Journal of Business and Social Science Review
Issue: Vol. 3; No.8; August 2022 (pp.47-50)
ISSN 2690-0866(Print) 2690-0874 (Online)
Website: www.jbssrnet.com
E-mail: editor@jbssrnet.com
Doi: 10.48150/jbssr.v3no8.2022.a3

An Early Example of Quantitative Security Analysis from the 1960s

Joel Rentzler
Professor of Finance
Baruch College, CUNY
E-mail: Joel.Rentzler@baruch.cuny.edu

Robert Ferguson
President
AnswersToGo

Abstract

Lacking computers, security analysis prior to the sixties was seldom quantitative. Toward the end of the sixties, time-sharing computers (using teletype machines as terminals and paper tape as backup media) became available. However, usable memory was tiny by today's standards (64 kilobytes was a lot). Basic statistical analysis became feasible, and a few security analysts with quantitative training began to try to gain an advantage over their colleagues by extracting what they viewed as private information from public data. This paper presents one example, an analysis of catastrophic insurance losses in the United States. Hurricane Betsy was an intense tropical cyclone that devastated Florida in 1965. Its destructive path, which caused $1.4B in losses made many security analysts following the insurance industry fearful that such catastrophic losses were a new norm and would adversely impact insurance stocks. This motivated the analysis presented here, which was written as an internal memorandum in 1969 at one buy-side boutique research stock brokerage firm to put the loss from Hurricane Betsy in perspective. The security analyst it was done for felt that this analysis gave him an information advantage versus competing firms without the technical expertise to do this.

Introduction

Insurance is purchased as protection against certain kinds of random events. Examples include losses due to fire, disease, crime, and accident. Policyholders find it desirable to insure against most events that can have a serious detrimental impact on them. When these events occur in large enough numbers, a reasonably accurate prediction can be made of the average per capita loss. This fact enables an insurance company to cover its losses with a high degree of certainty with a premium only moderately above the average loss. These kinds of events seldom produce wide swings in a company's earnings. In the case of infrequent large-scale events such as hurricanes, tornados, and other major catastrophes, an insurance company does not have the same level of protection from the law of averages. For example, in 1965 there were only 4 hurricanes with just one considered a major hurricane (Cat 3+), Hurricane Betsy. Hurricane Betsy made many insurance industry security analysts fearful that future average per capita losses would be much larger than before, adversely impacting insurance stock values. Because of this, it is important to know what the catastrophic losses experienced during a particular period might amount to. Also, to analyze earnings trends, it is necessary to know to what extent single large catastrophic losses are abnormal.

The purpose of this memorandum is to outline a method for determining to what extent historical catastrophic losses have been abnormal and for predicting what they might be in the future.

Definitions and assumptions

A catastrophe is defined as a single event, which produces a total insured loss of at least $1.0 million.[i] The kinds of events, which produce catastrophic losses, include hurricanes, tornadoes, disorders, hailstorms, and windstorms. In this memorandum, it is assumed that the number of catastrophes, which occur during a period, is independent of the number, which occurred prior to that period. It is also assumed that the size of a catastrophe is independent of the number and size of those, which occurred previously.

The size of a catastrophe

Estimating the probability distribution of the size of a catastrophe was accomplished by fitting a probability function to a sample of catastrophic losses. Other things equal, the larger the sample the more accurate the fitting process. In this case, a large sample could be obtained only by combining data from several different years. This is not proper unless the probability distribution is the same for these years. Because it is plausible to expect the size of a catastrophe to have a trend over time, it was necessary to test the reasonableness of assuming that the probability distribution did not change significantly from year to year before combining the data. The process is outlined below.

A record of all catastrophic losses from 1953 to 1966 was obtained from the Insurance Information Institute. The reasonableness of treating all or a major part of this data as coming from the same probability distribution was checked in two ways. First, the K'th order statistic was used as an estimate of the K/(N+1) fractile of each year's cumulative probability distribution (N being the number of catastrophes which occurred during the year). After plotting these estimates on chart paper, a smooth curve was drawn through the points, fitted by eye using french curves, to obtain an initial estimate of each year's cumulative probability distribution.

A visual inspection of the charts revealed little evidence for rejecting the hypothesis that the probability distribution of the size of a catastrophe is constant over the time interval examined. If there were a trend to the size of a catastrophe, the shapes of the curves would be expected to vary from year to year. However, some variation is to be expected due to random effects even if there is no trend. The key is to see if there is more variation between widely separated years than between adjacent years. To test this, one-year and eleven-year gap graphs were compared by eye. This comparison revealed little evidence of a long-term trend.

In view of the result of the visual examination, a formal test of the hypothesis that the probability distribution of the size of a catastrophe is constant over time was made using a non-parametric runs test. This tested the null hypothesis that two samples come from the same distribution. The concept is that, if the observations from two samples from two unknown distributions are combined and arranged in order of increasing size, the number of runs of observations from each sample conveys information about the likelihood that the two distributions are the same. The number of runs tends to be larger when the distributions are the same than when they are different. Here the procedure consisted of combining the catastrophic loss figure for two selected years, arranging them in order of increasing size, and determining whether the number of runs observed was consistent with the hypothesis that the two distributions are the same. This was done for all 91 possible combinations of the 14 years taken two at a time.

Results showing the number of runs actually observed for each pair of years and the expected number of runs under the null hypothesis were obtained. There were times when the null hypothesis, that the two samples came from the same probability distribution, was rejected at a stated level of significance, α. However, since the runs test was applied repeatedly, the occurrence of some rejections of the null hypothesis is to be expected.

Therefore, it is necessary to determine whether the number of rejections is consistent with the null hypothesis that the probability distribution of the size of a catastrophe is constant over time. "The following procedure can always be used to test whether two samples drawn from the same distribution are constant over time," (Mood and Graybill, 1963).

The probability of obtaining exactly k rejections of the null hypothesis at the α significance level in n applications of this test α using a binomial distribution is:

$$\binom{n}{k}\alpha^k(1-\alpha)^{n-k}$$

(1)

Thus, the probability of obtaining at least m rejections of the null hypothesis in n such applications of the binomial distribution is:

$$\sum_{k=m}^{n}\binom{n}{k}\alpha^k(1-\alpha)^{n-k}$$

(2)

An examination of the number of rejections showed that at an α = 0.05 significance level there was a 0.0385 rejection level. This is entirely consistent with the hypothesis that the probability distribution of the size of a catastrophe is constant over time. Therefore it is reasonable to combine all the data from 1953 to 1966.

Next, an initial estimate of the cumulative distribution was obtained by applying the fractile method to the combined data. This resulted in a scatter diagram representing an empirical cumulative probability distribution. Several kinds of probability distributions were fit to this distribution using an empirical maximimum likelihood estimation procedure to find suitable parameters. A gamma distribution was found to fit the emprical cumulative probability curve reasonably well, with the result that the size of a catastrophe is distributed approximately as follows.

$$f(y) = \frac{y^{a}e^{-\left(\frac{y}{b}\right)}}{\Gamma(a+1)b^{(a+1)}}$$

(3)

y = ln (x)

x = Insured loss (millions of dollars).

a = 1.3765

b = 0.6796

The number of catastrophes

The number of catastrophes occurring during a year was assumed to be independent of the number occurring during any prior year. Specifically, the number was conjectured to have a Poisson distribution with the parameter growing at a constant rate from year to year. Parameter estimates were obtained using log-linear regression. The result is that the number of catastrophes occurring during a year has approximately the probability distribution shown below.

$$f(x_i) = e^{-m_i}\left(\frac{m_i^{x_i}}{x_i!}\right)$$
(4)

$$m_i = m_n(1+R)^{(i-n)}$$
(5)

$i \equiv$ The year the distribution applies to.

$x_i \equiv$ The number of catastrophes in year i .

$n \equiv$ 1966

$m_n \equiv$ 14.46

$R \equiv$ 0.02984

The total loss

Knowing the probability distribution of the size of a catastrophe and the probability distribution of the number of catastrophes, it is possible to calculate the expected total loss for a particular year. This is done for the years from 1953 to 1966 and these theoretical losses then were compared with the actual losses to determine the probable error when these distributions are used to forecast future losses. The result is a probabilistic forecast of future total losses obtained by combining the two empirical probability distributions discussed above.

Future Applications

Whether the multiple 20 billion-dollar catastrophes that occurred in the United States in 2021 should be considered outliers or part of a new norm for insurance stocks may be determined using a similar technique applied to more recent data.

References

Mood, Alexander and Franklin Graybill. Introduction to the theory of Statistics. McGraw-Hill Book Company. Second Edition, 1963.

---

[i] The current definition of a catastrophe has increased to $25.0 million