Insureware
Software Solutions and eConsulting for P&C Insurance

Variability and Uncertainty

There is an important distinction between variability and uncertainty and the two should not be used interchangeably.

"***Variability is a phenomenon in the physical world to be measured, analyzed and where appropriate explained. By contrast uncertainty is an aspect of knowledge***."

Sir David Cox.

Uncertainty and variability are philosophically very different and it is common for them to be kept separate in risk analyses modeling.

Variability is the effect of chance and a function of the system. It is not reducible through either study or further measurement (may be reduced through changing the system).

Uncertainty is the assessor's lack of knowledge (level of ignorance) about the parameters that characterize the physical system that is being modeled. It is sometimes reducible through further measurement or study.

Uncertainly has also been called "fundamental uncertainly" or "degree of belief".

Simple examples

Suppose a symmetric coin is tossed 100 times and X denotes the no. of heads.

The mean number of heads (the mean of X) is 50. The SD of X is 5. The Binomial probability of each possible outcome of X (0, 1, 2,..100) is known precisely. There is no uncertainty about the coin's variability.

A 100% **confidence interval** for the mean is [50, 50]. There is no uncertainty in the mean and indeed in any of the probabilities of the outcome X. The probability that X=50 (the mean) is approximately 0.08. A 95% **prediction interval** for the outcome X is [40,60]. This 95% prediction interval cannot be shortened.

Suppose we do not know the true probability of a head, p, because the coin is mutilated. Suppose also that before the coin is tossed 100 times, it is tossed 10 times to get an estimate of the probability p. Suppose for the sake of argument 5 heads are observed. Now the estimate of .5 of the probability of a head (in one toss) is uncertain.

We can create a confidence interval (CI) for p, and also for the mean 100p of the number of heads in 100 tosses. The CI is an interval around the estimated mean, namely, 50=100*.5. The CI is not the same as a prediction interval for the outcome in 100 tosses.

A prediction interval includes process variability and is wider than a confidence interval. A confidence interval is for a parameter, which is a constant. A prediction interval is for a random variable. [The loss reserve is a random variable].

Now, a 95% prediction interval for the number of heads in 100 tosses will be wider than [40, 60], say, [35, 65]. This interval can only be reduced to at best [40, 60] by reducing the parameter uncertainty through more (prior) sampling. But you **cannot** make a 95% prediction interval narrower than [40, 60].

Consider another example with the same mean. A symmetric roulette wheel, numbered, 0, 1, 2, 3,....., 100 that is turned only once, and let X be the random variable that represents the outcome. The mean of X is 50, SD is 29. There is no uncertainty about the variability in the outcome X. The probability that X=50 is 1/101.

A 100% **confidence interval** for the mean is [50, 50] (just like the coin). A 95% prediction interval for the outcome X is [2, 97], for example.

Each process (symmetric coin and symmetric roulette wheel) has the same mean, or if you like the same "best estimate". Which one requires more capital?

So "best estimate" is pretty useless and a "range of estimates", also pretty useless, notwithstanding the fact that how do you know if an estimate is "best"?

In general, only in the presence of a probabilistic framework can you assess "best". Indeed it is only in a probabilistic framework that you identify (build) a model that represents the variability in the data.

Probabilistic Trend Family (PTF) of models described in the paper "*Best Estimates for Reserves*"

In the PTF modeling framework a model is identified (built) that quantifies the variability in the data.

Variability is decomposed into Trends + Process Variability. You cannot reduce process variability. You can reduce parameter uncertainty (for the past) by having a larger triangle or a related triangle so that you can do some credibility modeling. Future variability on a <u>log scale</u>= parameter uncertainty + process variability.

In general, parameter uncertainty increases the length of a prediction interval, alternatively, the skewness (both mean and coefficient of variation of the loss reserve distribution). In the case of no parameter uncertainty the prediction distribution's skewness is determined by only the process variability inherent in the data that cannot be reduced.