goal is to make a statement about overall loan quality -- in other words, to analyze a small sample of loans so that we can draw a conclusion about the quality of the loan population from which the sample is drawn. Statisticians call this a *sample inference.* For example, if a small sample is drawn from a population and the sample defect rate of the sample is determined to be 5%, we might infer that the population defect rate could be as high as 10% (a precision of + or - 5%). Note that an emphasis on sample inferences implies less importance on fixing the problems found in a given sampled loan - the real concern is the population of loans. To be effective, the QC manager must keep the focus on the forest rather than the trees.

If loan quality findings are organized so as to eliminate statistical bias, useful sample inferences can be made for a number of segments of the total loan population. These segments should be defined in a way that reflects *production (or process) defect risk*; in other words, the quality of loans within a given segment should be more similar, on average, than loans from different segments. This generally means dividing total production volume by origination source type (retail, wholesale, etc.) Each source type, in turn, can be further divided by origination source unit (retail branch, wholesale broker, etc.) Two final levels of detail can be obtained by targeting unit staff (individual underwriters, loan officers, and processors) and by targeting other loan characteristics (product type, geographic location, appraiser, etc.). The primary benefit of this kind of

organizational hierarchy is the opportunity to monitor overall quality trends at the higher levels, and then to trace the origin of poor quality spikes to the unit and individual levels.

Although they can overlap, production risk is quite different from *performance (or default) risk*. All loan product types, for example, generally have a distinct default risk. To illustrate, a group of 15-year, 80% LTV, single-family fixed-rate mortgages is expected to have fewer delinquencies over time than a similar group of 30-year, 90% LTV adjustable-rate mortgages. If sample selection were based on delinquency risk – a common industry practice – then relatively more 30-year ARMs would be selected for QC review. The problem with this approach, however, is that QC can only monitor production defects and, through the feedback loop, seek to stabilize production quality. Two loans with identical origination processes have, by definition, equal opportunity to incorporate a production defect – yet these same two loans, because of their different product types, FICO scores, or locations, can have different delinquency potential. Unless the QC focus is kept on production risk, sampling will be inefficient at best and, at worst, provide a misleading impression of production quality.

Determining how many loans to sample is the next critical issue. Obviously there is a cost associated with completing a QC review -- industry estimates range from $100 to $200 per sampled loan -- so there is a strong incentive to minimize sample size. The statistical precision of a sample inference is a direct function of sample size: the greater the precision requirement, the larger the sample must be. The question becomes, therefore, how much precision is necessary?

At the portfolio level, a minimum precision requirement has been established, more or less, for lenders that sell loans to the government agencies (Fannie Mae, Freddie Mac and Ginnie Mae). The old agency-mandated 10% random sample can be replaced with a smaller "statistical" sample so long as the sample inference provides 2% precision at 95% confidence. We will defer, for now, what this means -- but the consequence to a typical lender is an annual sample size of around 1% of annual production. This means that a mid-sized lender originating 30,000 loans per year (and not using statistical sampling) can reduce its number of purely random QC reviews by 2,700. With the cost of a QC review at $100 per file (at least), this lender might re-allocate as much as $270,000 in annual QC costs by using a statistical methodology. In general, these savings grow as volume increases and quality improves.

The critical advantage of a smaller, purely random sample is that the lender can then emphasize discretionary reviews of loans with the greatest potential risk. For discretionary reviews - selected on the basis of origination source, source unit, or targeted criteria - no agency-mandated precision level exists. The desired level of statistical precision for these production segments should reflect, therefore, each segment's unique risk potential to the lender.

In general, the Cogent Sampling Methodology allows lenders to allocate the total sample of loans so that proportionally more loans are reviewed from the highest risk sources. These may be new and untested brokers, branches, loan agents, loan products, etc. - or they may be sources known to have had quality problems in the past. These "intelligently targeted" reviews provide the most value to lenders in terms of risk management and mitigation.