Extrapolation Sampling — Model-Based Sampling in Litigation Support

Timothy Wyant, Independent Consultant

The applied statistician typically encounters problems involving some quirk that limits the applicability of standard results. Effective statistical consulting requires that the statistician address such quirks in a manner that is statistically reasonable, while meeting client needs in terms of cost, relevance, and timeliness.

One such quirk arises in toxic tort and product liability litigations. Such litigations — those involving asbestos, Agent Orange and other dioxin exposures, the Dalkon Shield contraceptive device, and DPT vaccine come to mind — typically begin with a few individuals suing the manufacturer. If some of these suits are successful (or appear to have that potential), litigations may snowball. Eventually, cases of many individuals may be consolidated into a class action.

In some instances, so many litigants come forward, each with sufficiently high probability of courtroom success, that the manufacturer declares bankruptcy. Such was the case with Johns Manville (asbestos) and A.H. Robins (the Dalkon Shield). In such cases, people with claims against the manufacturer become a special sort of creditor. Although such people "compete" in the legal arena with other creditors, their claims have no determinate value at the onset of bankruptcy proceedings. In either type of claim consolidation — bankruptcy or class action — deriving a reasonable estimate of the total dollar value of outstanding claims becomes a paramount issue.

Extrapolation Sampling

The statistical quirk that may face the statistician retained as an expert in consolidated proceedings is the need to sample from two populations simultaneously. Claims resolved prior to the consolidated proceeding provide a calibration population. In each such claim, the legal system assigned a value to the claim that in some sense weighted both the claim’s scientific merit (did the product “cause” the claimant’s injury) and the claim’s social merit (how much is the injury "worth").

A sample of resolved claims provides a database from which one can estimate not only the historic frequency of claim characteristics, but also the nature of the relationship between a claim’s characteristics and its likely value in the eyes of the legal system.

The consolidated unresolved claims constitute a target population. A sample drawn from this population provides a basis for estimating the distribution of claim characteristics for claims yet to be compensated (or dismissed). To the extent that the two populations — the calibration and the target — differ primarily in the frequency with which different claim characteristics occur, and not in the way in which the characteristics determine value, one can use the two samples to estimate the total value of unresolved claims.

Urgency (the injured need compensation, the manufacturers need their legal problems resolved, the Courts need to clear their dockets) may force the statistician to design and administer samples of both calibration and target populations simultaneously. Such extrapolation samples make the statistician parcel out sample units in the two populations based on anticipated characteristics of both. Standard sampling methods do not specifically address this situation.

Example — The Dalkon Shield

The Dalkon Shield was an intrauterine contraceptive device (IUD) sold by the A.H. Robins company in the 1970’s. A number of studies implicated the Dalkon Shield as a cause of pelvic infection, infertility, birth defects, and sometimes deaths. Though some controversy still exists as to the precise culpability of the IUD in causing such injuries, evidence accumulated in the 1970’s resulted in about $400 million being paid to resolve approximately 10,000 claims from 1975 to 1985.

In 1985, Robins sought protection from the growing number of claimant "creditors" by filing for Chapter 11 Bankruptcy. Advertisements required by the bankruptcy court elicited an additional 330,000
claims. As part of the bankruptcy proceeding, a group of statistical experts designed and administered extrapolation samples of resolved and pending claims. (The major parties to the proceeding — stockholders, creditors, etc. — each chose experts to serve with the group. I served with this group on behalf of the Dalkon Shield Claimants’ Committee.)

Experts for the major parties (all of whom used SAS™) presented analyses of the sample data in a Court hearing in November, 1987. As a result of that hearing, a $2.3 billion trust fund to compensate claimants was established in March, 1988.

Addressing Statistical Quirks

The Dalkon Shield example provides an instance where the need for extrapolation sampling arises, and where the manner in which the sampling is conducted can affect litigation involving many people and enormous sums of money. The consulting statistician has a variety of ways to address quirks such as the need for extrapolation sampling:

- Ignore them
- Find an approach in the existing literature
- Derive some useful formulas
- Use simulations or numerical approximations

Although ignoring the quirks seems at first glance to avoid the statistician’s main task, it may be an appropriate strategy. For example, one problem that arises in extrapolation sampling is that of cross-allocation. Given N sample units, how many should go to the target population, and how many to the calibration? Optimal allocation may be of primary interest to the statistical researcher. But the consulting statistician remains constantly aware of the the limitations imposed by the political and administrative context in which statistical decisions must be made. Sometimes, the difference between "optimal" and "sub-optimal" sample allocations is inconsequential in the context of the problem. Or, the situation may preclude obtaining the information necessary to determine "what is optimal".

In either case, the statistician may have more important battles to fight. He or she can't just "keep statistics and be a jerk".

Consulting statisticians generally check the standard literature for approaches to quirks. But such "literature surveys" tend to play a lesser role for the statistical practitioner than for the researcher. The practitioner must provide a reasonable solution at reasonable cost, and on time. Extending theory is not the immediate task. If the statistician projects little chance of finding an existing solution in time, yet feels secure that a reasonable simulation approach can be developed, concentration on the latter direction is warranted.

Of course, the statistician has to ask, "When is a quirk not a quirk?" If the seeming "quirk" arises again and again, and appears to be in fact fundamental to a class of problems, more extensive investigation is warranted.

Simulations and other numerical approaches provide the statistician with powerful and flexible tools given today’s computers. I ran many simulations using SAS™ Proc Matriz (now IML) in examining various extrapolation sampling approaches for the Dalkon Shield litigation. But simulation has not eliminated the need for the statistician to attempt some theoretical derivations. The attempt itself focuses the statistician on critical issues. And a simple, useful, and relevant formula (even if over-simplistic) can provide the critical tool for devising and achieving a statistical approach that solves a client’s problem.

For one thing, heuristics and rules of thumb leap more readily from a formula than from a collection of simulation runs. Decisions on sample allocations or similar statistical matters are often made in meetings or negotiations in which administrative or political issues loom larger than statistical ones. Sound, understandable heuristics and rules of thumb often prevent divergence from a reasonable statistical solution in such situations.

... all he ever did was keep statistics and be a jerk.

1043
In addition, there is always a need for "back of the envelope" calculations in such meetings and negotiations. A quick answer at the right time can be all-powerful in shifting the course of events. A few formulas can turn the tide. None of this discussion is intended to shortchange the power and usefulness of simulation approaches. But to be effective, the consulting statistician, in my experience, must do more than return to the computer for yet one more "what if" run.

Some Results for Extrapolation Sampling

Some derivations relevant to the particular quirks inherent in extrapolation sampling follow. They should not be taken as definitive theory, but as examples of the kind of calculations made by one consulting statistician in a particular situation.

Suppose a natural stratification exists in the calibration and target populations. For simplicity of exposition, assume two strata.

The calibration population:

<table>
<thead>
<tr>
<th>Stratum index</th>
<th>i, i=1,2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean claim value</td>
<td>μ_i</td>
</tr>
<tr>
<td>Variance of values</td>
<td>σ_i^2</td>
</tr>
<tr>
<td>Calibration sample</td>
<td>m units</td>
</tr>
<tr>
<td>Calibration allocation —</td>
<td>m_i units, stratum i</td>
</tr>
</tbody>
</table>

Target population:

<table>
<thead>
<tr>
<th>Population size</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportions by stratum</td>
<td>P_i</td>
</tr>
<tr>
<td>where p_1 (proportion in stratum 1)</td>
<td>= (1 - p_2)</td>
</tr>
</tbody>
</table>

The natural estimator for the total value of claims in the target population is

\[ \hat{T} = \hat{T}_1 + \hat{T}_2 = N(\hat{p}_1\mu_1 + \hat{p}_2\mu_2) \]

and the variance of \( \hat{T} \) is

\[ V(\hat{T}|\hat{p}_1,\hat{p}_2) = N^2\sigma_1^2/m_1 + N^2\sigma_2^2/m_2 \]

if \( \sigma_2 = c\sigma_1 \), the standard error of \( \hat{T} \) is

\[ \sqrt{N\sigma_1^2\hat{p}_1^2/n_1 + \hat{p}_2^2c^2/n_2} \]

Minimizing the standard error with respect to n gives (with the constraint \( m_1 + m_2 = m \)) gives optimal allocation occurring at

\[ m_2/m_1 = c(p_2/p_1) \]

Note that if the strata proportions \( p_1 \) and \( p_2 \) are identical in both the calibration and target populations, a random sample of the calibration population (disregarding strata) would tend to give an after-the-fact ratio of

\[ m_1/m_2 = p_2/p_1. \]

So the extent we would diverge from a random sample to achieve optimal allocation in this situation is driven by \( c \), the ratio of standard deviations.

The optimal allocation formula in this situation is (not surprisingly) basically "Neyman allocation" from standard sampling theory. But the formula here involves proportions from the target. Also, this informal derivation is based on a superpopulation model rather than a sampling distribution.

I prefer to base applied sampling formulas on probability models, rather than on classical sampling distributions. Beyond the conceptual appeal of the model-based approach, as documented in much current literature, the mathematics tend to make derivations simpler. In addition, model-based arguments often reinforce the value of "judgment" samples that extend pure random samples to provide "insurance" against unanticipated anomalies such as isolated, huge values in the population. (Of course, in the extrapolation example, the calibration sample is basically a purely model-based experimental design. It is not a finite population sample in the usual sense, unless there is a need to pin

1044
down historic distributions of claim characteristics.)

Usually, even within a stratum of the target or calibration population, statistical modeling breaks the population down into numerous cells, each cell containing homogenous units. The estimated target total is then:

\[ \hat{T} = \hat{\mu}_1 \hat{p}_1 + \hat{\mu}_2 \hat{p}_2 + \cdots + \hat{\mu}_k \hat{p}_k \]

Typically, although the \( \hat{\mu}'s \) are independent of the \( \hat{p}'s \), the vectors \( \hat{\mu} \) and \( \hat{p} \) have internal dependencies. If the covariance matrices of \( \hat{\mu} \) and \( \hat{p} \) are respectively, \( \varphi \) and \( \tau \), the formula

\[ V(\hat{T}) = \hat{\mu}' \varphi \hat{\mu} + \hat{p}' \varphi \hat{p} + (\text{vec } \tau)'(\text{vec } \varphi) \]

is useful.

For a simpler example that provides useful insight, consider the following:

<table>
<thead>
<tr>
<th>The calibration population:</th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>( m )</th>
<th>( \hat{\mu} )</th>
<th>( \frac{\sigma^2}{m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimate of the mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>( \delta_1 = \frac{\sigma}{\mu} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The target population:</th>
<th>( p )</th>
<th>( p(1-p) )</th>
<th>( n )</th>
<th>( \hat{p} )</th>
<th>( p(1-p)/n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propportion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td></td>
<td>( p )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
<td>( n )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimate of the mean</td>
<td></td>
<td>( \hat{p} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. error</td>
<td></td>
<td>( p(1-p)/n )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>( \delta_2 = \sqrt{p(1-p)/p} = \sqrt{(1-p)/p} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The constraint:

\[ m + n = N \]

The target total:

\[ \begin{align*}
\text{Actual} & : p \mu \\
\text{Estimate} & : \hat{p} \hat{\mu} \\
\text{Variance of the estimate} & : \frac{\sigma^2 p(1-p) + \mu^2 p(1-p) + p^2 \sigma^2}{m/n} \\
\end{align*} \]

Minimizing variance subject to the constraint (and ignoring finite population correction factors as well as replacing \( n-1 \) with \( n \) to simplify certain standard errors) results in the following relationship between \( m \) and \( n \) at the optimal cross-allocation:

\[ m^2/\delta_1^2 + m \approx n^2/\delta_2^2 + n \]

If the samples are large enough that the squares dominate, one gets

\[ m^2/n^2 \approx \delta_1^2/\delta_2^2 \]

It is easy to solve for the optimum ratio of \( m \) to \( n \) numerically — it is the solution to a quadratic — but more insight is probably gained by preserving the symmetry of the preceding equations. Among other things, we see that:

1) Equal \( \delta's \) imply equal samples
2) Larger \( \delta \) gets larger sample, in direct ratio
3) As \( p \) gets small (\( \delta_2 \) gets big), \( n \) gets large relative to \( m \)
4) As \( \mu \) (which can be thought of as the difference between the two target categories in mean value) gets large (\( \delta_1 \) gets small), \( n \) gets large relative to \( m \).

To the statistician with sound common sense, and unwavering confidence in his or her instincts, the preceding rules of thumb may seem obvious. But my experience indicates that some of the above results may be counterintuitive even to experienced practitioners — especially the fact that the target sample can have such
importance. (For $p = .1$, for example, $\delta_2$ equals 3.0.)

Modeling in the calibration population can easily reduce residual calibration error to the point where more intense sampling of the target is indicated.)

Staying in Context

To return to the initial points, the statistician should use heuristics such as the preceding only with full appreciation of the political, administrative, and ethical context of the problem. Cost always rears as a critical issue, as just one example, and costs of sampling the target population may differ markedly from costs of sampling the calibration population. Furthermore, "costs" should be considered in the largest sense. In litigation settings, sampling of a target population can impose significant burdens of time, expense, effort, and legal complication on people in the population. Even though such cost may not appear in the statistician's budget, he or she should remain sensitive to opportunities to reduce such cost either through administration of the sample to reduce individual burdens, or through smaller samples. The lure of statistical calculations should not distract us from doing what's right.

Timothy Wyant
Decipher
7202 13th Avenue
Takoma Park, MD 20912