DISCUSSION BY ROBERT N. TREMELLING, II

While receiving greater attention in the past few years, credibility standards still lack clarity and structure. This becomes all too apparent when it is realized the standards in common use today are intended for claim frequency only, but are routinely applied to pure premiums. In addition, the assumptions underlying the traditional Poisson claim frequency process are rarely fulfilled. These assumptions include homogeneity of risks and randomness of claims (implying both an "accidental" nature and mutual independence).

Concentrating on claim frequency, Mr. Hansen presents a rationale for the inadequacy of the basic Poisson distribution as a model by attacking the assumption of homogeneity in the risk population. It seems clear that a measure of increased variability must be taken into account if the risks are in fact heterogeneous. The measure described in this paper is a structure function. This, then, becomes the central topic: a standard for full claim frequency credibility through consideration of additional variation inherent in a non-homogeneous population.

The structure function is, of course, of primary importance and should be closely scrutinized. The general form of the gamma function is first developed as a structure function, but later abandoned in favor of the exponential which maximizes the population variability. Mr. Hansen defends the use of the exponential by stating the ideal credibility standards should be "generally conservative". In contrast, I believe the standards should be "generally exact". We do not need a ceiling or a floor, but a correct sample size level. Further, to suggest any one unique structure function for differing lines, covers, deductibles, and territories is rather optimistic.

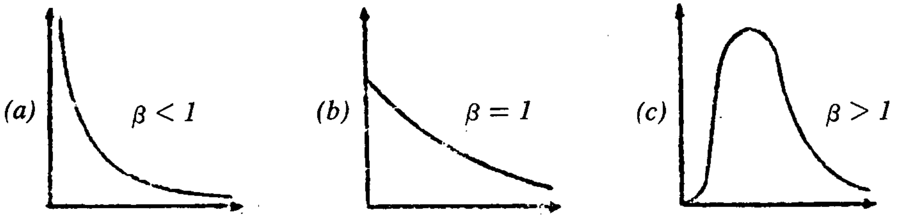
Continuing with the structure function concept, the author states that "gamma distributions which may be used as structure functions have increasing failure rates." In fact, the exponential is a special case of the gamma and has a constant failure rate. The gamma probability distribution function is given by:

$$f(\lambda) = \frac{\alpha^{\beta}}{(\beta - 1)!} e^{-\alpha\lambda} \lambda^{\beta - 1} \qquad \begin{array}{l} 0 < \lambda < \infty \\ \alpha > 0 \\ \beta > 0 \end{array}$$

By setting $\beta = 1$, we have the exponential:

$$f(\lambda) = \alpha e^{-\alpha\lambda} \qquad \begin{matrix} 0 < \lambda < \infty \\ \alpha > 0 \end{matrix}$$

In general, we may summarize the form of the gamma by the following graphs:



(a) Decreasing failure rate

(b) Constant failure rate (Exponential)

(c) Increasing failure rate (forms similar to the Chi-Square or Log-normal)

As Mr. Hansen notes, those distributions asymtotic to the vertical axis, such as graph (a) above, are not intuitively appealing as structure functions. But graph (c) represents the family of gamma distributions which are all useful as possible structure functions. They are only bounded by the exponential. Perhaps consideration should also be given to another distribution with similar characteristics, the Weibull. The Weibull is widely used for real life systems, and has many forms.[1] In fact, the Weibull distributions also becomes the exponential in the special case where $\beta = 1$.

It should further be noted that an implied characteristic of the constant failure rate is randomness, in contrast to the increasing failure rates which follow a specified pattern. In conclusion, much more work needs to be done on estimation of structure function parameters (especially the shape parameter), although I am sympathetic to the difficulties presented.

[1] See for example, Eugene L. Grant, *Statistical Quality Control,* McGraw-Hill, New York, 1964, pp. 505–507.

Turning to other areas presented in this paper, I offer the following comments:

(1) While dealing with heterogeneity of the risk population, Mr. Hansen accepts "the usual assumption of mutual independence among risks." This cannot be overlooked in a truly comprehensive study. Not only is the claim sample inter-related, there is also the possibility of auto-correlation between samples from different time periods. This auto-correlation becomes even more significant when the ratemaker does not have a large sample. The information can become biased under these conditions, depending largely on the extent to which the same policyholders report claims in different sample periods.

(2) New credibility standards should focus on criteria for pure premiums. Even though Mr. Hansen does present a potentially viable technique for claim frequency, claim severity should also be investigated to provide a more complete answer.

(3) The traditional claim number for full credibility given in this paper is 1,082. However, tables such as those used by the Insurance Services Office show 1,084 as that standard. True, the difference is very minor, especially when the true adequate number is no doubt hundreds of claims higher. But at least this basic number should be consistent. Relying on interpretation of the specific equation used, the required number must be "greater than or equal to" 1,082.4. It seems clear that this number should be the next higher integer value, or 1,083. While the difference in actual sample size is highly insignificant, the lack of agreement is significant. The basic calculation is given below.

$$n \geq \left(\frac{Z_{.05}}{\epsilon}\right)^2 \geq \left(\frac{1.645}{.05}\right)^2 \geq 1{,}082.4 \rightarrow 1{,}083$$

Lastly, I would like to propose the technique of stratified sampling as an alternative to an increase in sample size under simple random sampling from heterogeneous populations. If the risk population of a line is not homogeneous, the derived rates are inequitable. Credits and other rating factors are usually applied to correct this shortcoming, implying a definite partitioning scheme for refining the population into homogeneous subclasses. Stratified sampling is specifically designed for application to non-homogeneous populations. It is also one of the most powerful methods for dealing with skewed distributions (fast becoming a major consideration in credibility studies).

Briefly, the population is partitioned by significant differences so as to account for as much of the inherent heterogeneity as possible. Then simple random samples are taken from each subclass and combined by a specific weighting procedure. If the criteria for partitioning is totally without basis, the technique renders the same degree of information as simple random sampling. Effective partitions will bring about a distinct reduction in variability. In other words, a smaller sample size is required under stratified sampling for the same degree of precision as simple random sampling.

In conclusion, Mr. Hansen justly points out that the underlying assumption of risk homogeneity is not generally met. He further provides us with a solution to this shortcoming by introducing a structure function for the risk parameter. Parameter estimation work for the structure function should render this technique viable for claim frequency, although treatment of claim severity is still needed. Rather than increasing the sample size to some specified new level, perhaps serious consideration should be given to variance reduction techniques, such as stratified sampling.