Loss Reserve Estimates: A Statistical Approach for Determining "Reasonableness"

Mark R. Shapland, FCAS, ASA, MAAA

LOSS RESERVE ESTIMATES:
A STATISTICAL APPROACH FOR DETERMINING "REASONABLENESS"

ABSTRACT

With the NAIC's adoption of the Accounting Practices and Procedures Manual, the statutory accounting practices for the P&C insurance industry have now been codified in a series of Statements of Statutory Accounting Principles (SSAP's). Within the SSAP's, various terms such as "Management's Best Estimate," "Ranges of Reserve Estimates" and "Best Estimate by Line" have been defined. In addition, the Actuarial Standard of Practice (ASOP) No. 36, adopted by the Actuarial Standards Board in March 2000, provides definitions for terms such as "Risk Margin," "Determination of Reasonable Provision" and "Range of Reasonable Reserve Estimates." While they are both well designed and a definite improvement, these new principles and standards of practice provide only broad guidance to the actuary on what is "reasonable." This broad guidance is based on the principle that "reasonable" assumptions and models lead to "reasonable" estimates. Unfortunately, this broad guidance can leave the low end of a range of "reasonable" reserves open to an interpretation which could lead to unintended consequences in practice. This paper will review some current actuarial practices and examine how they relate to the question of what is "reasonable" from a statistical perspective. Moreover, it will review and further develop some statistical concepts and principles that actuaries can add to their repertoire when developing ranges of liability estimates and then evaluating the "reasonableness" of management's best estimate of reserves within those ranges. It is hoped that the Actuarial Standards Board and others will consider adopting a more definitive definition of "reasonableness" in order to help avoid the unintended consequences of allowing the reserves to get "too low" in practice.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the many thoughtful comments from Oakley E. (Lee) Van Slyke on early drafts of the paper, many of which formed the basis for the section on Practical Applications. Other individuals that provided valuable contributions that improved the quality of the paper include Glen Barnett, Robert Shapland, Dean Westpfahl, Claus Metzner, Kevin Madigan and some members of the Committee on Reserves. The opinions and conclusions expressed, as well as any errors or omissions that remain, are solely the responsibility of the author.

LOSS RESERVE ESTIMATES:
A STATISTICAL APPROACH FOR DETERMINING "REASONABLENESS"

"The work of science is to substitute facts for appearances and demonstrations for impressions."

- John Ruskin

1. INTRODUCTION

The work of the actuary in developing loss liability estimates is a relatively scientific process, yet it is guided by some very subjective terms like "reasonable." The purpose of this paper is to develop a more definitive framework for the term "reasonable reserve estimate" based on statistical principles that the actuary can use when developing ranges of liability estimates and then evaluating management's best estimate of reserves within those ranges. Along the way, it will show how the current broad guidelines could be "misinterpreted." The first step in developing any set of principles is to start with a solid foundation, so this paper will begin by reviewing some "codified" terms and their definitions, defining some terms for use in this paper, and reviewing various statistical measures of risk. Next, it will examine some of the current practices for determining "reasonableness" and suggest a framework for defining "reasonableness" more precisely. Then various risk concepts will be reviewed and, more importantly, how they relate to the question of "reasonableness." Once all of these definitions and concepts are outlined, some general models for calculating ranges will be examined and some practical applications will be reviewed to see how these principles might be applied in practice. Finally, the paper will conclude by suggesting some areas for further research and an overview of the findings.

2. DEFINITION OF TERMS

Throughout this paper, unless noted otherwise, loss reserves are intended to include both loss and allocated loss adjustment expense reserves.[1] The SSAP's and ASOP No. 36 contain some definitions related to the term "reasonable." From the SSAP's we have the following:

> **Management's Best Estimate** – Management's best estimate of its liabilities is to be recorded. This amount may or may not equal the actuary's best estimate.

[1] While many of the principles and analyses in this paper might also apply to unallocated loss adjustment expense reserves, they have been kept outside the scope of the discussion.

Ranges of Reserve Estimates – When management believes no estimate is better than any other within the range, management should accrue the midpoint.[2] If a range can't be determined, management should accrue the best estimate. Management's range may or may not equal the actuary's range.

Best Estimate by Line – Management should accrue its best estimate by line of business and in the aggregate. Recognized redundancies in one line of business cannot be used to offset recognized deficiencies in another line of business.[3]

From ASOP No. 36, we have the following:

Risk Margin – An amount that recognizes uncertainty; also known as a *provision for uncertainty*.

Determination of Reasonable Provision – When the stated reserve amount is within the actuary's range of reasonable reserve estimates, the actuary should issue a statement of actuarial opinion that the stated reserve amount makes a reasonable provision for the liabilities.

Range of Reasonable Reserve Estimates – The actuary may determine a range of reasonable reserve estimates that reflects the uncertainties associated with analyzing the reserves. A *range of reasonable estimates* is a range of estimates that could be produced by appropriate actuarial methods or alternative sets of assumptions that the actuary judges to be reasonable. The actuary may include risk margins in a range of reasonable estimates, but is not required to do so. A range of reasonable reserves, however, usually does not represent the range of all possible outcomes.

These definitions provide the actuary with only broad guidance on what is "reasonable." For example, is any reserve "reasonable," as long as it falls within any range of reserves based on any set of assumptions and models as long as those assumptions and models[4] are deemed reasonable by a competent actuary?[5] Of course any set of assumptions and models deemed reasonable by the actuary must also stand up to peer review scrutiny, but does this imply that two actuaries can create a quorum for determining reasonableness? Should the actuary's judgment about the assumptions

[2] Statutory guidance was silent on this point before the SSAP's, however when no estimate is better than any other within a range GAAP accounting standards state that the lowest estimate in the range should be accrued.

[3] Definitions of GAAP accounting terms may also be useful, but differences between GAAP and Statutory accounting principles are beyond the scope of this paper.

[4] Throughout this paper, the terms "method" and "model" are used interchangeably. However, preference is given to the term "model" to emphasize the need to think about actuarial reserve calculations as a model of the underlying process that is generating the claims rather than simply as a process for making calculations.

[5] A competent actuary could be defined as someone who is trained in the application of generally accepted actuarial methods and assumptions, but, interestingly, this creates a circular logic for determining "reasonableness."

and models be the only criteria for reasonableness, or do we need additional context to put these questions in perspective? In essence, these terms seem to imply a "reasonable person" standard much like you would find in a legal context.

This paper will argue that a statistical approach *should be added to* the "reasonable person" standard so that a more informed judgment, by both actuaries and users of the actuarial work product, about whether a stated reserve is "reasonable" or not can be made. In order to develop this approach, some basic definitions are offered. Consider the following:

- **Reserve** – an amount carried in the liability section of a risk-bearing entity's balance sheet for claims incurred prior to a given accounting date.
- **Liability** – the actual amount that is owed and will ultimately be paid by a risk-bearing entity for claims incurred prior to a given accounting date.[6]
- **Loss Liability** – the *value* of all estimated future claim payments.
- **Risk** *(from the "risk-bearers" point of view)* – the uncertainty[7] (deviations from expected) in both timing and amount of the future claim payment stream.[8]

3. MEASURES OF RISK

From statistics, actuaries often use a variety of measures that help define risk. These measures could include: variance, standard deviation, kurtosis, average absolute deviation, Value at Risk, Tail Value at Risk, *etc.* which are measures of dispersion. Other measures that help to define aspects of the distribution that might be useful in determining "reasonableness" could include: mean, mode, median, *etc.* The choice for measure of risk will also be important when considering the "reasonableness" and "materiality" of the reserves in relation to the capital position.

[6] The Statement of Principles Regarding Property and Casualty Loss and Loss Adjustment Expense Reserves define **Loss Reserve** as "a provision for its related liability." While reserves and liabilities are sometimes used interchangeably, they are given separate definitions in this paper, and used differently throughout, to help clarify the concepts discussed.

[7] In section 3.6.1 of ASOP No. 36, sources of uncertainty are described and include the following: random chance; erratic historical development data; past and future changes in operations; changes in the external environment; changes in data, trends, development patterns and payment patterns; the emergence of unusual types or sizes of claims; shifts in types of reported claims or reporting patterns; and changes in claim frequency or severity.

[8] If the loss liabilities are discounted, this would add an additional source of uncertainty to the expected value of the future payment stream. For purposes of the paper, "interest rate risk" will be ignored and liabilities are assumed to be undiscounted.

For insurance risks, actuaries often discuss the need to consider both "process" and "parameter" risk since both of these are part of the risk-bearer's burden.

> **Process Risk** – the randomness of future outcomes given a *known* distribution of possible outcomes.
>
> **Parameter Risk** – the potential error in the estimated parameters used to describe the distribution of possible outcomes, assuming the process generating the outcomes is *known*.

Statistically, both of these can be measured and used to calculate the distribution of possible outcomes. However, these calculations assume that the process that is generating the outcomes is known and the only requirement is to estimate the parameters of that process. Thus, for the purpose of describing a range of possible liability outcomes an additional type of risk could be defined as:

> **Model Risk** – a measure of the effect (*i.e.*, forecast error) given the model ("process") used to estimate the distribution of possible outcomes is incorrect or incomplete.[9]

While some models will allow us to capture the most salient characteristics of a set of data, the fact remains that no model is ever completely "correct" or "complete." [10]

Consider an example from gambling. In the game of Roulette, the casino knows exactly what the distribution of numbers and colors are on the roulette wheel, so determining the payouts (odds) involves only the process risk for the game since the parameters are certain (assuming a fair game). If we were to change the game so that the casino did not know the exact distribution of numbers and colors, then the casino could only determine appropriate payouts by continuous sampling of the outcomes.[11] In this case the casino, like the insurance risk-bearer, does not know the exact

[9] In common vernacular, actuaries and statisticians generally use the term "parameter risk" to include both parameter risk and model risk as defined in this paper. The two risks are separated here in order to distinguish the portion that is readily measurable (assuming a given model) from the portion that is not. They are also separated to emphasize the fact that all models used by actuaries make assumptions about the claim process that are critical to the estimates they produce.

[10] Model Risk could also be further divided into: i) model selection uncertainty, and ii) model specification risk. **Model Selection Uncertainty** is where you choose one model from a set of candidate models and forecast as if the chosen model was the "correct" one, when in fact there may be a variety of quite plausible models. **Model Misspecification Risk** is the contribution to forecast error from the fact that none of the candidate models is actually correct.

[11] If the numbers and colors could also change over time, this would make the example more "real" in terms of its applicability to insurance, but the point about "process" and "parameter" risk does not change.

parameters of the game, so excluding the "parameter" risk from their payouts could lead to potential bankruptcy and, at a minimum, less profit than was expected.

So far this example implicitly assumes that the game still resembles a game of roulette, except that the numbers on the wheel are not known in advance. If we were to change the game even more such that the casino did not know how the outcomes are produced, then the casino would also be forced to guess at the process used to create outcomes when they are estimating the odds from their continuous sampling. The observed outcomes may resemble the outcomes from one or more mathematical distributions, which can be used to estimate the parameters, but the actual process that is generating outcomes is still unknown. Again, the casino, like the insurance risk-bearer, would need to add in an additional "risk load" in order to include "model" risk and be properly compensated.[12]

Returning to the insurance world, if there were no risk there would be no need for insurance. Even if there were no parameter or model risk, the insurance risk-bearer would still have some chance of insolvency. Failing to recognize parameter and model risk increases the danger of insolvency.

Before moving on to look at how these various types of risk relate to the reasonableness of reserves, note that standard statistical techniques (and terminology) are already available and, hence, do not need to be reinvented. For example, standard deviation and standard error have slightly different formulae and different meanings. Standard deviation describes a characteristic of a known distribution and includes only "process" risk, while standard error is an estimate of that characteristic of the underlying distribution based on sample data and includes both "process" and "parameter" risk. Unfortunately, calculating model risk may not be possible.[13] While model risk is implied with the common definition of parameter risk and, therefore, implied to be included in

[12] Returning to the earlier definition of **Loss Liabilities**, this analogy would imply that all 3 types of risk (*i.e.*, process, parameter and model risk) should be included as part of the calculated expected value. Alternatively, some or all of these types of risk could be included in **Risk Margin** as defined under ASOP No. 36.

[13] In fact, some sources of model uncertainty can be estimated in some circumstances. For example, if selecting from a sufficiently flexible group of models that the bulk of the information in the data about the future has been captured, then one may estimate *Model Selection Uncertainty* from the data. Of course there are other sources of model uncertainty (*e.g.*, *Model Misspecification Risk*) that must still be included judgmentally.

standard error calculations, it would seem more prudent to include a separate measure or loading for model risk.

4. HOW DO WE DEFINE REASONABLE?

In accordance with the SSAP's and ASOP's, the actuary must opine on the reasonableness of management's reserves, but the definition of what constitutes "reasonable" simply refers to a range. Thus, the actuary, and management, needs to consider a range of estimates, but there seems to be no defined process for determining what is "reasonable" within this range or whether the range itself is "reasonable."[14] A range of estimates, by itself, creates several problems that need to be overcome in order to determine "reasonableness":

- A range (arbitrary or otherwise) can be misleading to the layperson – it can give the impression that any number in that range is equally likely.[15]
- A range can also give a false sense of security to the layperson – it gives the impression that as long as the carried reserve is "within the range" anything is reasonable (and therefore in compliance) as long as it can be justified by other means.
- There is currently no specific guidance on how to consistently determine a range within the actuarial community (*e.g.*, +/- X%, +/- $X, using various estimates, *etc.*).[16]
- A range, in and of itself, therefore has insufficient meaning without some other context to help define it.

[14] One of the few places where more specific guidance is found is in SSAP 55, which states, in part, "when no estimate within a range is better than any other, the midpoint of the range should be accrued."

[15] Another gambling example might be useful here. Let's start with a game of chance where you wager a certain amount ($X) and in return you receive the dollar amount for the number that turns up on a roll of a fair die, plus $10. The range of possible outcomes is $11 to $16 and expected value is $13.50, so a fair wager is $13.50. A higher wager would be "good" for the house (they would gain over time), while a lower wager would be "bad" for the house (they would lose over time). Converting this to an insurance example, suppose an actuary was to tell management that the expected value of the liability estimate is $13.5 million, but the estimated range is $11 to $16 million and that each value in that range is equally likely to occur. What values in that range are "reasonable" for management to accrue?

[16] This statement does not imply that there has been no discussion about how to calculate ranges within the actuarial community. Quite the contrary, there have been numerous valuable contributions on this topic from authors of papers, editorials in the Actuarial Review, committee research, *etc.* The point is that the current guidelines simply say that a range *may* be used and that it *could* be calculated in a certain way, but the actuary is *not required* to create one.

Rather than simply saying that the actuary should calculate a range of liability estimates, it is the contention of this paper that actuaries should generally focus on calculating a distribution of possible outcomes such that the carried reserves would be sufficient to cover all future estimated claim payments at least X% of the time.[17] Once we define a "reasonable" range of the distribution based on probabilities, it can be translated into a range of liabilities that correspond to these probabilities. For example, telling management the liability estimate is $100 +/- $20 lacks sufficient meaning because of the reasons noted above. Contrast this to telling management the liability estimate shows they need $100 in order to have sufficient reserves at least 50% of the time and if they would like to increase the probability of having sufficient reserves to 75% they will need $120 in reserves. The second approach will be much more meaningful to management and other users of actuarial reports.[18]

Using a probability range to define a range of reasonable liabilities has the advantage of using the "risk" inherent in the data to define the range instead of a simple constant percentage. For example, if we were to define "reasonable" as a probability range of 50-75%, then the corresponding range of reasonable reserves might be $97-115 for a line of business with a relatively consistent claim payment stream, while the corresponding range of reasonable reserves might be $90-150 for a line of business with a more volatile claim payment stream. Contrast this with the common approach of using the estimated liabilities +/- X% for each line of business.

Table 1: Comparison of "Reasonable" Reserve Ranges by Method

	Relatively Stable LOB			More Volatile LOB		
Method	Low	Expected	High	Low	Expected	High
Expected +/- 20%	80	100	120	80	100	120
50th to 75th Percentile	97	100	115	90	100	150

---

17 Conversely, we could also define the probability range such that the carried reserves would be insufficient to cover all future expected claim payments at most (1-X)% of the time, although this approach has less intuitive appeal.

18 Continuing the simple example from Footnote 15, the actuary could advise management that reserves of a least $13.5 million was required in order to insure at least a 50% probability that they were sufficient and that $14.75 million would be required in order to insure at least a 75% probability that they would be sufficient. This would give the range some "reasonability" context that management could use to set reserves.

Using a probability range will also add context to other statistical measures. For example, as most liability distributions are skewed to the right, the mean will usually represent a value that is greater than the 50th percentile and can be used to help illustrate how the potential for the actual outcome to be worse than expected is greater than the potential to come in better than expected. Some actuaries have argued that the mode or the median could also be considered when describing what is "reasonable" in this context but, like the mean, discussing these as part of a probability range will complete and tie these various measures together.

The argument for using the mode as the "reasonable" reserve is that it has the highest probability of actually occurring. However, since liability distributions are usually skewed to the right (as illustrated in Graph 1), the mode would generally be less than the 50th percentile. In the context of liability distributions, the mode is the least desirable option for the low end of the range. Looking at the median (50th percentile), this would appear to be a logical low end to a range of "reasonable" reserves, but care must be exercised when selecting reserves by line of business compared to the aggregate reserves for all lines combined.

When reserves are selected by line of business and then simply added together to arrive at the total for all lines of business combined, this process is the same as assuming 100% correlation between lines. Generally, there is some level of independence between lines (*i.e.*, less than 100% correlation) which means that the total of selected individual medians (or modes) will be less than the median (or mode) of the aggregate for all lines combined. This concept is illustrated in Graph 2. Thus, if the median (or mode) is considered to be a "reasonable" low end for a range of reserves, then the medians (or modes) for the individual lines of business will need to be adjusted so they sum to the median (or mode) for the aggregate of all lines.[19] Using the expected value as the low end of the "reasonable" range of reserves will avoid this problem.[20]

---

[19] These "adjustments" by line also seem consistent with the SSAP definition of **Best Estimates by Line** which implies consistency by line and in the aggregate.

[20] While acknowledging the usefulness of mode and median, and that it is a matter for the industry to define, the remainder of the paper will focus on the estimated expected value as the low end of a reasonable range. Indeed, section 3.6.3 of ASOP No. 36 states, in part, that "[o]ther statistical values such as the mode... or the median...may not be appropriate measures... such as when the expected value estimates can be significantly greater than these other measures."

The concept of a "reserve margin" is often discussed in terms of a prudent excess over the expected value.[21] This definition of reserve margin is consistent with using probability ranges for reserves. For example, if the carried reserve is greater than the expected value, then the reserve margin is the difference between the carried reserve and the expected value.[22] However, nothing in this paper should be construed as implying that a carried reserve margin is not reasonable. On the contrary, recognition of "process," "parameter" and "model" risk would imply that having a reserve margin is not only reasonable, but prudent.

At the high end of the range, considerations related to materiality[23] of the reserve compared to the resulting surplus come into play. One way to tie materiality to the probability range of liabilities would be to use dynamic risk modeling to estimate how liability outcomes relate to the probabilities of insolvency. Consider the following tables:[24]

Table 2: Comparison of "Reasonable" Reserve Ranges with Probabilities of Insolvency

"Low" Reserve Risk

		Corresponding Surplus Depending on Situation[25]					
Loss Reserves		Situation A		Situation B		Situation C	
Amount	Prob. Of Sufficiency	Amount	Prob. Of Insolvency	Amount	Prob. Of Insolvency	Amount	Prob. Of Insolvency
100	50%	80	40%	120	15%	160	1%
110	75%	70	40%	110	15%	150	1%
120	90%	60	40%	100	15%	140	1%

---

21 Further distinctions between the "actual reserve margin" (determined after all claims incurred prior to a given accounting date are settled) and the "estimated reserve margin" (using the estimated expected value) could also be examined. However, since the scope of this paper involves estimated liabilities all references to reserve margins will imply estimated margins.

22 A negative reserve margin could also be defined as the difference between the carried reserve and the expected value.

23 ASOP No. 36 provides some guidance for evaluating Materiality – *In evaluating materiality within the context of a reserve opinion, the actuary should consider the purposes and intended uses for which the actuary prepared the statement of actuarial opinion.*

24 The numbers in these tables are purely hypothetical and designed for illustration purposes only.

25 If all else were equal, increasing the amount of the carried reserves will directly decrease the amount of surplus (Surplus = Assets – Liabilities) and the probability of insolvency wouldn't necessarily change. However, in practice, if the higher outcome actually occurs then the possibility that surplus could be eroded due to such things as insufficient rates, non-recoverable reinsurance, *etc.* would normally increase somewhat.

"Medium" Reserve Risk

		Corresponding Surplus Depending on Situation					
Loss Reserves		Situation A		Situation B		Situation C	
Amount	Prob. Of Sufficiency	Amount	Prob. Of Insolvency	Amount	Prob. Of Insolvency	Amount	Prob. Of Insolvency
100	50%	80	60%	120	40%	160	10%
120	75%	60	60%	100	40%	140	10%
140	90%	40	60%	80	40%	120	10%

"High" Reserve Risk

		Corresponding Surplus Depending on Situation					
Loss Reserves		Situation A[26]		Situation B		Situation C	
Amount	Prob. Of Sufficiency	Amount	Prob. Of Insolvency	Amount	Prob. Of Insolvency	Amount	Prob. Of Insolvency
100	50%	80	80%	120	50%	160	20%
150	75%	30	80%	70	50%	110	20%
200	90%	-20	80%	20	50%	60	20%

The relationship between reserve risk and the risk of insolvency is a complex issue. As illustrated in the tables above, there is a very strong interrelationship between how well an insurance enterprise is capitalized and the magnitude of the reserve risk. For example, if two companies have the same distribution of loss liabilities but Company A has only half the surplus as Company C, the range of reasonable reserves is the same for both companies even though the probability of insolvency for Company A is significantly higher. Alternatively, if Companies A and C both change their mix of business over time in such a manner that it increases their reserve risk (from, say, "low" to "high" risk), then the probability of insolvency will also increase for both but not to the same degree.

Of course, insolvency risk also depends on several other types of risk such as asset default risk, interest rate risk, reinsurance risk, catastrophe risk, *etc.* However, when all else is equal, the probability of insolvency decreases as the amount of surplus increases.

Interestingly, statistical analysis using ruin theory shows that pricing and reserving to the expected value every year, without any margin for risk loading, will eventually lead to insolvency with

[26] A particularly interesting example in these tables is the "high" risk situation A. In theory, the probability of insolvency wouldn't change if the company booked reserves of 200 instead of 100 even though the balance sheet would show negative surplus. Conversely, there would be pressure to book less than 100 to give the false impression that the company is more secure than it actually is.

probability of 100%.[27] This suggests that a prudent lower bound to the "reasonable" probability range for reserves should be *at least* the expected value, *if not higher.*

From the tables and discussion above, we might assume that a probability range from the expected value to 90% is "reasonable" so that every company can recognize the impact of reserve risk on their balance sheet and be properly compensated for risk in their pricing. Since market considerations related to "perceived" undercapitalization and the distortion of earnings that occur when a company strengthens their reserve position within this range put a natural economic barrier on the high end of the range, it seems like most regulators would be mainly concerned with keeping carried reserves above the low end of the range. Alternatively then, we might consider any carried reserves above the expected value to be "reasonable."[28]

Relating the concept of materiality to a probability range of liabilities could also prove useful in other related areas such as discussions of risk based capital and other solvency measures. For example, in a recent paper by Herbers [14] the viewpoints of different users of Statements of Actuarial Opinion are considered and a variety of sources for defining materiality are identified. Among all the different interests identified, the common goal among them is to make sure that risk is adequately disclosed. Conversely, the differences seem to be related to what level of risk needs to be disclosed. In order to satisfy the needs of all different users of actuarial opinions, the author suggests using the:

> **Principle of Greatest Common Interest** – the "largest amount" considered "reasonable" when a variety of constituents share a common goal or interest, such that all common goals or interests are met; and the
>
> **Principle of Least Common Interest** – the "smallest amount" considered "reasonable" when a variety of constituents share a common goal or interest, such that all common goals or interests are met.

---

[27] For example, see: Beard, Robert E., Pentikäinen, T. and Pesonen, E., "Risk Theory," Chapman & Hall, 1984, 3rd Edition.

[28] In order to help identify strong reserve positions, categories for subsets above the expected value could also be added. For example, the range from the expected value to 75% could be "reasonable and prudent" and the range above 75% could be "reasonable and conservative."

These principles could be used separately or in conjunction with each other, depending on which goal or interest is being considered. For example, at the low end of a probability range the principle of greatest common interest would imply using the highest minimum such that the requirements of all constituents are met. For materiality, the principle of least common interest would imply using the least amount of surplus change considered "reasonable" by all constituents concerned with materiality.

5. OTHER RISK CONCEPTS, ASSUMPTIONS AND CONSIDERATIONS

Before discussing the practical aspects of actually calculating these probability distributions, it is important to review other risk concepts, assumptions and considerations that will be relevant to the discussion. For example, covariance becomes important, both by year and LOB:[29]

- **Concept 1:** For each (accident, policy or report) year, the coefficient of variation (standard error as a percentage of estimated liabilities) should be the largest for the oldest (earliest) year and will, generally, get smaller when compared to more and more recent years.
- **Concept 2:** For each (accident, policy or report) year, the standard error (on a dollar basis) should be the smallest for the oldest (earliest) year and will, generally, get larger when compared to more and more recent years.[30] To visualize this, remember that the liabilities for the oldest year represent the future payments in the tail only, while the liabilities for the most current year represent many more years of future payments including the tail. Even if payments from one year to the next are completely independent, the sum of many standard errors will be larger than the sum of fewer standard errors.
- **Concept 3:** The coefficient of variation (standard error as a percentage of estimated liabilities) should be smaller for all (accident, policy or report) years combined than for any individual year.

---

[29] These covariance standard error concepts assume that the underlying exposures are relatively stable from year to year – *i.e.*, no radical changes. In practice, random changes do occur from one year to the next which could cause the actual standard errors to deviate from these concepts somewhat. In other words, these concepts will generally hold true, but should not be considered hard and fast rules in every case.

[30] For example, the total reserves for 1990 might be 100 with a standard error of 100 (coefficient of variation is 100%), while the total reserves for 2000 might be 1,000 with a standard error of 300 (coefficient of variation is 30%).

- **Concept 4:** The standard error (on a dollar basis) should be larger for all (accident, policy or report) years combined than for any individual year.
- **Concept 5:** The standard error should be smaller for all lines of business combined than the sum of the individual lines of business – on both a dollar basis and as a percentage of total liabilities (*i.e.*, coefficient of variation).
- **Concept 6:** In theory, it seems reasonable to allocate any overall "reserve margin" (selected by management) based on the standard error by line after adjusting for covariances between lines.

To simplify the calculations, claim payments by period are often assumed to be normally distributed in many of the commonly used models for to estimating liabilities. This can be a useful assumption when working through fictitious examples, but the actuary must be very careful when using these assumptions with real data:

- **Assumption 1:** For lines of business with small payment sizes (*e.g.*, Auto Physical Damage) this might be a reasonable simplifying assumption.[31]
- **Assumption 2:** For most lines of business, the distribution of individual payments, or payments grouped by incremental period, is skewed toward larger values. Thus, it would be better to model the claim payment stream using a Lognormal, Gamma, Pareto, Burr or some other skewed distribution function that seems to fit the observed values.
- **Assumption 3:** Estimating the distribution of loss liabilities (in total or by accident or payment period) assuming that the claims are normally distributed could produce misleading results for management whenever the actual claims are not normally distributed. The relevance of this distortion compared to the cost of improving the estimates needs to be considered.
- **Assumption 4:** Estimating the standard error in the claim payments assuming a normal distribution and then simulating the total loss distribution using a lognormal distribution (or some other skewed distribution) is marginally better, but it will require much greater skill and

[31] Even though using the normal distribution might be a reasonable simplifying assumption, the actuary must still exercise caution. For example, for some combinations of mean and standard error (*e.g.*, low mean, high standard error) the calculated range could include negative values.

care than using an assumption based on parameters assuming a lognormal (or some other skewed distribution) and testing to see how well this fits the actual data.

Since the projection of incurred losses does not directly measure the variability of the future payment stream, its usefulness in determining liability distributions should be considered:

- **Consideration 1:** The "extra" information in the case reserves is generally believed to add value by giving a "better" estimate of the expected mean. The exceptions to this are well documented in the actuarial literature. However, does this "extra" information really change the estimate of the *expected value* of the payment stream (by year), or does it give a better "credibility adjusted" estimate of the *likely outcome* (by year) as the additional information comes to light and leave the expected value of the payments unchanged?[32]
- **Consideration 2:** Consider two identical books of business with two different insurance companies.[33] They are identical except that one company sets up case reserves on the claims and the other does not. The estimates of the total liabilities (IBNR vs. case plus IBNR) are identical. Will the deviations of actual from the expected value of the future claim payments be any different?
- **Consideration 3:** Since measuring the variations in the incurred claims does not directly measure the variations in the payment stream, should risk measures based on incurred claims be used to quantify risk for management? With consistent levels of case reserves, the variations in the incurred claims might be more stable and might converge more quickly towards the actual outcome, but would this measure mask some of the true volatility? On

[32] The appropriate question here is whether the case reserve information can be used "optimally" in the sense that an appropriate credibility-weighted estimate is produced from the paid data and the case reserves. Let us assume that there is a small amount of information in the case reserves, but the additional information it contains about the payments requires the use of a model (at a minimum, you'll need to work out the mean, variance and covariance of the forecasts given the case reserves). That is, if A is some forecast of payments (whether an individual forecast or some total), P is the set of past payments, C is the set of past case reserves, and let's say we want the distribution of A, f[A | P,C]. Then if there is no parameter uncertainty (and ignore all the kinds of model uncertainty), it is true that f[A | P,C] must have a smaller (or no larger) standard error than f[A | P]. However, the moment you look at a predictive distribution, this is no longer true, because you have additional parameter uncertainty (and model uncertainties will compound the problem). For case reserves to help you forecast, any additional information would have to be larger than the additional predictive uncertainty the larger model introduces.

[33] This thought exercise also applies to the same book of business before and after the addition of case reserves to the claim settlement process.

the other hand, with case reserve strengthening or weakening, the variations in incurred claims may be less stable than for paid claims and could possibly overestimate volatility.

6. MODELS FOR CALCULATING RANGES

Historically, the problem of quantifying a probability distribution for a defined group of claim payments has been solved using "collective risk theory."[34] Actuaries have built many sophisticated models based on this theory, but it is important to remember that each of these models make assumptions about the processes that are driving claims and their settlement values. Some of the models make more simplifying assumptions than others, but none of them can ever completely capture all of the dynamics driving claims and their settlement values. In other words, none of them can ever completely eliminate "model risk."

For example, consider this thought exercise. Do claim adjusters base their individual claim payments on the cumulative value of past payments for each claim? No, they base each incremental payment on the circumstances at the time.[35] Thus, claim payments are not generally correlated to the cumulative payments to date. However, a convenient simplifying assumption is made when using models based on link ratios that the cumulative payments are correlated, but this creates a bias whereby "unusually" low cumulative values tend to under-predict the ultimate and "unusually" high cumulative values tend to over-predict the ultimate. Every actuary recognizes this bias (either implicitly or explicitly) and quite often the Bornhuetter-Ferguson model and informed judgment are used to adjust for this bias.

In fact, Venter [24] has shown that models based on link ratios often fail to be good predictors when you test the underlying assumptions. The chain ladder model (*i.e.*, weighted average of all link ratios) is actually a form of regression through the origin. Venter showed that quite often a better

---

34 There are a number of good books on the subject, including, but not limited to, Buhlmann, "Mathematical Models in Risk Theory"; Gerber, "An Introduction to Mathematical Risk Theory"; and Seal, "Survival Probabilities".

35 A possible exception to this might be cases involving annuity type claims, but even here if the circumstances change then the future claim payments could change or stop altogether. Quite often, claim adjusters make one payment on a claim and not multiple payments. When evaluating that payment, similar cases are considered at that time. It might be the timing of when the payments on similar type cases are made that matters more, but this still implies that the timing of when the payment is made is more significant than the cumulative history of other payments.

predictor is an average plus a constant (*i.e.*, slope not through the origin) or perhaps just a constant term.

A range of estimates using models based on link ratios should necessarily exclude using link ratio models when the assumptions underlying the models aren't strictly met – *i.e.*, they fail tests of their predictive value as described by Venter. In other words, if you have "bad" estimates, they are "bad" estimates and shouldn't enter into the determination of the "reasonable" range.[36] In the discussions that follow, all estimates using link ratio models are assumed to pass these tests.

Models based on incremental payments get around this "limitation" of the link ratio models and also have the advantage of more directly measuring the fluctuations in the timing and amount of the future claim payment stream. On the other hand, incremental models are less well known (or at least seem to be used in practice and discussed less often) and can be more difficult to apply for certain data sets. As always, the practicing actuary needs to be familiar with the advantages and disadvantages of each model used to estimate liabilities.

For purposes of this paper, the models used to calculate liability ranges will be grouped into four general categories: multiple projection models, statistics from link ratio models, incremental models, and simulation models.

A. Multiple Projection Models

In this category, the actuary uses multiple models and possibly various assumptions for each model to come up with a variety of possible estimates. Usually this involves models based on link ratios (at least in part) and it is assumed that these various estimates are a good proxy for the variation of the expected outcomes. This is inconsistent with the process underlying the concepts set forth in this paper in several important respects:

[36] While common sense and various sections of ASOP No. 36 would seem to imply this type of testing of the assumptions in a loss estimation model, the Actuarial Standards Board may wish to consider adding language to more directly address this issue.

- The projected estimates produce a range, but it does ***not*** provide a measure of the density of the distribution for the purpose of producing a probability function – it simply produces a range of estimates for the mean, but only to the extent that the actuary varies the models and assumptions.[37]
- The "distribution" of the projected means is a distribution of the models and assumptions used, ***not*** a distribution of the expected future claim payments.[38]
- While models based on link ratios are often assumed to be estimating the expected value of the reserves, in point of fact they only produce a single point estimate and there is no statistical process for determining if this point estimate is close to the expected value of the distribution of possible outcomes or not.
- Since there are no statistical measures for these models, any overall distribution for all lines of business combined will be based on the addition of the individual ranges by line of business with judgmental adjustments for covariance, if any.

While there are serious statistical limitations and drawbacks to using multiple projections to determine a liability range, we must recognize that producing any range is better than no range at all. Also, data limitations may prevent the use of more advanced models. Unfortunately, multiple projections don't provide a true probability range based on statistics, so the more sophisticated models described later would normally need to be used in practice or appropriate caveats will need to be included in the actuarial report.

Unfortunately, a strict interpretation of the guidelines in ASOP No. 36 would generally lead the actuary to use this model to create a "reasonable" range. In addition, one may wonder how often the tests outlined by Venter are actually being used to remove estimates that fail these tests from these "reasonable" ranges in practice. Given these limitations, therefore, it would seem

[37] Perhaps a better description for a range of estimates of the mean is "scenario testing."

[38] With enough estimates a nice bar chart showing the number of estimates that fall into selected intervals can be produced. However, while it may look rather like a probability distribution, it is just a bar chart that looks like a histogram and it wasn't generated by any random process. It was generated by the principle that underlies all scientific investigation: If something is quite reasonable, it can be justified in a lot of different ways. But if something is almost unreasonable, then it can be justified in only a limited number of ways, often only one.

prudent for the actuarial profession to consider adding language similar to the following to ASOP No. 36:

"Whenever a range of expected values is produced as the range of reasonable estimates, and the actuary has no further means of producing a reasonable distribution of possible outcomes, then the midpoint of the range of expected values should be used as the minimum acceptable reserve."

This would add language to ASOP No. 36 which is consistent with the definition used in the SSAP's for "Ranges of Reserve Estimates."

B. *Statistics from Link Ratio Models*

In this category, the models described by either Mack [16, 17] or Murphy [20] and others, can be used by the actuary to calculate the standard error in the payment stream using the variation in the link ratios. The actuary can use the standard error to calculate the distribution of the liabilities using the cumulative normal distribution or use logs to get a skewed distribution. These models are better than using Multiple Projections, but they are still inconsistent with some of the concepts set forth in this paper:

- The expected value used in these models is still based on multiple models and is subject to most of the same limitations described above for multiple projections.
- The standard error calculations in these models often assume that the distribution of the link ratios is normally distributed and is constant by (accident) year – this violates three concepts: 1) link ratios are a measure of the cumulative claim payment variations not the incremental variations (definition of risk), 2) the claim payments are usually not normally distributed (Assumption 2), and 3) the standard errors should not be constant across (accident) years (Concept 1).
- The standard error values from these models provide a process for calculating an overall probability distribution for all lines of business combined. However, this will require making assumptions about the covariances between lines or assuming independence among lines. Further research is needed to develop additional formulas for calculating the covariances between lines of business.

Using statistics from link ratio models is a significant improvement over ranges based on multiple projections since the variations in the underlying data are more directly modeled and used in the results. In other words, it is focused on calculating a distribution of possible outcomes given an estimate of the expected value. For these models, it would also seem reasonable to apply the language suggested above for ASOP No. 36 to the expected value portion of the calculations.

If data limitations prevent the use of models based on incremental values, then this model will need to be used. Otherwise, incremental models would normally be preferable.

C. Incremental Models

Models based on the incremental values of claims paid from one period to the next have been under development for quite some time.[39] These models generally overcome the "limitations" of using cumulative values and have the advantage of modeling calendar year inflation (along the diagonal) using a separate parameter(s). They also generally comply with the concepts set forth in this paper, with only a few exceptions:

- Several of the models in general use assume that the distribution of incremental claims is lognormal. The actual distribution of incremental payments may or may not be lognormal, but this is a significant improvement over models that assume normality and generally this provides a good fit to the actual data. Other skewed distributions are also used, but they generally add complexity to the formulations.
- Like for the other categories, when adding liability estimates for individual lines of business the correlations between lines will need to be considered when they are combined. Recent papers by Brehm [6] and Kirschner, et. al. [15] are good examples of how incremental models can be correlated and combined. Research in this area is ongoing.
- An added bonus is that some of these models allow the actuary to thoroughly test the model parameters and assumptions to see if they are supported by the data. They also allow the actuary to compare various goodness of fit statistics to evaluate the reasonableness of

[39] A brief sampling from the actuarial literature could include papers by Finger [11], Hachemeister [12], Zehnwirth [3, 28], England [9, 10] and Verrall [9, 10] to name but a few.

different models and/or different model parameters. Essentially, they allow the actuary to tailor the model parameters to fit the characteristics of the data.

For the purpose of calculating a distribution of possible outcomes, incremental models are a significant improvement over models based on link ratios since they are focused on directly calculating the distribution and then the expected value is determined from the distribution itself. The main limitation to these models seems to be only when some data issues are present.[40]

D. *Simulation Models*

Because of the complex interactions between claims, reinsurance, surplus, *etc.*, a dynamic risk model may be needed in order to more fully test the reasonableness of the range of liabilities. Models from all of the previous three categories can be used to create such a risk model, but in order to evaluate them we need one more concept:

- **Concept 7:** Whenever simulated data is created, it should exhibit the same statistical properties as the real data. In other words, the simulated data should be statistically indistinguishable from real data.

Unfortunately, simulation models based on link ratios tend to be the least useful since they quite often exhibit statistical properties not found in the real data being modeled. Whenever link ratios are shown to be worse predictors than a constant, or link ratios plus a constant, data simulated using link ratios *will be distinguishable* from real data. While this problem may not invalidate the conclusions from a liability simulation study, it will certainly reduce the reliability of the results.[41]

---

40 A good example is when separate data for Salvage & Subrogation is not available. In this case, when the "tail" of the loss development pattern contains a significant amount of negative values they cannot be modeled using logs.

41 While taken out of context, the following quote is still relevant. "The bone of contention will be whether a model, to be of any use, must be 'essentially' realistic, or whether an admittedly unrealistic model may have its purposes. I hold that, so long as we don't forget the unrealistic assumptions we have made, we are free to make what models we will and then see what insight, if any, they yield."; Gene Callahan, "Choice and Preference", Ludwig von Mises Institute, www.mises.org, article posted Feb. 20, 2003.

This problem with "link ratio simulations" is usually overcome with models based on incremental values. It can also be overcome with ground-up simulations using separate parameters for claim frequency, severity, closure rates, *etc.* As with any model, the key is to make sure the model and model parameters are a close reflection of reality.[42]

7. PRACTICAL CONSIDERATIONS

Up to this point, the discussion has been mainly focused on theoretical and philosophical issues related to using probability ranges. Before the paper is concluded, it will also be useful to focus on some considerations of using probability ranges in practice.

A. *Are Reasonable Assumptions Enough?*

Some actuaries may find themselves not agreeing with the conclusion that the phrase "a reasonable range" is meaningless without some other context. Their reaction may be that context is provided by the phrase, "that could be produced by appropriate actuarial models or alternative sets of assumptions that the actuary judges to be reasonable." In other words, the sentence, "The reasonable range is from $A to $B" must make sense in light of reasonable statements about the history of cost drivers (such as premium, exposure, and benefit changes) and about the history of loss development (such as age-to-age factors or severity trend rates).

Turning to what is "reasonable" under the definition in ASOP No. 36, it seems safe to say that "reasonableness" is determined by the actuarial culture. By talking to other actuaries, attending conferences, talking with clients, reading the newspapers, and reading some of the actuarial literature, we maintain a culture that reflects actuarial expertise. Assumptions and statements that are consistent with this culture are necessarily reasonable, even if we personally disagree with them. Assumptions and statements that would be considered misleading in the context of

---

[42] Actually, there is a very real sense in which "unrealistic" models are to be preferred when forecasting. A model should tend to under-parameterize somewhat, if one wants a minimum mean square prediction error forecast – one should, for example, tend to over-smooth rather than fully fit all changes in trend, *even where you know for certain* there is a change. Often a substantial reduction in the effect of parameter uncertainty on the variance of the forecast comes at the price of a smaller increase in (squared) bias.

that culture are usually unreasonable – but one exception is statements that are well argued and supported with data, because that is how the culture is changed over time.

The author would certainly agree that culture is an appropriate context for our guidelines, but the use of probability ranges will add a new dimension to the guidelines. For example, even if every actuary in the world were to agree that all of the assumptions and models used to develop the range $A to $B are reasonable, we are still left with the question, from a solvency point of view at least, of "What makes selecting $A as the final reserve any more or less 'reasonable' than $B or any other number in between?"[43] Without any further guidance do we, as a profession, have any basis for selecting one number in the range over another?

What if two or three actuaries with appropriate training and experience estimate that a given liability has an expected value of $100 million[44] but the range of expected values is $70 to $140 million based on the information and support their conclusion with reasonable models and assumptions. Is $70 million a reasonable estimate? Based on current standards, unless there are assumptions that are "unreasonable," or data they have overlooked, or a mistake in their work, then the $70 million must be considered reasonable since it is "within the reasonable range" as currently described in our guidelines.

On the other hand, what if those same actuaries develop a distribution of possible outcomes with an expected value of $100 million and the end points of the range noted above correspond to the 25th and 80th percentiles, respectively. If there is only a 25% chance that $70 million is sufficient to cover all future claims, then is it still a "reasonable" estimate? It is not up to the author alone to determine at what percentile an estimate changes from reasonable to unreasonable, but it sure seems like it should be much closer to the expected value (or higher) than the 25th percentile. Since no model can ever remove all of the subjectiveness from the estimation process, setting an absolute percentile that the actuary cannot go below may not be a good idea. But theoretically at least, the expected value seems to be a logical minimum for a reasonableness standard.

---

[43] More or less adequate is a different question than where to draw the line on "reasonableness."

[44] As with previous examples, the time value of money is being ignored to simplify the discussion.

A standard that is less than the expected value would be akin to recommending to a casino that they set the odds at something less than in their favor.[45] While some constituents may consider a percentage lower than the expected value to be a reasonable lower bound, the principle of greatest common interest would suggest that other interested parties, such as stockholders, policyholders and solvency regulators, who would likely insist on at least an expected value standard as the minimum for the reasonable probability range.

Stated differently, the current guidelines seem to be saying that as long as the actuary can document the reasonableness of the models and assumptions used to arrive at a "possible outcome" then, ipso facto, that "possible outcome" is reasonable. Rather than *only* reviewing the reasonableness of the underlying models and assumptions, in and of themselves, the contention of this paper is that the actuary *also* needs to look at the reasonableness of that "possible outcome" in relation to all other possible outcomes. In other words, no matter how reasonable a given model and assumptions are, is that "possible outcome" reasonable if it is less than the expected value given a reasonable distribution of possible outcomes?

Turning to Statements of Actuarial Opinion, how should the actuary respond to the example described above if management wishes to book $70 million? Some actuaries may say "I can't find a way to shoot down the 'optimistic' assumptions that resulted in an estimate of $70 million as being unreasonable, I just think there is a lot of uncertainty." Should the actuary then give a "clean" opinion because management made a good case, but unless something changes, include a sentence in the "risks" section of the opinion that there is a 75% chance this will prove to be inadequate? Or, should the actuary give a qualified opinion? This will need to be answered by the actuarial profession and other constituents that are the intended audiences for the actuarial work product. On the other hand, if management does book the expected value, at what point does the actuary need to report the high end of the liability range in the "risk" section of the opinion?[46]

[45] Actually, the casino would not want to set their odds at less than the expected value, plus a risk margin based on the process risk.

[46] This point has been debated among actuaries for at least 25 years as attested by the following quote from Bailey, Robert A. *The Actuarial Dilemma*, The Actuarial Review, Volume 5, No. 1, January, 1978, p. 7. "Loss reserving is about as actuarial as any work can be because it involves an estimation of an unknown quantity which is subject to future contingencies (inflation,

It is hoped that clarifications to the standards of practice will provide answers to these questions. In addition, the Committee on Property-Liability Financial Reporting may wish to define "risk" for purposes of a Statement of Actuarial Opinion in relation to the range of possible liability outcomes. For example, it could be "recommended that, if possible, the actuary disclose the 95th percentile for their estimated range of possible liabilities."

Another problem with the current definition of a "reasonable range" is the way it is implemented in practice. In theory, if actuary A says that the liability is \$X, and actuary B finds that this is in the reasonable range as measured by ASOP No. 9 (Documentation), ASOP No. 36 (Reserves) and the CAS principles, then actuary B should give a clean opinion. That is, actuary A, who presumably knows the situation better, is to be believed unless there is a problem. In practice, insurance companies can use the existence of the "reasonable range" as currently defined to create space to manage earnings. Using a "probability standard," actuary A would then be required to report where they believe \$X is with respect to the probability distribution of possible outcomes. In addition, actuary A could also be required to treat any material change in this percentage from one year to the next as a change in assumptions.[47]

It is easy to see how well-intentioned experienced actuaries could follow the standards of practice to the letter and end up signing a clean opinion on reserves that have a "high" probability of being deficient. In addition, in practice some of the modeling deficiencies described in the previous section could be compounding this issue by distorting the quality of the actuary's calculated range.

The wording in the ASOP's was worked out by actuaries who were familiar with mathematical models and yet decided that such models did not provide the solution. It may be safe to surmise they were concerned that mathematical models alone do not create a wide enough "safe harbor"

---

court settlements, *etc.*) based on past experience and informed judgment. But if estimating the value of unpaid claims is actuarial, certainly the appraisal of the degree of uncertainty associated with that estimate is at the very core of actuarial work. What could be closer to the theory of risk? If we succeed in avoiding the appraisal of the uncertainty in loss reserves, by simply stating that in our opinion the reserves are 'reasonable,' which means, I suppose, that the reserves have a 50% likelihood of being adequate, don't we leave a vacuum to be filled by some other profession?"

47 Materiality for these purposes will need to be related to the concept of materiality in other contexts noted earlier in the paper. For example, a "material change" could be defined as "an increase or decrease of more than 10 percentage points in the probability that the carried reserves are adequate."

for actuarial practice. Yet, given the questions raised by looking at probability ranges, one has to wonder if we might have inadvertently created a "safe harbor" that is potentially too wide at the low end? While there are many references to "uncertainty" in the ASOP's, additional guidance on what should be disclosed at the high end of the range also seems appropriate.

B. *The Evolution of Information*

It can be said that a range of reasonable reserves is a function of evidence, not just possible outcomes. For example, if the only information about a block of business is that it was priced to produce an 80% loss ratio, then the only reasonable liability estimate one can make is 80% of earned premium. The range widens and shifts as, and only as, other evidence emerges showing that other outcomes are reasonable (and perhaps that 80% is no longer reasonable).[48]

For a new block of business, the only evidence for setting reserves is the pricing documentation used to produce the rates (let's call this anecdotal evidence). As this block of business is observed over time, more and more evidence (let's call this physical evidence) emerges about how it is performing relative to the initial estimates and to any new updated pricing estimates (more anecdotal evidence). However, even if an 80% loss ratio is reasonable throughout this entire process that does not mean that other outcomes are not possible at every point along the way. As time passes, the physical evidence leads us toward the actual outcome and less weight is given to the anecdotal evidence, but in general 100% weight is not given to the physical evidence until all claims are closed.[49] While the physical evidence is leading toward the actual outcome for each year, statistically the *a priori expected* outcome may not be moving or may be moving in the opposite direction from the actual outcome (See Graph 3).

This discussion can be summarized using one of the questions noted earlier in the paper. Namely, does this "extra" evidence really change the estimate of the *expected value* of the payment stream (by year), or does it give a better "credibility adjusted" estimate of the *likely outcome* (by year) as the additional evidence comes to light and leave the expected value of the payments

[48] This does not mean that there is no range to start with. Quite the contrary, historical data or other anecdotal evidence could be used to calculate a reasonable *a priori* estimate of the range.

[49] A nice feature of the Bornhuetter-Ferguson model is that it shifts the weight over time using a nice mathematical (Bayesian) process.

unchanged? While the earlier question was aimed at the merits of determining risk using paid claims vs. incurred claims, it is equally relevant here.

This question, in turn, leads us to the realization that reserves are accounting fictions – they are estimates of liabilities, not the liabilities themselves.[50] Thus, we might also look to the accounting profession for some additional principles that might be relevant. For example:

At the high end of the range, according to a general principle of accounting, a liability should not be recorded for an "event" that has not yet occurred. It is a settled issue that an "event" is the claim itself, but how far does it go to include the conditions under which the claim will be settled? For example, if inflation (CPI) has historically been about 3% and the data for a line of business is consistent with the CPI, it seems reasonable to estimate the high end of the range assuming inflation of 3% in the future. Would the high end of the range only increase if inflation actually increased above 3%? Or, is it reasonable to assume that inflation could increase above 3% and include that possibility as part of the reasonable range? Another area where these questions are relevant is with emerging theories of law or legislated changes that are allowing new claims to be filed which were not anticipated in years past. A good example here is newly emerging legal theories of asbestos liability that were not known years ago.

At the low end of the range, according to a general principle of accounting, a business should not record a profit on a particular activity until it has data to support the estimation of that profit. Accordingly, the low end of the range should be selected in order to produce zero profit in the period if there is insufficient data to establish that a profit has been earned. Recording a liability any less than $X would create the incorrect impression that the business was known to be profitable. This principle seems consistent with keeping the minimum probability for the reasonable reserve range at the expected value or above.

[50] As noted earlier, this "realization" is already recognized in the Statement of Principles definition of **Loss Reserves** as "a provision for its related liability."

C. Who is the Audience?

While it is the contention of this paper that a probability range should be used to determine what is "reasonable," we must also recognize that precisely defining what a "reasonable" probability range is may depend on the audience and, if possible, the audience should define what is "reasonable" to them. For example, solvency regulations and organizations concerned mainly about solvency (*e.g.*, state regulators, A. M. Best's, S&P, *etc.*) may feel that prudence would require a range with a minimum corresponding to the expected value and a maximum of, say, 85 or 90%. Other regulatory bodies might define the "reasonable" probability range differently (*e.g.*, the IRS might consider a range from 50% to 75% to be reasonable for tax considerations and the NAIC might have different ranges for statutory reserves compared to rate filing regulations). However, all of these different constituencies could use a probability range as a consistent starting point or perhaps even agree on a consistent lower bound to the probability range.

The *principles of least (greatest) common interest* apply when there are multiple parties that have an interest in a certain outcome. This is almost always true of actuarial reports, which means that there can be conflicting goals from the different audiences. It is easy to identify direct users of the report (*e.g.*, management, the Board of Directors, regulators, *etc.*), but it is not always clear who might indirectly use or benefit from the report (*e.g.*, stockholders, policyholders, consumer groups, *etc.*).[51]

We should also recognize that these two principles have the potential to cause ranges from two difference audiences to not intersect (*e.g.*, the high end of the range for one party is below the low end of the range for another party). If this should occur, it is hoped this approach to determining "reasonableness" will provide both parties with a method for working out their differences. Alternatively, it could be used to more clearly define difference between accounting standards used for different audiences (*e.g.*, GAAP vs. Statutory vs. Tax Accounting rules).

[51] The *principles of least (greatest) common interest* are ***not*** intended to suggest that the actuary should attempt to identify all possible users of their work. This would be an onerous requirement. What it does suggest is that the actuary should not be able to select an end point for their liability range that is acceptable to one of the users of their work when it would clearly not be acceptable to other ***readily identifiable*** users of their work.

The final phrase in ASOP No. 36's definition of the range of reasonable reserves is "A range of reasonable reserves, however, usually does not represent the range of all possible outcomes." While the use of a probability range is not in conflict with this statement, the example discussed in Section 7.A. shows that it is subject to interpretation. In that example, it could be used to simply state that the range from $70 to $140 million does not include all possible outcomes. However, under a probability range approach it would be used to say "Of course outcomes less than $100 million are possible, but they are not reasonable since the probabilities that they are insufficient are too high. On the other hand, there is a 20% chance that outcomes above $140 million are also possible and the 20% probability may be too low given model risk that is incalculable or other unforeseen events."

Given the wide range of possible audiences for an actuarial work product, it seems prudent to err on the side of including more information rather than less. While in some cases this could increase the actuary's exposure to malpractice, in most cases this exposure should be reduced. For example, if the unexpected happens (let's say payments end up equaling $200 million in the example from Section 7.A. and the company ends up in bankruptcy), the actuary may be *exposed* to a claim of malpractice no matter what they said.[52] If the actuary simply told management the range ends at $140 million then there will be some explaining to do. But, if the actuary provided management with a probability range and also noted that there was a 5% chance that it could reach $200 million, then management will be in a much better position to make a decision on what reserves to carry and will not be able to say that this outcome was unforeseeable.

Using a probability range for liabilities, there seems to be two main reasons that actuarial malpractice could occur (excluding other potential reasons, like fraud):

1) If the actuarial models, assumptions and/or calculations used to create the overall expected outcome (within the distribution of possible outcomes) are faulty, or

[52] Being exposed to a claim of malpractice and actually being guilty of malpractice are a far cry from each other. Within the actuarial profession, the possible reasons for being guilty of malpractice have been the subject of considerable debate and are the purview of the Actuarial Board for Counseling and Discipline (ABCD). It is hoped that a statistical approach for determining reasonableness will help bring additional focus to the debate.

2) If the distribution of possible outcomes is "correct" given fully tested models and assumptions, but the actuary failed to alert the proper authorities that management was booking an amount that was less than the "reasonable" minimum, whatever percentage that turns out to be.

It doesn't seem right that getting the distribution of possible outcomes "correct," but years later finding out that the actual outcome is higher than the expected value, would be grounds for malpractice in and of itself. However, the public perception of getting it right and actually getting it right are two different things (especially in the hands of a skilled attorney). How much longer can the actuarial profession risk telling our constituents what is expected and not also telling them what is possible?

4. *When Does Insolvency Occur?*

The previous discussions about how probability ranges for liabilities are related to materiality can naturally lead to the question: "When is an insurer insolvent?" Does an insurer become insolvent when their surplus was actually inadequate or when a regulator finds out about it?

For instance, suppose a "clean" loss reserve opinion is given on the company described in Section 4 as "medium" risk in scenario A (*i.e.,* carried reserves of $100 million, surplus of $80 million and probability of insolvency is 60%). Years later it turns out that the paid losses for claims represented by those reserves are likely to exceed $200 million. Was the company actually insolvent when the opinion was given? Or, does it become insolvent when the "higher than expected" claim payments indicate that the likely outcome will exceed $180 million? What if subsequent years improve such that cash flow never becomes an issue? What if subsequent years get worse?

At one extreme it could be reasoned that the insolvency actually took place when the clean opinion was given or even as early as when the business was written that resulted in the eventual insolvency. The rationale for this view rests on the assumption that insolvency is a technical condition not a human discovery of that condition. This would also be distinguished from actions taken by management and/or regulators in response to their discoveries.

At the other extreme, it could be reasoned that insolvency doesn't take place until the insurer reaches the point where it can't meet current cash flow needs. Unfortunately, at this extreme the identified liabilities will usually far exceed the current assets. It's not surprising then that regulators have set solvency requirements, via Risk-Based Capital requirements, so that they can take action before the insurer gets into cash flow difficulties. Therefore, a more reasonable alternate extreme might be that the insolvency has taken place at the time the information becomes available to value the company's surplus below RBC standards.

While both of these extremes are useful in framing the discussion, both of them rest on the assumption that future liabilities are known (or knowable with a very high degree of certainty). Until the liabilities are completely run-off no actuary can tell exactly what they will be. At either point in time (original valuation date or retroactive discovery date), two different actuaries will have two (or more) different estimates of what the liabilities are. If one estimate indicates that liabilities exceed assets and the other one doesn't, which one is right? The answer is neither of them is right.

If liabilities are viewed as a distribution of possible outcomes, instead of an actuary's best estimate or even a range of best estimates, at any point in time there is some probability that the future liability payments will exceed current assets (or more accurately future assets). So, from this perspective, the question becomes how high must this probability become in order for insolvency to occur or regulatory action to be triggered? Perhaps the added perspective of probability ranges will prove useful to actuaries and regulators as they continue to fine tune and improve the RBC formulas.

8. AREAS FOR FUTURE RESEARCH AND ANALYSIS

Throughout the paper, several areas for future research have been identified (or at least hinted at). For easy reference, they are summarized below:

- One of the suppositions in this paper is that measures of reserve risk should be based primarily on paid data, although some potential information from incurred data was also discussed. Research of measures of risk based on paid claims vs. incurred claims would be necessary to reach any definitive conclusions. Research papers to develop models that quantify the predictive value of case reserves and credibility weight that information with estimates based on paid data would also be a valuable addition to our literature.
- Various models for calculating probability ranges are discussed in the paper along with advantages and disadvantages of each. A research project involving retrospective testing of various models used to calculate ranges would yield insights into how significant these advantages and disadvantage are. To accomplish this, the author suggests a "blind" test with old data from multiple companies and multiple lines of business. The data should be at least 10 years old so that the final results are already known, but the tests should be run using only the triangles that would have been known 10 (or more) years ago.
- Continuing research on covariance calculation methods is a significant feature of any model used to calculate probability ranges of liabilities for an entire company.
- Further research on the relationship between reserve risk and insolvency risk could lead to additional insights on how to define a "reasonable" probability range. It might also lead to some RBC insights or triggers for when a company should consider increasing its capitalization or have enough "extra" capital before paying dividends.
- Research on the quantification of "model" risk would be a welcome addition which could help move this from a judgmental to a calculated amount. Even when calculated amounts aren't a possibility, it would help improve informed judgment.
- Research on the differences between measures of reserve risk based on quarterly data vs. annual data should be performed in order to help guide actuaries when dealing with issues related to quarterly vs. annual accounting statements.

9. CONCLUSIONS

This paper started by reviewing some of the professional standards for determining the "reasonableness" of loss reserves and proceeded to examine how various statistical concepts might be used in conjunction with current standards. The main conclusions of this analysis are that using a probability range has the following benefits:

- Users of actuarial liability estimates based on probability ranges will get much more information for risk evaluation and decision-making,
- The width of the dollar range will be directly related to the potential volatility (uncertainty) of the actual data,
- The concept of materiality can be more directly related to the uncertainty of the estimates,
- Risk-Based Capital calculations could be related to the probability "level" of the reserves,
- Both ends of the "reasonable" range of reserves will be related to the probability distribution of possible outcomes *in addition to* the "reasonableness" of the underlying assumptions,
- The concept of a "prudent reserve margin" could be related to a portion of the probability range and will then be directly related to the uncertainty of the estimates, and
- The users of actuarial liability estimates would have the opportunity to give more specific input on what they consider "reasonable."

In order to implement the advantages of the statistical approach, the actuarial profession should consider adding wording similar to the following to ASOP No. 36:

> *"Whenever the actuary can produce a reasonable distribution of possible outcomes, a reasonable reserve estimate should not be less than the expected value of that distribution."*

Essentially, this paper is NOT proposing that we eliminate the "what a reasonable person might do" standard and replace it with probabilities. What it is suggesting is that we can improve the "reasonable person" concept by adding some additional context. There must be no illusions here. Adding a probability measure to the "reasonable person" standard will not provide a magic solution to define the exact number where the minimum "reasonable" reserves should be. Calculating the mean of the distribution is no less difficult. However, adding "probability standards" can make the "reasonable person" standards more meaningful.

In addition, the ASOP No. 36 definition of **Risk Margin** could be improved by adding wording similar to the following:

> *"A risk margin should include an amount(s) to reflect 'process,' 'parameter' and 'model' risk. Whenever possible, it should be statistically calculated, otherwise a judgmental amount can be included."*[53]

Other issues mentioned in the paper that should also be addressed in our standards include: 1) the need to consider language to more directly require testing of the assumptions for different models, 2) a more definitive solution for how to consistently disclose the relative reserve risk, and 3) a more precise definition of "material change" as it relates to reserve risk.

Finally, we must not forget that calculating a distribution of possible outcomes is not always possible. In that event, adding wording similar to the following to ASOP No. 36, as suggested earlier in the paper, would be consistent with the SSAP's:

> *"Whenever a range of expected values is produced as the range of reasonable estimates, and the actuary has no further means of producing a reasonable distribution of possible outcomes, then the midpoint of the range of expected values should be used as the minimum acceptable reserve."*

In closing, ask yourself the following question: "WHAT IF you knew the EXACT distribution of possible liability outcomes, would you feel comfortable giving a clean opinion to a company that wanted to carry less than the expected value on their books?" As a profession we want the outside world to rely on our "actuarial judgment" to determine what is "reasonable." Will your answer give the public added confidence in the profession? Doesn't it make sense to strengthen our standards in order to increase public confidence?

---

[53] Definitions or "process," "parameter," and "model" risk consistent with the definitions in this paper may also need to be added.

REFERENCES / SELECTED BIBLIOGRAPHY

[1] Actuarial Standards Board, Actuarial Standard Of Practice No. 9, "Documentation and Disclosure in Property and Casualty Insurance Ratemaking, Loss Reserving and Valuations," Adopted January, 1991.

[2] Actuarial Standards Board, Actuarial Standard Of Practice No. 36, "Statements of Actuarial Opinion Regarding Property/Casualty Loss and Loss Adjustment Expense Reserves," Adopted March, 2000.

[3] Barnett, Glen and Zehnwirth, Ben, "Best Estimates for Reserves," Casualty Actuarial Society *Forum* 3, Fall 1998, pp. 1-54.

[4] Berquist, James R. and Sherman, Richard E., "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," PCAS LXIV, pp. 123-184.

[5] Bornhuetter, Ronald L. and Ferguson, Ronald E., "The Actuary and IBNR", PCAS LIX, 1972, pp. 181-195.

[6] Brehm, Paul J., "Correlation and the Aggregation of Unpaid Loss Distributions," Casualty Actuarial Society *Forum*, Fall 2002, pp. 1-24.

[7] Butsic, Robert P., "Determining the Proper Interest Rate for Loss Reserve Discounting: An Economic Approach," 1988 CAS Discussion Paper Program, pp. 147.

[8] Christofides, S., "Regression Models Based on Log-Incremental Payments," Claims Reserving Manual, Institute of Actuaries, London, 1990.

[9] England, Peter D. and Verrall, R. J., "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving," Insurance: Mathematics and Economics, pp. 281-293.

[10] England, Peter D. and Verrall, R. J., "Stochastic Claims Reserving in General Insurance," Presented to the Institute of Actuaries, 28 January 2002.

[11] Finger, Robert J., "Modeling Loss Reserve Development," PCAS LXIII, p. 90.

[12] Hachemeister, Charles A., "Credibility for Regression Models with Application to Trend" From CREDIBILITY: THEORY AND APPLICATIONS edited by P.M. Kahn. Copyright 1975 by Academic Press, Inc. Reprinted by permission of the publisher in Casualty Actuarial Society *Forum*, Spring 1992, pp. 307-342.

[13] Halliwell, Leigh, "Loss Prediction by Generalized Least Squares," PCAS LXXXIII, p. 190.

[14] Herbers, Joseph A., "Materiality and Statements of Actuarial Opinion," Casualty Actuarial Society *Forum*, Fall 2002, pp. 103-138.

[15] Kirschner, Gerald S., Kerley, Colin and Isaacs, Belinda, "Two Approaches to Calculating Correlated Reserve Indications Across Multiple Lines of Business," Casualty Actuarial Society *Forum*, Fall 2002, pp. 211-245.

[16] Mack, Thomas, "Distribution-Free Calculation of the Standard Error of Chain-Ladder Reserve Estimates," ASTIN Bulletin, 23, pp. 213-225.

[17] Mack, Thomas, "Measuring the Variability of the Chain Ladder Reserve Estimates," Casualty Actuarial Society *Forum* 1, Spring 1994, pp. 101-182.

[18] Mack, Thomas and Venter, Gary, "A Comparison of Stochastic Models that Reproduce Chain-Ladder Reserve Estimates," Insurance: Mathematics and Economics, 26, pp. 101-107.

[19] Mango, Donald, "Risk Load and the Default Rate of Surplus," CAS Discussion Paper Program, May, 1999, pp. 175-222.

[20] Murphy, Daniel M., "Unbiased Loss Development Factors," PCAS LXXXI, 1994, pp. 154-222.

[21] Philbrick, Stephen, "Accounting for Risk Margins," Casualty Actuarial Society *Forum*, Spring 1994, Vol. I, pp. 1-87.

[22] Sherman, Richard E., "Extrapolating, Smoothing and Interpolating Development Factors," PCAS LXXI, pp. 122-192.

[23] Van Slyke, Oakley E., "Regulatory Standards for Reserves," Casualty Actuarial Society Discussion Paper Program, May 1987, pp. 368-421.

[24] Venter, Gary G., "Testing the Assumptions of the Age-to-Age Factors," PCAS LXXXV, 1998, Book 1, pp. 807-847.

[25] Weber, Robert A., Van Slyke, Oakley E. and Russo, Guiseppe, "Loss Reserve Testing: Beyond Popular Methods," Casualty Actuarial Society *Forum*, Summer 1997, pp. 381-448.

[26] Wright, T. S., "Probability Distributions of Outstanding Liability from Individual Payments Data," Claims Reserving Manual, Institute of Actuaries, London.

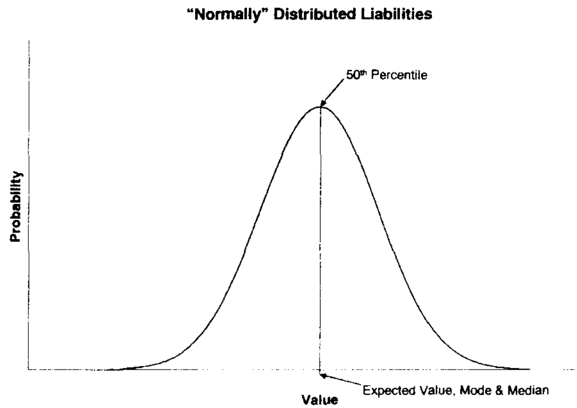
[27] Zehnwirth, Ben, "ICRFS Manual and Users Guide," Version 9.1, Insureware Pty. Ltd.

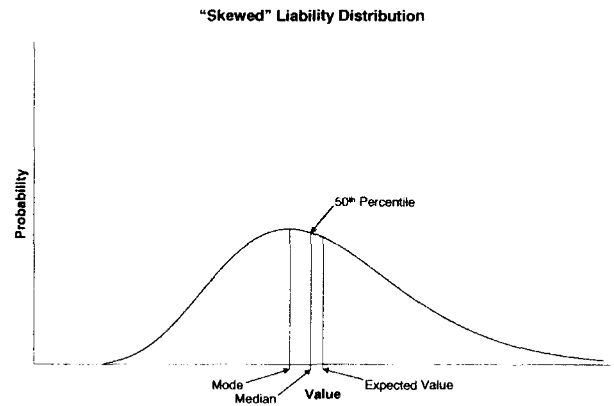
[28] Zehnwirth, Ben, "Probability Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals, and Risk Based Capital," Casualty Actuarial Society *Forum 2*, Spring 1994, pp. 447-606.

LOSS RESERVE ESTIMATES:

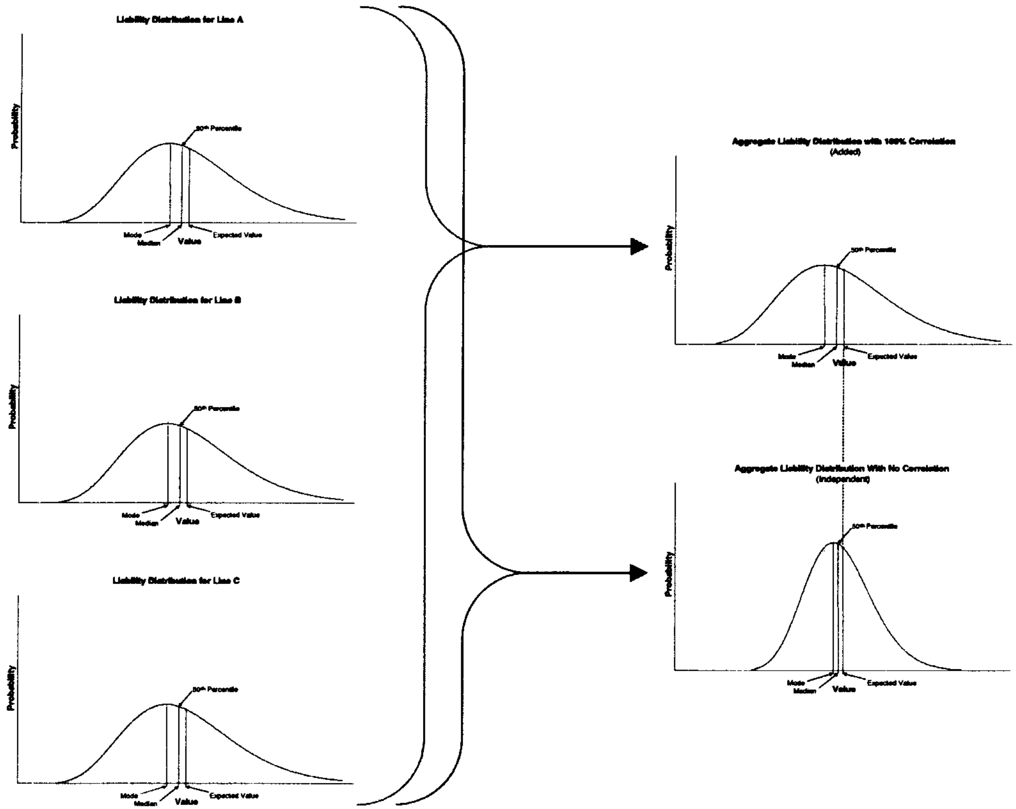
A STATISTICAL APPROACH FOR DETERMINING "REASONABLENESS"

Graph 1: Comparison of "Normal" vs. "Skewed" Liability Distributions

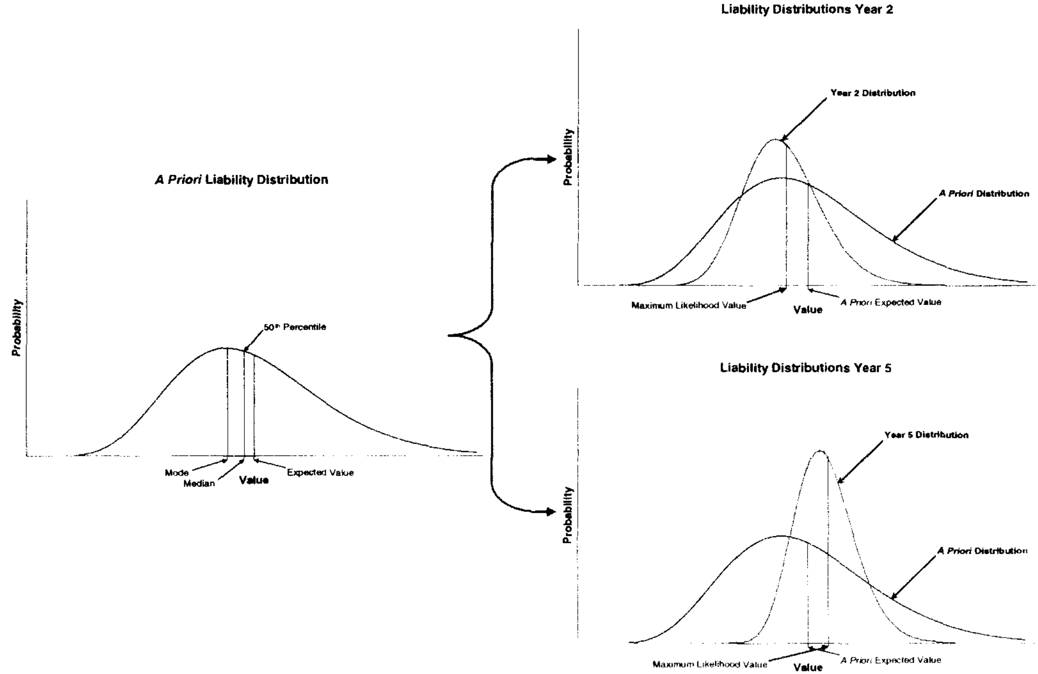




Graph 2: Comparison of Aggregate Liability Distributions



Graph 3: Comparison of *A Priori* vs. Credibility Adjusted Liability Distributions



Prior to the incidence of a cohort of claims, their distribution and expected value can be estimated (the *a priori* distribution). Once the claims have occurred and the settlement process begins, the new estimates of their ultimate value will gradually become more certain over time until all claims are completely settled and their value is known with 100% certainty. As an example illustrated here, in year 2 the claims paid to date are less than anticipated in the first two years and, therefore, the remaining expected value plus the current paid to date results in a new distribution with a total expected value (maximum likelihood) which is less than the *a prior* expected value. Continuing the example, in year 5 the claims paid to date are now greater than anticipated in the first five years and, therefore, the remaining expected value plus the current paid to date results in a new distribution with a total expected value (maximum likelihood) which is greater than the *a prior* expected value. At both 2 years and 5 years, the remaining uncertainty is getting smaller.