Sampling – The Key to Process Validation
Christopher H. Paskach, Partner, KPMG LLP
Michael J. Carter, Manager, Six Sigma Black Belt, KPMG LLP
May 13, 2011

Abstract

The increasing volume and complexity of electronically stored information and the cost of its review continues to drive the need for development of sophisticated, high-speed processing, indexing and categorization -- or "predictive coding" -- software in response to litigation and regulatory proceedings. Since the majority of these tools rely on sophisticated, proprietary algorithms that are frequently referred to as "black box" technologies, there has been a reluctance to exploit their expected productivity gains for fear that the results they produce may be challenged and rejected as not meeting the required standard of "reasonableness." Effective use of sampling can overcome this concern by demonstrating with a stated level of confidence that the system has produced results at least as consistent and reliable as those obtained by having attorneys review the documents without sophisticated technology support. Through testing, based on statistical sampling, the quality improvements and cost savings promised by "predictive coding" technology can be realized.

Current State

Determining the reasonableness of a document search and review process has been based on whether there was sufficient "attorney review" of the documents to ensure that the results will be reliable. While "attorney review" has been accepted as the "gold standard" for the adequacy of a document review process, the consistency and reliability of the results produced by the attorneys has rarely been questioned or tested. The presumed effectiveness of "attorney review" is generally accepted to meet the reasonableness standard so that sophisticated sampling and testing of the attorney review is rarely performed. The sheer volumes and unforgiving production deadlines of today's e-discovery efforts demand ever increasing review capacity and throughputs. Simply scaling up the review process with more people to handle these demands is clearly at odds with cost control initiatives that are of utmost importance to corporate law departments.

New Technologies

Recent technology development efforts have focused primarily on helping review teams manage the efficiency and cost of large scale reviews. Of particular interest are the tools that help categorize and cluster documents based on document content, or identifying near-duplicate documents and grouping them for review. These "predictive coding" tools can help reviewers speed through non-responsive or similar sets of documents by bulk tagging and more quickly isolating relevant material that has to be more carefully reviewed for privilege before being produced. The very latest technologies aim to automate the review process by minimizing the need for human reviewers in a first-pass review for relevance. But regardless of where an organization falls on the automation continuum in its adoption of technology -- from traditional linear review to concept-based clustering, leveraging technology or human review -- the goal of a faster, more consistent, more predictable and less costly review requires

more than basic efficiency gains. A cost-effective document review project requires more sophisticated technology and proven Quality Control (QC) processes to demonstrate its effectiveness.

Technology vs. Human Review

While an argument for cost effectiveness of technology-based processes has been largely established, the consistency of the results "versus" human review remains a topic of ongoing discussion. For many, the validation of review quality is often subordinate to the review itself and consists of informal or casual observations that lack the scientific rigor and quantifiable measures necessary to defend the quality process. More sophisticated quality methods that rely on sampling can provide the much needed assurance that the results are at least as good as human review and when used appropriately can result in significantly improved consistency and productivity.

Over the past three years, KPMG has conducted four test projects that compared the results of an "attorney review" process with results obtained by reprocessing the same document collection with a predictive-coding software tool. The tool used in these tests uses a series of randomly selected sample batches of 40 documents that are reviewed by a subject matter expert (SME) to train the software. Based on the SME's decisions on the training batches, the software calculates the relevance of the remaining documents in the collection. In all four test cases, the software was more consistent in categorizing documents than were the human review teams.

Although the software produced more consistent results than the review attorneys, the proprietary algorithm used to produce the relevance ranking is not publicly available. However, the results it produces can be effectively tested with sampling to determine the efficacy of the automated relevance ranking process.

Assuring Process Quality

Assuring process capability, explicitly or implicitly, is a requirement for defensibility. Having a defendable, and therefore accepted, process is a matter of sound design, transparency and predictable results. Process sampling delivers all three requirements. Sampling is a well proven, scientifically rigorous method that can give the Project Manager much needed flexibility to demonstrate effectively the quality of the review process. Carefully selecting a sample, and from it inferring the condition of the larger population with high confidence in the reliability of the inference, is a powerful tool with tremendous eDiscovery utility. The process of establishing review QC using statistical sampling enables the review team to determine appropriate sample size, quantify the process risks, and determine process acceptance and rejection criteria. Then, should questions arise concerning the quality of the results, a meaningful discussion of the QC methodology can take place without the need to explain, justify or alter unproven judgmental QC practices.

Objections to Statistical Sampling

If statistical sampling can provide all of these benefits to QC in discovery review, why isn't it more widely used? There are several possible reasons, including a lack of familiarity with the method or its perceived complexity and the anticipated time investment required to understand and achieve proficiency in it. Another concern may be that a small error found in sampling could render the entire review results unacceptable. Likewise, in the discovery review process there is no clear legal precedent that confirms

the acceptability of statistical sampling methods for eDiscovery. Whatever the reasons, although sampling is widely accepted as a basic QC methodology in numerous other product and service industries to manage and quantify quality risk, it has not been widely adopted in eDiscovery review projects.

Overcoming the Objections

How can the issues that prevent wider use of statistical sampling be addressed? Overcoming the lack of familiarity with sampling can be addressed through training and the use of experts. Involving those who understand QC sampling in the process of eDiscovery can be a very effective approach to achieving the benefits and overcoming project managers' unfamiliarity. These sampling experts can assist with data stratification, determining sample sizes and calculating confidence levels for statistical inferences. One objection to this approach would be the added cost of these sampling experts. This can be addressed with a straight-forward cost-benefit calculation comparing the cost of the experts to the avoided costs of more extensive testing with non-statistical approaches. Another objection would be the risk of the supervising attorneys not being sufficiently knowledgeable to assess the quality of the sampling experts' work. This can be addressed through careful questioning and review of the experts' approach and results.

Another option to support using statistical sampling would be to programmatically integrate generally accepted QC sampling methods into widely-used eDiscovery applications. Carefully designed user interfaces for selecting samples, testing them and reporting the results could guide users through the sampling process, thereby minimizing, if not eliminating, most common sampling mistakes. Increased consistency, repeatability and reproducibility of the QC process would result.

Additionally, the sampling methodology could include periodic batch sampling throughout the review process with a mechanism for dealing with review error as soon as it is detected to reduce the need to re-perform a significant portion of the review process. Likewise, sampling error could be addressed with a set of tools that would enable sample results to be adjusted and reinterpreted in light of sampling error to reduce the risk of having to significantly expand the sample or restart the sampling process.

The final objection regarding a lack of a clear legal precedent is likely to be addressed soon by the courts, which are becoming increasingly aware of the benefits of statistical sampling in dealing with the challenges posed by very large populations of documents. Without clear legal precedent there is some additional risk to applying new technologies and relying on statistical sampling to demonstrate their efficacy. However, the benefits in terms of quality and cost of QC sampling the results from these new technologies can more than offset these risks until the legal precedents supporting their use are clearly established.

Note: The preceding commentary relates solely to process control sampling as applied in the performance of document review in connection with electronic discovery and is NOT a commentary on the maturity of sampling techniques relative to financial statement auditing.