ESTIMATING PROBABLE MAXIMUM LOSS WITH ORDER STATISTICS

MARGARET E. WILKINSON

VOLUME LXIX

DISCUSSION BY JOHN S. MCGUINNESS

It is always refreshing to see new thinking. This paper introduces to our *Proceedings* an approach—order statistics—that has not before been mentioned there as far as I can find. It would be most welcome if all our papers, as this and many others do, contained clear examples of how to apply practically the new ideas or techniques they propose. Miss Wilkinson is to be complimented for giving us this example of how to use examples.

The paper also has a number of other major values. It points up several important needs and raises several important questions from which we all can learn. The essentials of the paper can be summarized as:

1. It introduces and explains an unfamiliar method which seems to be a generalization of rank correlation and use of dummy variables in correlation problems into a broader realm in which quantitative values of a variable do not exist, or are not for some reason handy to use.
2. It shows the need for, and benefits of, thorough presentation and thorough research.
3. It shows the need by American actuaries to expand their horizons to foreign actuarial work and references.
4. It raises several important questions, the answers to which can be very instructive:
   a. What does "probable maximum loss" mean, most particularly to the author?
   b. What shape does a curve of percentage losses really have?
   c. What work has been done in investigating partial losses and their frequency distribution?
   d. How well do order statistics work with a U-shaped curve and with others that depart materially in shape from that of the normal curve?

The first point needs no further comment, but it will be useful to look at each of the other points in turn, the better to appreciate the paper. Prior to this, however, basics such as the purpose and nature of PML require consideration.

The Purpose of PML

Noting the purpose of computing a PML can be helpful in keeping our eyes on the essentials. As the author mentions, the concept originated in property insurance (more specifically, in connection with fire insurance on fixed-location properties). It was observed that the preponderance of losses were only partial, due in large measure to public fire protection. Total losses were so rare in categories such as protected, fire-resistive structures that, instead of limiting the amount of liability retained on such structures according to the rules followed by ordinary risks, it was safe to anticipate a partial loss and base the retention on that amount instead. A real world PML is thus always relative: to a limit of insurance; to the value of a property; to the value of another insurable interest.

Few underwriters are actuaries,[1] so the concept of probable maximum loss among underwriters has usually been a matter of feel and completely unaided by the concept of a confidence interval. Errors such as a 10 percent estimate of PML on Chicago's late McCormick Place and a 25 percent estimate of PML on the totally destroyed Lake Charles, Louisiana refinery have been the not infrequent result.

The term "PML" appeared only after World War II, although a rough "theory of lines" and elaborate retention schedules or "line sheets" existed in fire insurers prior to 1900. Some time ago the Reinsurance Offices Association in London, after lengthy discussion, recognized the very imprecise nature of PML estimates that are actually used. It has instead standardized on EML (estimated maximum loss), a far more accurate name that is reflective of the judgmental nature of customary practice. The use of EML does not indicate that accurate, fact-based PML's cannot be calculated, but only that they are not being calculated in common practice.

Thorough Presentation and Thorough Research

Frequency distributions of losses by percentage of the limit of insurance are not easily come by. This reviewer first thought that the author had the distinction of obtaining a new one. Nowhere are the data in the paper's Exhibit I labelled as synthetic or hypothetical. The exhibit calls them "Sample Data." The third

[1] One leading underwriter who merited the designation was Benjamin Rush, whose researches substituted facts for underwriting feelings, and thereby changed the marine business of the Insurance Company of North America from a disaster to a profitable operation, and revolutionized ocean marine insurance ratemaking in the process. Another was Francis C. Moore of the Continental Insurance Company, whose *Fire Insurance and How to Build* (New York: Bacon and Taylor Company, 1903) was for fifty years the leading work on how to set underwriting retentions.

paragraph in the section on application of order statistics states "Exhibit I contains a list of 100 claims that are representative of a particular problem in which a PML estimate is needed." But the distribution in the paper was seen to be quite different from any that the reviewer had ever encountered, so the author was queried as to the source of the "data." The author was completely forthcoming in response and acknowledged that they are not real data but synthetic numbers derived by assuming a particular (humped) frequency distribution. Clear and complete labelling should be a *sine qua non*. All evidence the reviewer has seen indicates that the author's assumption is wrong, thus throwing into question the results and conclusions of the paper. Checking easily available references could supply real data that can test these conclusions.

Insurance to Value by Dr. George Head has been in our syllabus of required reading for examinations for several years. Dr. Head cites a paper[2] by one of our charter members, Professor A. W. Whitney, and quotes from it some actual data of the type needed.[3] A second ready source is Ruth Salzmann's "Rating by Layer of Insurance" in our 1963 *Proceedings*.[4] A third source that was not available to Miss Wilkinson when she wrote her paper is one by Gunnar Benktander, presented to the 1982 ASTIN meeting.[5] Although residing in Europe, Dr. Benktander had been able to find Miss Salzmann's paper and referred to it with admiration.

Foreign Actuarial Work and Reference

All of the references in the CAS Syllabus are to works published in America. None of the American or foreign contributions to the *Transactions* of the International Congresses of Actuaries, the ASTIN *Bulletin,* the proceedings of foreign national actuarial bodies, or foreign books, are being used. There is much material in these other publications that could be of real value to our students, even as much of the work in our *Proceedings* is being read and used by foreigners. Miss Wilkinson's paper shows that use by this Society of foreign references as study materials would acquaint our newer members with sources

[2] A. W. Whitney, "The Actuarial Theory of Fire Insurance Rates as Depending on the Ratio of Insurance to Sound Value, Hence the Determination of the Rates for Use With the Coinsurance Clause," *Transactions,* VI International Congress of Actuaries, Vol. 2, 1909, pp. 395–403.

[3] G. Head, *Insurance to Value*, (Homewood, Illinois: Richard D. Irwin, Inc., 1971) pp. 83–88.

[4] R. Salzmann, "Rating by Layer of Insurance," *PCAS* L (1963), pp. 15–26.

[5] G. Benktander, "First and Second Risks," 1982 ASTIN Colloquium, Liège.

they do not know about, and would lead ultimately to a considerably broadened perspective within our membership.

Defining Probable Maximum Loss

The author starts her paper by saying: "In the past there has been much discussion about the definition of probable maximum loss (PML), but little attention has been given to its quantification." She quotes a little later an integrated set of three definitions that were designed to quantify the term quite precisely and thus to permit derivation of PML estimates directly from measured facts; she then quotes a contrasting definition that ignores facts and the concept of a confidence interval and makes a PML estimate a pure judgment or feeling; she then declines to adopt a definition for her paper. This leaves the reader puzzled as to what she is writing about.

By her later use of confidence intervals, she implicitly seems to adopt the first set, but gives no hint of (1) why she mentioned the underwriter's definition (which cannot serve as a base for her statistical analysis), (2) why in view of this limitation she treats it as of equal importance or validity with the former, (3) why she does not state the definition on which she is basing her paper, and (4) why she is *not* explicitly adopting a definition. This source of confusion for the reader of the paper seems to have adversely affected the paper itself.

Failure to make another fine but important distinction also causes trouble. The statement is made:

> The PML depends upon (i) estimates of the likelihood that losses of various sizes will occur, (ii) the amount of losses and associated probabilities that the insured is willing to accept, and (iii) the amount of losses and associated probabilities that the underwriter is not willing to accept. Thus, the insured and the underwriter can have different estimates of the PML for the same loss exposure.

Correctly stated, this would read:

> The PML depends upon the probabilities that losses of various proportions of the relevant limit of loss will occur, and upon the confidence level selected.

The author's statement ignores the fact that a PML is always relative to some limit of loss; the term cannot have meaning otherwise. It also fails to embrace the fact that a PML is always fact-based and fact-related. *Estimates* of the PML—not the actual PML—are what the author describes in the quote above. Naturally, by making different assumptions and guesses, insured and underwriter will make different *estimates* of PML, but that does not change the

actual PML. The quotation also fails to allow for the fact that the insured and underwriter (or any two persons) can also make different PML estimates simply by having different confidence levels, assuming that both know what a confidence level is and think in such terms, or implicitly without knowing it by using different confidence levels.

It is hoped that the author will in her reply set down precisely what she is writing her paper about: what she is measuring or estimating and calling "PML."

Frequency Distribution of Proportional Losses

Ignoring the essential proportional character of PML has led the author to some questionable conclusions. This proportional character is evidenced by more than the data on losses to fixed-location property that were previously cited. For further example, the existence of "total loss only" insurance on waterborne hulls would make no sense but for the U-shaped curve of proportional partial losses. The land-based fire insurance underwriter's customary assumption that properties not under public fire protection are "total-loss" risks is an exact parallel, based on the observation that if a fire gets beyond a minor stage in such property it is generally extinguished only by the burning of all the combustible material that is present.

Contrasting the numbers used by the author with some of Miss Salzmann's real data can be highly informative, so this has been done in Exhibits I and II. The table in Exhibit I sets side by side the number of claims, total monetary amount of claims, and decimal fraction of claim dollars in each class. Each class contains claims of a particular size. Size is shown as a proportion of the limit of insurance ("insured value"), consistent with the purpose of determining a PML. Miss Wilkinson's top loss is, for want of a stated loss limit, used therefor. The statistics from Miss Salzmann's first table, covering protected frame homes, are used, although any of her other three sets or those from the other references could be.

The great difference between the two frequency distributions is apparent. The graph in Exhibit II presents the contrast pictorially. These exhibits show the basis for the curves given in one reference Miss Wilkinson cited[6] and the basis for the cautions appearing in the same source about setting confidence intervals,[7] both of which seem to have escaped her attention.

[6] J. S. McGuinness, "Is Probable Maximum Loss (PML) a Useful Concept?", *PCAS* LVI (1969), pp. 34–35.

[7] *Ibid.*, pp. 32–33.

Order Statistics and Non-Normal Distributions

The paper is based on the assumption of at least asymptotic normality in the underlying data. All available evidence indicates this assumption to be incorrect, so the paper needs to be reworked.

One fact that the Salzmann data reveal is a hump at the left extreme of the data when they are finely enough divided there (in this case, by tenths of one per cent). This would suggest the possibility of a Poisson distribution were it not for the rising right-hand tail.[8] The Salzmann data were, unfortunately, not split by individual percentages between 90 and 100 per cent, but clearly would display the typical rise were this done; only the extent is left unknown.

Since proportions of a single limit are the relevant numbers in determining PML, assuming them to be mutually independent seems at least questionable. Robert Hurley's cautions (in the first four paragraphs of his review of the Salzmann paper)[9] about dealing with this type of data are well taken and to the point here.

In her Exhibit III Miss Wilkinson acknowledges that at least four of her six estimates are not distribution-free. Consequently, at least these are made erroneous by the incorrect assumption of normal data. The 4th estimates, which exceeds by 71 per cent the upper limit of the numbers she presents as data, is thereby also inconsistent with any rational concept of PML, and any actual data so far revealed.

Potential for Order Statistics

The unrealistic results displayed in the paper are due to faulty data and not necessarily to the use of order statistics. Although order statistics hide information in the data that is relevant and important, they possibly can be useful in work on PML estimates. It is sincerely hoped that the author will accept the challenge to apply them to real world data and let us see in her reply whether she has given us a tool that is both new to us and practically useful. I hope she has.

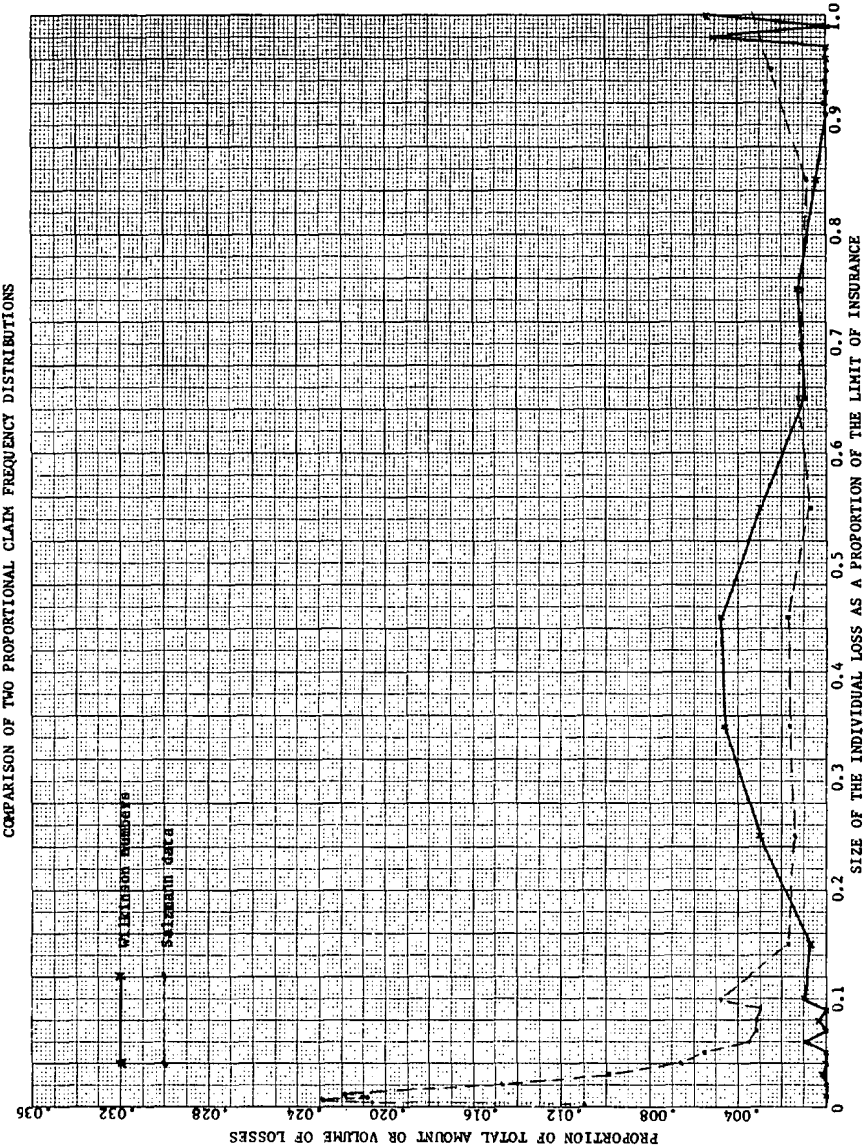
[8] For a relevant discussion of curve shapes see H. Buehlmann, *Mathematical Methods in Risk Theory* (New York: Springer-Verlag, 1970), pp. 4–12.

[9] R. Hurley, Discussion of "Rating by Layer of Insurance," *PCAS* L (1963), p. 27.

EXHIBIT I

COMPARISON OF TWO PROPORTIONAL CLAIM FREQUENCY DISTRIBUTIONS

Upper Class Limit as a % of Insurance Amount	Number of Claims		Amount of Claims in Dollars		Proportion of Total Amount of Claims	
	Wilkinson	Salzmann	Wilkinson	Salzmann	Wilkinson	Salzmann
1	0	3,310	0	194,386	0	.097196
2	0	671	0	146,114	0	.073732
3	0	275	0	98,098	0	.049502
4	1	132	19,874	65,746	.000935	.033177
5	0	86	0	54,913	0	.027710
6	3	46	96,884	35,328	.004559	.017827
7	0	34	0	31,578	0	.015935
8	1	31	40,660	31,793	.001913	.016043
9	0	20	0	30,192	0	.015235
10	2	31	110,051	47,294	.005178	.023865
20	9	94	733,184	168,544	.034499	.035050
30	22	37	3,214,792	135,034	.151269	.068140
40	24	27	4,854,234	155,985	.228412	.078713
50	19	16	4,922,466	168,850	.231622	.085204
60	10	8	3,178,831	72,536	.149577	.036603
70	3	10	1,136,200	122,774	.053463	.061954
80	3	9	1,322,259	104,923	.062218	.052946
90	1	6	482,259	78,378	.022692	.039551
91	0	—	0	—	0	—
92	0	—	0	—	0	—
93	0	—	0	—	0	—
94	0	—	0	—	0	—
95	0	19	0	239,237	0	.120723
96	0	—	0	—	0	—
97	0	—	0	—	0	—
98	1	—	563,899	—	.026534	—
99	0	—	0	—	0	—
100	1	—	576,525	—	.027128	—


COMPARISON OF TWO PROPORTIONAL CLAIM FREQUENCY DISTRIBUTIONS
Pareto Models
Salzburg data
PROPORTION OF TOTAL AMOUNT OR VOLUME OF LOSSES
SIZE OF THE INDIVIDUAL LOSS AS A PROPORTION OF THE LIMIT OF INSURANCE
0
.004
.008
.012
.016
.020
.024
.028
.032
.036
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0