Substituting these values in equations (F) and (E) we obtain $\lambda = 35{,}287$ and 33,258 respectively. The introduction of the third moment here increases the credibility requirement by 6%.

The number of claims required for full credibility under the assumptions above is strikingly reduced by the introduction of the million dollar limitation. The values of λ in fire lines certainly contrast sharply with those in automobile and workmen's compensation found in the original paper. In closing, let me emphasize that the fire loss distribution data found herein is an approximation and should not be considered a precise nor final answer on this subject.

DISCUSSION BY CHARLES C. HEWITT, JR.

This review will have two principal parts:

(1) A focusing of attention upon the recent general definition of credibility by Buhlmann (1), and

(2) A commentary upon the true meaning of "full credibility" in view of the insight that Buhlmann's generalization provides.

(1) Partial Credibility — the Buhlmann Definition

Buhlmann restates the familiar

$$z = \frac{n}{n+K}$$

when n is the number of observations, but goes on to prove that

$$K = \frac{\textit{Expected value of the process variance}^{*}}{\textit{Variance of the hypothetical means}}$$

* This conclusion was reached with respect to both the Gamma-Poisson process and the Beta-Binomial process in Mayerson's earlier work (2) on a Bayesian treatment of credibility, but was not recognized in this most general form by either Mayerson or the author of this current review in his earlier review of Mayerson's Bayesian approach (3). In the latter review this author even went to the trouble of pointing out Albert W. Whitney's fifty-year-old statement (4) of this formulation for the (essentially) Beta-Binomial situation without achieving the insight contained in Buhlmann's analysis. (In failing to recognize K in the Buhlmann format, this reviewer was fooled by his own constant dependence on the Gamma-Poisson process and the coincidence that the mean and variance in the Poisson process are identical.) Finally (for those who prefer numerical values attached to ideas) the Appendix includes an application of the Buhlmann definition to Canadian private passenger auto statistics.

If it is assumed that the *variance of the results of prior observations* is admissible as an estimator for the *expected value of the process variance,* then Buhlmann's work may be generalized far beyond actuarial, mathematical, or even mensural limitations:

$$\text{Credibility} = \frac{(n)\ \text{Number of observations}}{(n)\ \text{No. of observations} + (K)\ \dfrac{\text{Variation in results}}{\text{Variation of hypotheses}}}$$

Take the simple statement of faith, "There is but one God." This permits no variation in hypothesis; therefore the denominator of "*K*" is zero, "*K*" is infinite, and credibility is zero. Thus no observation, no matter how often repeated, can shake the faith of the persons who make this assertion.

There are, then, three variables which can affect credibility:

(i) number of observations,
(ii) variation in results (estimator for process variance), and
(iii) variation of hypotheses (variance of hypothetical means).

Credibility will increase from zero towards unity as:

(i) the number of observations increases, or
(ii) the variance of the results of prior observations decreases, or
(iii) the variance of the hypothetical means increases.

These statements will be illustrated with examples, each slightly more complex and unfamiliar than the preceding. Full credibility occurs —

(i) *When the number of observations increases without limit.*
This is the most easily understood situation — by laymen and mathematicians alike. The classic example is the coin toss in which the proportion of heads or tails becomes more believable the more often the observation is repeated.

(ii) *When there is reason to expect that repeated trials will produce the same result.*
This is the situation in which the immediate observation produces essentially the same result as prior observations. The best known examples are in the physical sciences — the time of rising and setting of the sun and moon, the position in the skies of the planets and stars, the time of high and low tides, the temperature at which water boils or freezes — and so on.

(iii) *When the possible hypotheses are not finite, and one hypothesis is substantially as likely as another.**

This is the situation in which there is no a priori knowledge and no clue as to any favored prediction(s). Because we live in an advanced civilization it is difficult to conjure a good illustration. However, let us use our imaginations to suppose that we are the first sentient beings placed upon Earth, and that awareness occurs for the first time during the night. Imagine our awe when the first sunrise (that we have ever been conscious of) occurs; we do not know whether it is a ball of fire that will be snuffed out in an instant, or on the other hand whether it will remain in the sky forever. Finally, the sun does set on our first observed daytime, but we still don't know whether or not we'll ever see the sun shine again. However, if it does come up again we now think we know how long daylight will last.

(2) Full Credibility — the Classicist's Definition

We have just seen that, lacking (i) an infinite sample, (ii) absolute invariance of results, or (iii) infinite variance of hypotheses, there is no such thing as full credibility. There is a certain percentage of human beings, including a substantial number of mathematicians and actuaries, which finds this thought intolerable.

The classical statistician (Neymann-Pearson School) does not trust a priori judgments, because he says they are "biased" — a word apparently more horrid than "spit." The classicist has achieved a definition of full credibility by a contrived device that runs something like this, "Full credibility exists when an-observation-should-be-within-100k%-of-the-expectation-with-probability, P." But in dodging a priori judgments the classical statistician creates two new parameters k and P, both of which may be varied to suit the judgment or practical necessity of the statistician using them. The sterility of this concept becomes evident when one tries to assign partial credibility, having decided upon full credibility without any real understanding of the meaning of credibility itself. A number of approaches have been

* The essence of these three statements appeared in the *Proceedings* of this Society as long ago as 1950 (5) in a discussion by the late A. L. Bailey. This reviewer was strongly tempted to credit (A. L.) Bailey rather than Bühlmann with the general definition of partial credibility. If this review has erred in giving credit to Buhlmann, it is because the Buhlmann definition is not obscured by the often confusing symbols which the pioneer American actuary unfortunately selected for expressing his (otherwise) lucid thoughts.

tried; one such is the square root of the number of claims (presumably because the variance of the mean increases in proportion to the expected number of claims).

Fortunately the authors of the paper being reviewed here make no claim that their effort is productive of an approach to partial credibility. In fact, early in the paper they point toward the Buhlmann definition. Thus, within the self-imposed restrictions, the Mayerson-Jones-Bowers paper is a worthy attempt to come to grips with the often perplexing problem of assigning full credibility to pure premiums by contemplating both the frequency and severity of claims, and the distributions of frequency and severity. If one "buys" the classical standard for full credibility, referred to in the preceding paragraph, then the authors have achieved their goal of establishing a distribution-free approximation of a standard for full credibility, which utilizes the *relationship* between the higher moments and the mean of the distributions of the number and size of claims. (So we have Mayerson the Bayesian (2) and Mayerson the classicist, and an unregenerate Bayesian may only ask, "Will the real Allen Mayerson please stand up?")

At this point in the discussion it becomes necessary to point to a practical weakness in the solution offered by Mayerson-Jones-Bowers. If one reads this paper carefully he notes that, although the authors emphasize the distribution-free nature of their standard, the three examples which illustrate the standard *all assume specific distributions for the number of claims*. This is *not* merely for convenience, as the authors seem to imply, but a necessary substitute for the fact that one cannot obtain higher moments (than the first) of the distribution of the number of claims without retreating into some specific assumption concerning exposures. Even partitioning the number of claims for a particular risk, or group of risks, on a year-by-year basis (a possible device for estimating higher moments of the number of claims) implies the use of one "risk-year," or "class-year," as an exposure base. Those familiar with workmen's compensation insurance will recognize that even this restriction is not sufficient when the payroll (exposure base) of a risk, or group of risks, fluctuates from one year to the next.

APPENDIX

Buhlmann (1) indicates that the problem of estimating the expected value of the process variance and the variance of the hypothetical means has not yet been attacked. But it has, although Buhlmann would have had no way of finding this out. Using data in his own 1960 paper on Canadian private passenger automobile merit rating (6), this reviewer presented the following estimators at a panel session on credibility in Boston in November, 1965. Rephrased to fit the Buhlmann definition of partial credibility, the data is again presented below:

	Canadian Private Passenger Car Experience			
Classification	Expected Value of Process Variance	Variance of Hypothetical Means	K	z
	(1)	(2)	(3)	(4)
	(Exposure basis—one car year)		(1)/(2)	$\frac{1 \text{ (car year)}}{1+(3)}$
1 — Adult — pleasure use	.087	.00288	30.2	.032
2 — Young driver — limited use	.120	.00337	35.6	.027
3 — Business use	.142	.00487	29.2	.033
4 — Unmarried young owner (or principal operator)	.162	.00599	27.0	.036
5 — Married young owner (or principal operator)	.110	.00263	41.8	.023

The process is Gamma-Poisson as described in detail in (6).

BIBLIOGRAPHY

(1) Buhlmann, Hans — "Experience Rating and Credibility," *ASTIN Bulletin* Vol. IV (1967), pp. 199-207 (particularly p. 207)

(2) Mayerson, Allen L. — "A Bayesian View of Credibility," *PCAS* Vol. LI, pp. 85 et seq.

(3) Hewitt, Charles C., Jr. — Discussion of "A Bayesian View of Credibility," *PCAS* Vol. LII, pp. 121-127

(4) Whitney, Albert W. — "The Theory of Experience Rating," *PCAS* Vol. IV, pp. 274 et seq.

(5) Bailey, Arthur L. — "Credibility Procedures" (shortened title), *PCAS* Vol. XXXVII, pp. 7-23 & discussion thereof with author's reply; ibid pp. 94-115 (particularly p. 114)

(6) Hewitt, Charles C., Jr. — "The Negative Binomial Applied to the Canadian Merit Rating Plan" (shortened title), *PCAS* Vol. XLVII, pp. 55-56.