ESTIMATION OF THE DISTRIBUTION OF REPORT LAGS BY THE METHOD OF MAXIMUM LIKELIHOOD

EDWARD W. WEISSNER

VOLUME LXV

DISCUSSION BY JERRY A. MICCOLIS

Perhaps the most important contribution the actuarial profession can make to the industry which it serves is the representation of complex insurance phenomena by means of coherent mathematical models. The intelligent formulation of a mathematical model tends to strip away much of the mystery surrounding a given insurance problem. It makes explicit the many assumptions that may be taken for granted in less rigorous approaches. It allows the actuary to make verifiable numerical statements about the most convoluted of insurance problems by building in logical progression upon basic mathematical foundations. Most importantly, though, it aids the actuary in his future research by prompting him to ask the correct questions about the issue under study. For these reasons, Ed Weissner's modelling of the report lag phenomenon is a worthy addition to our *Proceedings*.

This reviewer, after making a few (rather pedestrian) comments on some of the technical aspects of the paper, will concentrate on actual applications of the author's model to real-world situations. The reader is urged, while considering the few minor criticisms which follow, not to lose sight of the overall importance of Mr. Weissner's fine paper.

Reinforcing a Point

It should be stressed that while the data used in the formulation of the model is truncated at various report lags, the parameter that is estimated is not only the parameter of the fitted *truncated* distributions but is also the parameter of the fitted *complete* (untruncated) distribution as well. This is an important point. It is one that the author makes but one that, I feel, bears reinforcement. This technique of fitting complete distributions using incomplete (truncated, censored, etc.) data is a powerful one and has found use in other areas of actuarial work.

The Search for a Maximum

The crucial operation in maximum likelihood estimation is the finding of a maximum of the likelihood (or log-likelihood) function. I must admit to a pet peeve here. Several authors, in their search for a maximum (or minimum) of a given function, set the first derivative of the function equal to zero and automatically assume that the root of this equation is the desired maximum (or minimum). This is, of course, not necessarily so. I am afraid the author is guilty of this assumption in the case of both $g(\theta)$ and $g^*(\theta)$. It would have been a minor task to verify that $\hat{\theta}$, in both cases, provides the maximum.

Interestingly, proof exists for the case of $g(\theta)$ in the Figure in the paper. Note that $g(\theta) = Y_1 - Y_2$. Note also that $Y_1 > Y_2$ for $\theta < \hat{\theta}$ and $Y_1 < Y_2$ for $\theta > \hat{\theta}$. Since $g(\theta)$ is the first derivative of the log-likelihood function, we have a maximum because $g(\theta) = Y_1 - Y_2 > 0$ for $\theta < \hat{\theta}$; $g(\hat{\theta}) = 0$; and $g(\theta) = Y_1 - Y_2 < 0$ for $\theta > \hat{\theta}$.

Domain of Convergence

The author mentions the use of Newton-Raphson iteration. In the examples given, swift convergence to a reasonable result was apparently obtained. In some applications, however, divergence, or convergence to the wrong root, may result. I would have preferred that the author had pointed out these potential covergence problems and shared with us any hints he had on the selection of a proper seed.

Goodness of Fit

Once we have decided upon the form of the theoretical distribution we would like to fit to our data, and estimated the parameters of this distribution (by means of maximum likelihood estimation, for example), we should then test how well the distribution fits our observations. The Kolmogorov-Smirnov (K-S) test is a simple, yet powerful, test for this purpose. A description of the K-S test may be found in [1].

In actual applications, other tests should suggest themselves naturally. For example, using the exponential model in Section 2 of the paper, we are able to compute the estimated number of claims emerging during any calendar month. Comparing this number with the actual number of emerged claims during that calendar month (a diagonal in Table II) provides a good practical test of fit.

Sensitivity

All parameter estimation techniques and tests of fit operate on observed data. One of the uses the author suggests for his model is the estimation of claims incurred but not reported (IBNR). IBNR estimation is one example of *projection* based on the model, i.e., using the model to estimate the future unobserved portion of the data. If one of the reasons for developing a model is to use it for projection, then the testing of the model is incomplete unless it includes some form of sensitivity analysis. Utilizing the IBNR example: if the use of a log-normal distribution, say, over an exponential, results in vastly different IBNR estimates, then much more care in the choice of a distribution function is warranted. Perhaps the most appropriate estimate would be a range of values generated by a family of reasonable distributions.

A Word of Caution

The author mentions in his opening sentence that IBNR estimation is aided by knowledge of the report lag distribution. Indeed, he gives an example of IBNR calculation at the end of Section 1. I believe that it is dangerous to apply the model as it stands to the estimation of IBNR claims. This is my only substantial reservation about the paper.

A crucial assumption made in Section 2 of the paper is that θ, and hence the average report lag, is constant by accident month. Let us assume, alternatively, that the average report lag is increasing by accident month. Let us further assume, as did the author, that the number of (ultimate) occurrences is increasing by accident month. It is clear that these two phenomena will tend to significantly increase the actual IBNR over what would be the case if neither trend were present. It should also be clear that each of these trends will successfully mask the other in the data we have available (i.e., data in the form of Table II). In other words, if October occurrences are greater than April occurrences, we will not notice that fact since they will emerge, on the average, at later lags than did the April occurrences, and will more likely fall in the future unobserved region of Table II. But Table II is all the model has to work with! Hence it cannot distinguish between the "double trend" and "no trend" scenarios above. The model will give accurate results for the "no trend" case but will seriously underestimate IBNR in the "double trend" instance.*

With reasonable effort, the author's model can be generalized to accommodate the assumption of changing report lags and changing number of occurrences by accident month. Hints on how to proceed may be found in [2].

A different phenomenon from continuously changing report lags is the case of an abrupt one-time change in average report lag (due to, say, the implementation of on-line computer claims reporting). This would occur during a particular calendar period and would affect all accident months along a Table II diagonal (doing violence to the implicit independence-by-accident-month assumption necessary to the

* The model will also work in the case of varying occurrences and constant θ, as the author has shown.

formulation of the likelihood function L^*). In this case, I would suggest manual adjustment of all data above the diagonal to put it on the new accelerated-reporting basis rather than adapting the model. This would also be the procedure for other non-recurring type phenomena.

Where from Here?

This model is flexible enough that it may also be used to estimate the lag between claim reporting and claim payment. Both of these lag models, in combination with a model describing claim size amounts by occurrence date and payment date, may then be used to build a complete model of the claim payment process.

Conclusion

This is a significant paper. While the comments above argue against the immediate use of the author's unmodified report lag model as a practical tool, the paper remains important in two respects. First, any responsible attempt, such as this, to mathematically model a complex insurance phenomenon should be heartily welcomed by the actuarial fraternity. Second, the specific fitting technique employed (i.e., estimating a complete distribution function with incomplete, biased data) is extremely useful and has much wider application than the estimation of report lags. The recent technical advances in the field of increased limits pricing owe much to this technique.

This paper should provide a firm foundation for the study of report lags; its techniques should find broad application in other areas of actuarial endeavor; and, in prompting actuaries to "ask the right questions," it should enhance the future state of our science.

REFERENCES

[1] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Company, 1956.

[2] Charles A. Hachemeister, "IBNR Claims Count Estimation with Static Lag Functions," presented to Risk Theory Seminar, American Risk & Insurance Association, April, 1975.