Discussion Of
"A Bayesian Credibility Formula For IBNR Counts"

By Gary Venter

Discussion of
A Bayesian Credibility Formula for IBNR Counts

by Gary Venter

If an indicator of a significant paper is that it opens the door for further research, Dr. Robbin's paper should stand the historical test. This review will emphasize generalizing the Poisson assumptions of the paper; attention to optimal parameter estimation and other model assumptions may also prove fruitful, as may the quantification of uncertainty in the IBNR estimates.

The three way credibility weighting for IBNR is an interesting result of the paper. Credibility weights are specified for three estimators of IBNR:

(i) the original (e.g., pricing) expected claims less the observed claims to date

(ii) the observed claims to date times a development factor

(iii) the original expected claims less the expected claims to date.

To see the origin of these credibility weights, a slightly more general framework will be used here. A vector of parameters, u, is postulated to determine the distribution of N, the ultimate number of claims, M, the observed claims to date, and R, the IBNR claims.

It is assumed that M and R are conditionally independent given u. Further, n and q are functions of u, and s^2 is a positive constant with:

$$E(N|u) = n$$
$$E(M|u) = n(1-q)$$
$$E(R|u) = nq$$
$$EV(M|u) = s^2$$

This last assumption generalizes the Poisson assumption of the paper, where the expected conditional variance of M was $EnE(1-q)$.

It is also assumed that u is a vector of random variables such that n and q are independent.

The fundamental credibility formula from Robbin, section III.1, is then invoked to estimate R:

$$R^* = ER + (M-EM)C(M,R)/VM.$$

From the assumptions, $ER = EnEq$ and $EM = EnE(1-q) = En(1-Eq) = En-EnEq$. Also $VM = EV(M|u) + VE(M|u) = s^2 + V(n(1-q)) = s^2 + E(n^2(1-q)^2) - E(n(1-q))^2$. Then by the reasoning of B.3.(ii) of the paper, $VM = s^2 + E(n^2)V(1-q) + E(1-q)^2Vn$.

These three components of the variance of the observed claims, when divided by that variance, will turn out to be the three credibility weights to be applied to the three IBNR estimators (i) - (iii) above. To see this, a general formula on covariances is used to compute C(M,R):

$$C(M,R) = EC(M,R|u) + C(E(M|u),E(R|u)).$$

Because of the conditional independence of M and R, the first term is zero, and so $C(M,R) = C(n(1-q),nq)$

$$= E(n(1-q)nq) - E(n(1-q))E(nq)$$

Then, by the reasoning of B.3.(i) of the paper, $C(M,R) = VnEqE(1-q)-E(n^2)V(1-q)$. Plugging all of this back into the original credibility formula gives:

$$R^* = EnEq + (M + EnEq - En) [VnEqE(1-q) - E(n^2)V(1-q)]/VM.$$

This is regrouped into Robbin's three way credibility formula as follows: first combine the EnEq terms; apply M-En to the second term in brackets to yield $\underline{(En-M)E(n^2)V(1-q)/VM}$. When applied to the first term in brackets the M and En are separated, giving

a) En combined with Eq and adding to the EnEq component; and

b) $\underline{M[Eq/E(1-q)]VnE(1-q)^2/VM}$. The underlined terms are the IBNR estimators (i) and (ii) times credibility weights, where the weights are the second and third components of the variance VM above, divided by VM.

This interprets Eq/E(1-q) as a development factor, and in fact by the hypotheses above, ER/EM = Eq/E(1-q) and EN/EM = 1/E(1-q). This corresponds to the method of estimating LDF's from several accident years' data by $\sum N_i \,/\, \sum M_i$, as recommended by Stanard (PCAS 1985). With this definition of the LDF, the mathematically imprecise estimate of the development factor used by Dr. Robbin becomes unnecessary.

Finally the remaining terms of R^* can be algebraically combined to yield the credibility weight of s^2/VM applied to EnEq. Writing EnEq as En - EnE(1-q) shows this term to be the original expected claims less the expected claims to date.

The assumption that M and R are conditionally independent may be somewhat limiting. The possibility that some claims come in earlier than usual, so fewer come in later, or vice versa, suggest that R and M are not unconditionally independent. Assuming they are conditionally independent then attributes their correlation to non-independent parameters. But this suggests that the parameters are different from year to year. If the claims reported before and after a given point are each modelled as conditionally independent draws from a fixed, possibly unknown, report lag distribution, a negative correlation between reported and unreported claims would not be anticipated.

Dr. Robbin is to be congratulated for this thought provoking and potentially useful paper. He has proven his main point: a Bayesian credibility formula for IBNR does count.