Replicated Stratified Sampling

A New Financial Modeling Option

By Jay Vadiveloo

Products that serve consumers are increasingly complex, creating a real need to help management understand risk exposure in real time. New modeling techniques such as replicated stratified sampling may help executives better understand risk.

“The four broad areas of efficient modeling techniques are grouping, replicating portfolio, scenario reduction and statistical sampling.”

The financial services industry is presented with a dilemma: how to offer consumers products that serve their needs but are increasingly more complex, and also help management to accurately understand all risk exposures in real time. It is struggling with this major issue. A lack of robust and real-time modeling tools has contributed to management failure to understand and measure risk exposure, a reason for the financial crisis we recently experienced.

The root of the dilemma is that modeling the entire population, which is feasible with sophisticated computing technology, simply does not offer a practical response time. Many financial models are stochastic in nature and, even with the most advanced computing systems, can take days to produce results. Simplifying the modeling process is generally not an option because the financial products are complex in design and require sophisticated models.

In response to this dilemma, actuaries have developed several efficient modeling techniques.

Current Efficient Modeling Techniques

Other efficient modeling techniques besides grouping methods have been developed and are currently in use. The four broad areas of efficient modeling techniques are:

- **Grouping.** Group policies into risk classes or model cells, and every policy in a given risk class is replaced with an “average” policy.
- **Replicating portfolio.** Replace the projected liability cash flows with a representative portfolio of assets with similar cash-flow characteristics. Once these assets have been selected, the liability projections are replaced with asset cash-flow projections using the same number of stochastic scenarios.
- **Scenario reduction.** Use mathematical algorithms to select a few of the thousands of stochastic scenarios that could reproduce the tail quantiles of interest.
- **Statistical sampling.** Analyze a random sample of contracts to make inferences about changes in the risk metric for the in-force population.

Grouping techniques are the oldest form of efficient modeling techniques. They are very intuitive, adapt well to any actuarial modeling system and produce good estimates in predicting mean values of a risk metric. Replicating portfolio techniques have gained in popularity because of the investment-oriented nature of newer insurance products, and they are extremely efficient in capturing market and interest rate changes. Scenario reduction techniques focus on identifying scenarios that reproduce tail values of the risk metric, which is typically what is needed for capital and reserving requirements.

Use of statistical sampling is the most recent efficient modeling technique, even though statistical sampling is a well-established field in statistics and heavily utilized outside financial modeling. Why has it taken this long?

Statistical Sampling and Financial Product Modeling

Statistical sampling techniques are not used in financial product modeling for several reasons:

- A sample of the population is not needed since companies have easy access to the underlying population of in-force policies typically found in a valuation or administrative database.
- Concerns exist about introducing sampling error if statistical sampling techniques are used where exact and accurate values are needed for the financial metrics being calculated. Such instances could include minimum reserves or the various Greek measures used in hedging models.

- Deterministic (versus stochastic) calculations allow for individual policy calculations to be easily and quickly done on the entire in-force population.
- The advent of high-speed computers and sophisticated processing techniques enables efficient performance of complex stochastic calculations.

Since sampling error can be eliminated by accessing the entire population, the two worlds of financial modeling and statistical inference have never had to be combined.

Replicated stratified sampling (RSS), a modification of traditional stratified sampling, is an efficient modeling technique that attempts to overcome some of the drawbacks of sampling. Before we get into the details of the RSS technique, we need to analyze what would be the ideal attributes of an efficient modeling system. These are:

- Speed and accuracy
- Easy to understand, implement and maintain
- Error in estimation can be measured and controlled
- Well-established mathematical basis as to why the technique works
- Consistent and robust results that do not require frequent recalibration
- The ability to capture both changes in market assumptions (e.g., equity and interest rate changes) as well as actuarial changes (e.g., changes in policyholder behavior)

With these attributes in mind, let us get into more detail about the RSS technique.

RSS Technique Overview

Towers Watson's RSS technique uses a stratified sampling statistical approach to estimate the population distribution. Stratified sampling provides unbiased and consistent estimates of population parameters, although it alone does not eliminate the risk of sampling error. Sampling error can be reduced by increasing the number of strata or the sample size, but that would add to processing time.

To eliminate sampling error, RSS complements the stratified sampling with replicated samples. This is what makes the RSS technique unique enough to have led Towers Watson to seek a patent on the entire modeling process. RSS is generally not a technique that is used in statistical inference because of the cost and effort required to draw just a single sample. The area of statistics most similar to replicated sampling is "bootstrapping," the creation of replicated samples by drawing with replacement from a single sample. RSS differs from bootstrapping because each replication draws a new sample from the population. RSS is certainly not practical for most disciplines that use sampling techniques to make inferences about the population. But in financial product modeling, the industry's ease of access to an entire population of data makes obtaining repeated samples a simple and low-cost process.

Statistical theory is not built around speed of convergence of repeated stratified samples drawn from a population. However, classical convergence theorems like the Law of Large Numbers or the Central Limit Theorem naturally extend to multiple samples.

The steps in an RSS application are:

1. Determine the risk metric of interest (e.g., embedded value for a block of business) and the number of sensitivity tests to be undertaken.
2. Calibrate by determining the optimal number of risk classes, the optimal sample size from each risk class and the optimal number of replications based on the risk metric being estimated, and the size and characteristics of the population. RSS is designed to produce a convergence of results at a certain confidence level in the shortest model run time.
3. Draw a stratified sample to produce estimates of changes in the risk metric of interest under different sensitivities. Use an automated process, and then run a baseline scenario and each of the sensitivity tests for the sample using your company's existing actuarial modeling software. This process will be repeated for the number of replications required.
4. Feed the actuarial modeling software results for each of the samples into the RSS processor. Results are combined and the estimate of the change in the risk metric of interest for each of the sensitivity tests is determined.

"Use of statistical sampling is the most recent of efficient modeling techniques, even though statistical sampling is a well-established field in statistics and heavily utilized outside financial modeling."

> "Ideal efficient modeling attributes include speed and accuracy, ease in understanding, a well-established mathematical basis, and consistent and robust results."

A schematic of the process is shown in the figure below.

The speed of convergence is tied to how the information from the replicated samples is combined to estimate the population distribution. There are three steps in this process:

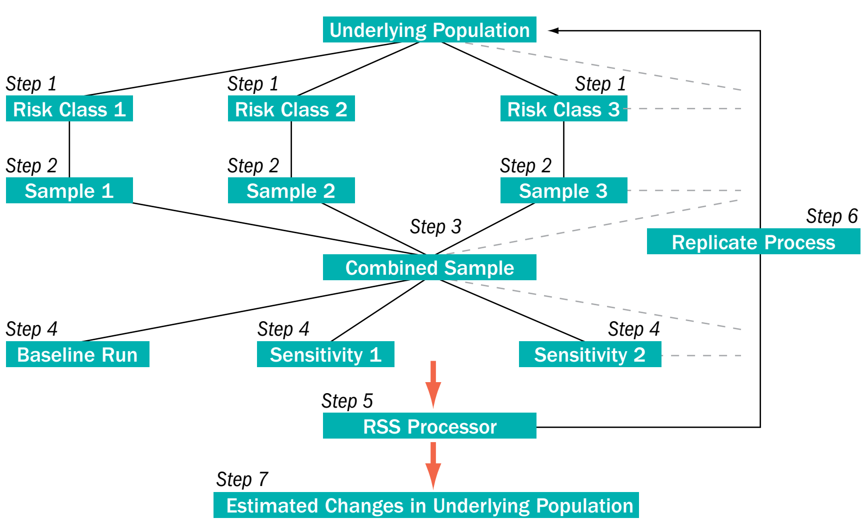
1. Eliminate unrepresentative samples or outliers.
2. Create cumulative estimates of the population distribution from the replications through different techniques to combine the remaining samples.
3. Determine the number of replications required to achieve the *predetermined* estimation error level.

RSS Technique Attributes

The RSS technique shares several of the ideal attributes for an efficient modeling technique. Based on extensive testing of RSS, we have observed the following:

- **Benefits of scale.** The RSS speed of convergence is independent of the underlying population size. For example, it takes about the same time to model 100,000 policies as it would 500,000 policies using the RSS technique, while the full population run would take about five times longer.
- **Predetermined level of accuracy.** The RSS parameters (risk classes, sample size and replications) can be calibrated to achieve a predetermined level of accuracy for estimating changes in a risk metric. This is a positive attribute in making this technique acceptable to auditors and regulators to determine changes in capital and reserving requirements on a more frequent basis.
- **Adaptation to existing modeling system.** RSS works directly with a company's existing financial modeling systems.
- **Consistent and robust results with limited maintenance required.** RSS is a robust management tool that produces consistent results. Well-established statistical sampling principles are the reason for consistency, while robustness is achieved by generating new samples from current in-force data each time the technique is applied. RSS parameters do not have to be recalibrated on a regular basis, thus requiring little maintenance for the technique.
- **Broad applicability.** The RSS technique does not attempt to simplify or approximate the underlying model complexity. By applying the existing model to the randomly selected stratified samples, it preserves the model complexity, while the replication process ensures convergence to a desired level of accuracy.
- **Easy to understand.** The technique is very intuitive, as statistical sampling is used in every other discipline and in market research surveys. This makes it easy to communicate to management, regulators and other stakeholders.
- **Mathematically rigorous.** RSS has a well-established theoretical foundation based on fundamental statistical sampling principles such as the Law of Large Numbers and the Central Limit Theorem.

RSS process



Speed of Convergence of RSS

The recent regulatory requirement for determining reserves for variable annuities with guarantees (the Variable Annuity Commissioners' Annuity Reserve Valuation Method or VA-CARVM) is possibly one of the most complex and computer-intensive calculations that insurance companies have to perform. Besides being a stochastic calculation, cash flows and accumulated statutory surplus have to be projected for each market scenario, and the greatest present value of accumulated deficiencies over each projection year has to be determined. The scenarios are then ranked, and the average of the largest 30% of the ranked values is calculated to determine the VA-CARVM reserve. Even with the most sophisticated hardware support, a company with a sizable block of variable annuity contracts could need several days to complete a single VA-CARVM run.

While the calculation of the base reserve level is laborious, what is possibly even more important to analyze is how VA-CARVM results vary as the liability and investment assumptions change. The model not only reflects all contractual guarantees, but also incorporates assumptions including mortality, lapses, policyholder behavior, partial withdrawals, fees and expenses, fund projection assumptions, reinsurance cash flows and hedging instrument cash flows. Management can fully understand and manage the risk exposure of these complex insurance products only if they are able to vary and stress-test these model assumptions to assess their impact. Currently, this can only be done on a limited basis, if at all, because of the sheer time constraints involved with current modeling tools.

RSS was implemented as part of a pilot study for a major insurer to estimate changes in VA-CARVM results based on different sensitivity tests. It estimated the impact on the VA-CARVM reserves of an immediate 15% and 35% drop in equity funds. The pilot study covered three legal entities, and analyzed results both before and after reinsurance. So in total, there were 12 estimated values produced by the RSS technique.

The results showed that for all 12 cases, in order to maintain accuracy levels within 1% of the true population change, the technique required calculations for between 1% and 2% of total in-force policies. This represents between 98% and 99% improvement in processing time (based on the total number of policies sampled relative to the in-force population) compared to a full population analysis. This dramatic reduction in run times will redefine the way management will be able to analyze business and make decisions.

RSS and Hedging

For hedging applications, RSS is more than just a technique to reduce processing time. In comparison to a full run, it is possible to increase the number of stochastic simulations for each sample and develop better estimates of the relevant Greek measures used in hedging, resulting in more accurate and cost-effective hedging decisions.

Where Do We Go From Here?

The RSS technique provides all kinds of exciting opportunities in financial modeling to provide timely and accurate management information. RSS does not add computer hardware or simplify the modeling process. Rather, it brings in well-established statistical inference techniques using a combination of stratified sampling and sample replication. And yet, it can be further developed and enhanced in two broad areas:

- Theoretical enhancements
- Application enhancements

Theoretical enhancements include the following:

- Optimizing the combination of number of risk classes, sample size per risk class and number of replications to maximize the speed of convergence
- Using techniques in numerical analysis to determine the best-stopping-time algorithm for the number of replications to achieve the prescribed level of accuracy
- Using advanced statistical inference techniques like bootstrapping and trimmed means to further refine RSS estimates

The potential application enhancements are numerous and extend beyond insurance applications to other kinds of financial modeling: valuing mortgage-backed securities, options and derivatives. Any modeling technique that uses grouping to speed up processing time could be enhanced to use seriatim modeling coupled with the RSS technique. This would better measure the tail risks without sacrificing processing speed.

In general, the RSS technique will benefit any modeling process that has easy access to the underlying population being analyzed and involves significant processing time. While the entire financial services industry (including insurance, banking, hedge funds and mutual funds) fits these criteria, there could also be creative applications of the RSS technique to economic forecasting, econometric modeling, projection of health care costs under different economic scenarios, and any complex modeling that uses large amounts of data and complex underlying design and assumptions.

For comments or questions, call or e-mail Jay Vadiveloo at +1 860 843 7073, jay.vadiveloo@towerswatson.com.

Jay Vadiveloo
Specializes in applied actuarial research projects with academics and actuarial professionals.
Towers Watson, Hartford