We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 4

An Inequality for Reinsurance Contract Annual Loss Standard Deviation and Its Application

Frank Xuyan Wang

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.76265

Abstract

For reinsurance contract simulated annual losses, an inequality relating their standard deviation and mean is found, $\sigma_f \geq m_f \sqrt{\frac{\mu(A^C)}{\mu(A)}}$, where the coefficient in the inequality is the square root of the ratio of numbers of zero losses years to numbers of non-zero losses years. The largest such coefficient is also proved to be the universal upper bound. As direct application of this inequality, bounds for other risk measures of reinsurance contract, the *TVaR* (average of the annual losses that are larger than a given loss), the probability of attaching (greater than a given attachment loss), and the probability of exceeding (the annual loss limit) are obtained, which in turn reveal the capability upper limit of the simulation approach.

Keywords: reinsurance contract, simulation, standard deviation, coefficient of variation, inequality, ratio distribution, model risk

1. Introduction

In reinsurance industry, simulated losses from catastrophe events combined with reinsurance contract financial terms are used to calculate the contract expected annual loss, the standard deviation of the expected annual loss, and quantiles of the losses (such as the AEP: Aggregate Exceedance Probability, or TVaR: Tail Value at Risk, in Ref. [1]). These numbers are in turn used for pricing or risk management of the contract. There are two kinds of model risks in this approach: independent simulations may give different results without any model change, and simulations vary before or after model change such as from the yearly catastrophe events sets or parameters updates. Empirically, the distribution of the contract expected annual loss may be more like a Beckmann, MaxStable, Gamma, Inverse Gaussian, or even a Lognormal Distribution

IntechOpen

© 2018 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

than a Normal Distribution, but consider that the mean annual loss is the average of a large number of losses, and especially for simplicity, if we assume they obey a Normal Distribution, to quantify the model risk, we can use Hinkley formula [2] or Marsaglia formula [3] to calculate the probability of two simulations have expected annual loss deviated from each other by more than say 50%. Their formulas, in our model risk quantification context and the simplest scenario, depend on only two factors: the correlation coefficient ρ of the distribution in the two simulations, and the coefficient of variation (CV) of the distribution, that is, the ratio of the standard deviation to the mean. Most reinsurance contracts, due to the carefully selected financial terms, have many or the majority of simulated year's losses zero. Thus call for a study of the CV range or bounds of those scarcely payout contracts.

2. Results

2.1. CV range

Starting from Hölder's inequality (https://en.wikipedia.org/wiki/Hölder's_inequality):

$$\|fg\|_1 \leq \|f\|_p \|g\|_q, \tag{1}$$

suppose f in the formula is the contract annual loss with mean m_f deducted, that is, is of the form $f - m_f$ and $p = q = 2$, g is some nonnegative weights on the discrete probability space of the simulated years Ω, for example, the set {1,2,…,100,000} for 100,000 years of simulations, each element with a probability of 1e-5 (a typical setting in practice). Then we get:

$$\int |f - m_f| g d\mu \leq \left[\int |f - m_f|^2 d\mu\right]^{1/2} \left[\int g^2 d\mu\right]^{1/2} = \sigma_f \left[\int g^2 d\mu\right]^{1/2} \tag{2}$$

Suppose f is nonnegative and zero outside of a subset A of Ω and g is constant a on A and constant b on A^C: $f|_A C = 0$, $f|_A > 0$, $g|_A C = b \geq 0$, $g|_A = a \geq 0$. Then we can deduct that:

$$\sigma_f \geq m_f \frac{a\mu(A^C) + b\mu(A^C)}{\sqrt{a^2\mu(A) + b^2\mu(A^C)}} \tag{3}$$

due to

$$|f - m_f| \geq f - m_f \tag{4}$$

and

$$\int |f - m_f| g d\mu = \int_A |f - m_f| a d\mu + \int_{A^C} m_f b d\mu \tag{5}$$

$$\geqslant \int_A f a d\mu - \int_A m_f a d\mu + m_f b\mu(A^C) \tag{6}$$

http://dx.doi.org/10.5772/intechopen.76265

$$= m_f a - m_f a\mu(A) + m_f b\mu\left(A^C\right) \tag{7}$$

The maximum of the right-hand side of Eq. (3) is achieved by $\frac{a}{b} = \frac{\mu\left(A^C\right)}{\mu(A)}$, increasing on $[0, \frac{\mu\left(A^C\right)}{\mu(A)}]$ and decreasing on $\left[\frac{\mu\left(A^C\right)}{\mu(A)}, \infty\right)$, and then we get:

$$\sigma_f \geq m_f \sqrt{\frac{\mu\left(A^C\right)}{\mu(A)}} \tag{8}$$

As a corollary, if we use the 2-sigma or 3-sigma rule for the confidence interval of the estimate of the mean, for those contracts that have most years 0 losses, these interval will be very large. For example, if we have only 100 years have nonzero losses out of the 100,000 simulated years, then we will get [math]::sqrt((1e5-100)/100) = 31.6069612585582:

$$\sigma_f \geq 31.6069612585582 m_f \tag{9}$$

And if we have 1000 years nonzero losses, we get the constant [math]::sqrt((1e5-1000)/1000) = 9.9498743710662:

$$\sigma_f \geq 9.9498743710662 m_f \tag{10}$$

Checking against some concrete examples, for a contract we see 2220 nonzero losses records, suppose all of them are in different years, then we get the ratio [math]:: sqrt((1e5-2220)/2220) = 6.63664411016931, and we have its mean loss m_f = 3848 and standard deviation σ_f = 37,149 from the simulation, 37,149/3848 = 9.65410602910603 > 6.63664411016931.

Another contract have unique 41,143 nonzero losses years out of the 100,000 simulated years, with mean loss m_f = 1,874,487 and standard deviation σ_f = 2,357,787, σ_f/m_f = 1.25783054243641 > [math]:: sqrt((1e5-41,143)/41,143) = 1.19605481319037.

From these examples, we see that our lower bound formula for *CV* is relatively close.

Typically, more than half of the contracts may have nonzero losses years count below 10,000, for which we can see their $\sigma_f \geq 3m_f$, since [math]::sqrt((1e5-1e4)/1e4) = 3. More than 80% of the contracts may have nonzero losses years count below 50,000, for which there is the simple inequality $\sigma_f \geq m_f$, since sqrt((1e5-5e4)/5e4) = 1. These bounds are collected in **Table 1**.

This inequality Eq. (8) can explain the observation that when we sort the contracts by the mean annual loss, the lower quarter of contracts may have more than tens of percent deviation from different simulations, since smaller mean loss usually corresponding to fewer years of nonzero losses and higher *CV* (more explanation in the following section).

To get an upper bound for *CV*, suppose the total simulated years is n and each year has the loss $x_i \geq 0$. Then:

$$CV \equiv \frac{\sigma_f}{m_f} = \sqrt{n \frac{\sum_{i=1}^{n} x_i^2}{\left(\sum_{i=1}^{n} x_i\right)^2} - 1} \tag{11}$$

For the expression $\frac{\sum_{i=1}^{n} x_i^2}{\left(\sum_{i=1}^{n} x_i\right)^2}$, by taking the partial derivative with respect to x_1, we can know that it is decreasing in $\left[0, \frac{\sum_{i=2}^{n} x_i^2}{\sum_{i=2}^{n} x_i}\right]$ and increasing in $\left[\frac{\sum_{i=2}^{n} x_i^2}{\sum_{i=2}^{n} x_i}, \infty\right)$. So the maximum must be attained at either 0 or ∞, that is, is the value of $\frac{\sum_{i=2}^{n} x_i^2}{\sum_{i=2}^{n} x_i}$ or 1. Recursively, we know that the final maximum is 1. So we get:

$$CV \leq \sqrt{n-1} \tag{12}$$

For the extreme case when only one year have nonzero losses, we thus verified that the coefficient [math]::sqrt((1e5-1)/1) = 316.226184874055 is exact, that is, we have:

$$\sigma_f = 316.226184874055 m_f \tag{13}$$

Nonzero losses years	***CV* lower bound** $\sqrt{\frac{\mu(A^c)}{\mu(A)}}$
1	316.226184874055
10	99.9949998749938
100	31.6069612585582
1000	9.9498743710662
2000	7
5000	4.35889894354067
10,000	3
20,000	2
30,000	1.52752523165195
40,000	1.22474487139159
50,000	1
80,000	0.5
90,000	0.333333333333333
99,000	0.100503781525921
99,900	0.0316385998584166
99,990	0.0100005000375031
99,999	0.00316229347167527

For given years of nonzero losses out of 100,000 simulated years, the lower bound as given by our formula Eq. (8).

Table 1. *CV* lower bound.

The overall upper bound $\sqrt{n-1}$ for CV is approachable when we let all but one of the x_i arbitrarily close to 0.

From the fact that the minimum of the expression in Eq. (11) is attained at $\frac{\sum_{i=2}^{n} x_i^2}{\sum_{i=2}^{n} x_i}$. we can see that:

$$\frac{\sum_{i=1}^{n} x_i^2}{\left(\sum_{i=1}^{n} x_i\right)^2} \geq \frac{\frac{\sum_{i=2}^{n} x_i^2}{\left(\sum_{i=2}^{n} x_i\right)^2}}{\frac{\sum_{i=2}^{n} x_i^2}{\left(\sum_{i=2}^{n} x_i\right)^2} + 1} \tag{14}$$

More generally, if year *i*s have possibly unequal probability p_i of occurrence (such as when using importance sampling), then:

$$\frac{\sum_{i=1}^{n} x_i^2 p_i}{\left(\sum_{i=1}^{n} x_i p_i\right)^2} \geq \frac{\frac{\sum_{i=2}^{n} x_i^2 p_i}{\left(\sum_{i=2}^{n} x_i p_i\right)^2}}{p_1 \frac{\sum_{i=2}^{n} x_i^2 p_i}{\left(\sum_{i=2}^{n} x_i p_i\right)^2} + 1} \tag{15}$$

The minimum value of the right side is attained when $x_1 = \frac{\sum_{i=2}^{n} x_i^2 p_i}{\sum_{i=2}^{n} x_i p_i}$. This can be used inductively to give an "elementary" proof of our lower bound results, and additionally can show that the lower bound is attained when all the nonzero losses are equally valued which using the Hölder's inequality cannot arrive. Similarly we can show that the upper bound is $\sqrt{\max\left(\frac{1}{p_1}, \frac{1}{p_2}, \ldots, \frac{1}{p_n}\right) - 1}$.

We summarize our deduction and discussion into the following:

Theorem 1. *For reinsurance contract simulated annual losses f, the standard deviation σ_f with respect to the mean m_f is bound below by:*

$$\sigma_f \geq m_f \sqrt{\frac{\mu\left(A^C\right)}{\mu(A)}} \tag{16}$$

where $\mu(A^C)$ and $\mu(A)$ are the measure of the numbers of zero losses years and the numbers of non-zero losses years, respectively. The lower bound $\sqrt{\frac{\mu(A^C)}{\mu(A)}}$ is attended:

$$\sigma_f = m_f \sqrt{\frac{\mu\left(A^C\right)}{\mu(A)}}, \tag{17}$$

if and only if all the non-zero losses are of the same value. The standard deviation σ_f with respect to the mean m_f is bound above by:

$$\sigma_f \le m_f \sqrt{\max\left\{\frac{1}{p_1}, \frac{1}{p_2}, \ldots, \frac{1}{p_n}\right\} - 1} \tag{18}$$

where the p_i is the probability of occurrence of year i. The upper bound is attained if and only if the smallest occurrence probability year is the only year of non-zero losses. And when only year i have nonzero losses:

$$\sigma_f = m_f \sqrt{\frac{1}{p_i} - 1} \tag{19}$$

For not necessarily nonnegative loss contracts (such as contracts with complex layers structure and hedging design), and for contracts that have significant concentration on the upper bound (due to limit and annual limit), replacing f by $f - m$ or $M - f$, where m and M are the minimum and the maximum annual loss, from the theorem we get the following lower bounds:

Corollary 1. *For arbitrary reinsurance contract simulated annual losses f, the standard deviation σ_f with respect to the mean m_f, minimum annual loss m, and maximum annual loss M, is bound below by:*

$$\sigma_f \ge (m_f - m) \sqrt{\frac{\mu(L^C)}{\mu(L)}} \tag{20}$$

$$\sigma_f \ge (M - m_f) \sqrt{\frac{\mu(U^C)}{\mu(U)}} \tag{21}$$

where $\mu(L^C)$ and $\mu(L)$ are the measure of the numbers of minimum losses years and the numbers of not-minimum losses years, $\mu(U^C)$ and $\mu(U)$ are the measure of the numbers of maximum losses years and the numbers of not-maximum losses years, respectively. The equality hold if and only if f is a bivalued distribution.

From Theorem 1, we can get an upper bound for the average annual loss on an arbitrary subset of the years:

Corollary 2. [1]*For a nonnegative random variable f on a probability space Ω, an arbitrary subset $B \subset \Omega$, the average $\frac{\int_B f d\mu}{\mu(B)}$ is bound above by the standard deviation σ_f and the mean m_f by:*

$$\frac{\int_B f d\mu}{\mu(B)} \le \sigma_f \sqrt{\frac{\mu(B^C)}{\mu(B)}} + m_f. \tag{22}$$

Proof:

Define two functions f_1 and f_2 from f such that they are the restrictions of f on the subset B and B^C: $f_1|_B C = 0$, $f_1|_B = f|_B$, $f_2|_B C = f|_B C$, $f_2|_B = 0$. Then we have $f = f_1 + f_2$ and $f_1 f_2 = 0$. The standard deviation:

[1]The nonnegative condition can be relax to f is bound below.

$$\sigma_f^2 = E\left[\left(f_1 + f_2\right)^2\right] - \left[E\left(f_1 + f_2\right)\right]^2 \tag{23}$$

$$= \sigma_{f_1}^2 + \sigma_{f_2}^2 - 2m_{f_1}m_{f_2} \tag{24}$$

$$\geq m_{f_1}^2 \frac{\mu\left(B^C\right)}{\mu(B)} + m_{f_2}^2 \frac{\mu(B)}{\mu\left(B^C\right)} - 2m_{f_1}m_{f_2} \tag{25}$$

$$= \left(m_{f_1}\sqrt{\frac{\mu\left(B^C\right)}{\mu(B)}} - m_{f_2}\sqrt{\frac{\mu(B)}{\mu\left(B^C\right)}}\right)^2 \tag{26}$$

from Theorem 1 and the fact that the domain with zero value for f_1 include the set B^C and the domain with zero value for f_2 include the set B. Hence:

$$m_{f_1}\sqrt{\frac{\mu\left(B^C\right)}{\mu(B)}} \leq m_{f_2}\sqrt{\frac{\mu(B)}{\mu\left(B^C\right)}} + \sigma_f \tag{27}$$

The inequality Eq. (22) is arrived by the fact that $m_f = m_{f1} + m_{f2}$ and

$$\sqrt{\mu(B)}\sqrt{\mu\left(B^C\right)}\left(\sqrt{\frac{\mu\left(B^C\right)}{\mu(B)}} + \sqrt{\frac{\mu(B)}{\mu\left(B^C\right)}}\right) = \mu\left(B^C\right) + \mu(B) = 1 \tag{28}$$

□

If we let the subset B be $\{x | f(x) > 0\}$, then Eq. (22) become Eq. (16). If we let the subset B be $\{x | CDF_f(x) \geq q, 0 \leq q \leq 1\}$, we get the so called AEP TVaR upper bound for the given quantile q or the return period $r \equiv \frac{1}{1-q}$: $TVaR(q) \leq \sqrt{\frac{q}{1-q}}\,\sigma_f + m_f$. For the usually used return period, the $TVaR$ upper bound (now simply $\sigma_f\sqrt{r-1} + m_f$) is in **Table 2**.

Numerical example shows that our $TVaR$ upper bound is relatively close in the quantile range [0.8,0.9], with the theoretical upper bound deviated from the simulated value by less than 20%, no matter what the distribution of the annual loss is.

Notice that the measure of the numbers of nonzero losses years is also called the probability of attaching in insurance, we can rearrange the terms in the formula Eq. (8) to get a lower bound for the probability of attaching:

Corollary 3. *For a reinsurance contract simulated annual losses f, the probability of attaching, ProbA* $\equiv Prob\{f > 0\}$, *with respect to the CV is bound below by:*

$$ProbA \geq \frac{1}{CV^2 + 1} \tag{29}$$

As an application of Corollary 3, we see that if $CV \leq 3$, then $ProbA \geq 0.1$, the 0.9 quantile of f is larger than zero. Equivalently, if the 0.9 quantile of f (the so called AEP in insurance) is zero, we

Return period	Quantile	*TVaR* upper bound $\sigma_f \sqrt{\frac{\mu(B^C)}{\mu(B)}} + m_f$
100,000	0.99999	$316.226184874055\sigma_f + m_f$
10,000	0.9999	$99.9949998749938\sigma_f + m_f$
5000	0.9998	$70.7036066972541\sigma_f + m_f$
1000	0.999	$31.6069612585582\sigma_f + m_f$
500	0.998	$22.3383079036887\sigma_f + m_f$
250	0.996	$15.7797338380595\sigma_f + m_f$
200	0.995	$14.1067359796659\sigma_f + m_f$
100	0.99	$9.9498743710662\sigma_f + m_f$
50	0.98	$7\sigma_f + m_f$
30	0.966666666666667	$5.3851648071345\sigma_f + m_f$
25	0.96	$4.89897948556636\sigma_f + m_f$
20	0.95	$4.35889894354067\sigma_f + m_f$
10	0.9	$3\sigma_f + m_f$
5	0.8	$2\sigma_f + m_f$
4	0.75	$1.73205080756888\sigma_f + m_f$
2	0.5	$1\sigma_f + m_f$

For given year of return period, the upper bound as given by our formula Eq. (22).

Table 2. *TVaR* upper bound.

CV upper bound	*ProbA* lower bound $\frac{1}{CV^2+1}$	Beginning quantile with nonzero loss	Return period $\left(\equiv \frac{1}{ProbA}\right)$
99.9949998749938	0.0001	0.9999	10,000
9.9498743710662	0.01	0.99	100
3	0.1	0.9	10
2	0.2	0.8	5
1.73205080756888	0.25	0.75	4
1	0.5	0.5	2
0.577350269189626	0.75	0.25	1.33333333333333
0.5	0.8	0.2	1.25

For given range of *CV*, the lower bound as given by our formula Eq. (29)

Table 3. *ProbA* lower bound.

know $CV > 3$: then we will less prone to think that those zero quantiles is due to simulation inaccuracy. The *CV* bounds for commonly used *AEP*, related to probability of attaching by the formula $\sqrt{\frac{1}{ProbA} - 1}$, is in **Table 3**.

http://dx.doi.org/10.5772/intechopen.76265

Similarly, from Corollary 2, we can easily rearrange terms to get an upper bound for the probability of exceeding a given loss, which is called the Cantelli's inequality in the literature (https://en.wikipedia.org/wiki/ Chebyshev's_inequality):

Corollary 4. *For a reinsurance contract simulated annual losses f, the probability of exceeding a given loss x, ProbE ≡ Prob{f ≥ x}, with respect to the mean loss m_f and the standard deviation σ_f, when $x \geq m_f$, is bound above by:*

$$ProbE \leq \frac{1}{\frac{(x-m_f)^2}{\sigma_f^2}+1} \tag{30}$$

Specifically, if $x = \frac{\sigma_f^2}{m_f} + m_f$, then:

$$Prob\left\{f \geq \frac{\sigma_f^2}{m_f} + m_f\right\} \leq \frac{1}{CV^2+1} \tag{31}$$

This bound gives a limitation on simulation with a given number N of simulated years where each year have equal probability of occurrence $\frac{1}{N}$. If $\frac{1}{CV^2+1} < \frac{1}{N}$, that is, $CV > \sqrt{N-1}$, then in theory no simulated loss can reach to $\frac{\sigma_f^2}{m_f} + m_f$, the lowest permissible exposure to allow the given mean m_f and given standard deviation σ_f (to be shown in Lemma 1). In other words, if $CV > \sqrt{N-1}$, no such simulation can match both the given mean and the given standard deviation closely (see also inequality Eq. (12)).

Lemma 1. *For a reinsurance contract simulated annual losses f that are bound up by M and with a given mean loss m_f and a given standard deviation σ_f, we must have:*

$$M \geq \frac{\sigma_f^2}{m_f} + m_f \tag{32}$$

On the other hand, with the given max loss M and mean loss m_f, the standard deviation σ_f must satisfy:

$$\sigma_f \leq \sqrt{(M-m_f)m_f} \tag{33}$$

The maximum standard deviation given m_f and M is attained only by a bivalued distribution of values either 0 or M, with probability $q \equiv 1 - \frac{m_f}{M}$ and $p \equiv \frac{m_f}{M}$, respectively, whose CV is then $\sqrt{\frac{M}{m_f}-1} = \sqrt{\frac{1}{p}-1}$. Similarly, the minimal exposure given m_f and σ_f is attained only by a bivalued distribution of values either 0 or $\frac{\sigma_f^2}{m_f} + m_f$, with probability $q \equiv \frac{\sigma_f^2}{\sigma_f^2+m_f^2}$ and $p \equiv \frac{m_f^2}{\sigma_f^2+m_f^2}$, respectively, whose CV is then $\frac{\sigma_f}{m_f} = \sqrt{\frac{1}{p}-1}$.

Proof: Let $g = \frac{f}{M}$, then g is a random variable with values in interval [0,1]. So:

$$g^2 \le g \tag{34}$$

and

$$\int_\Omega g^2 d\mu \le \int_\Omega g d\mu \tag{35}$$

$$\int_\Omega g^2 d\mu - \left(\int_\Omega g d\mu\right)^2 \le \int_\Omega g d\mu - \left(\int_\Omega g d\mu\right)^2 \tag{36}$$

$$\sigma_g^2 \le m_g(1 - m_g) \tag{37}$$

$$\frac{\sigma_f^2}{M^2} \le \frac{m_f}{M}\left(1 - \frac{m_f}{M}\right) \tag{38}$$

This proves both of our inequalities. Without loss of generality, suppose any nonempty subset of Ω have nonzero measure, the equality hold in Eq. (34) and its subsequent inequalities if and only if $g = 0$ or $g = 1$. □

Because of the probability of $\frac{m_f}{M}$ of taking value M, we cannot solve the limitation on *CV* by increasing M. The only solution is then by increasing N or using unequal probabilities (please refer to Eq. (18)), otherwise we may have to choose to only match the mean loss, and reduce the simulated standard deviation.

By examining the proof of Corollary 2 and Theorem 1 Eq. (17), forcing the inequality in Eq. (25) to be an equality, we can prove that:

Corollary 5. *For a nonnegative random variable f on a probability space Ω, an arbitrary subset $B \subset \Omega$, assuming $\sigma_f > 0, \mu(B) > 0, \mu(B^C) > 0, m_f > 0$, the average $\frac{\int_B f d\mu}{\mu(B)}$ attain its upper bound with respect to the standard deviation σ_f and the mean m_f:*

$$\frac{\int_B f d\mu}{\mu(B)} = \sigma_f \sqrt{\frac{\mu(B^C)}{\mu(B)}} + m_f \tag{39}$$

if and only if f is a nonzero constant function on the subset B and a constant function on the subset B^C.

So the maximum *TVaR* distribution is bivalued, this corollary provides a guide for implementing relatively high *CV* distribution simulation: for *CV* close to $\sqrt{N-1}$ which do not simulate well, a conservative and simple selection is using bivalued distribution. Similar conclusion about bivalued distribution can be made for the maximum *AEP* distribution which are bound by *TVaR's* bound and attain the same upper bound given in Eq. (39). Both conclusions give clue for a risk measure of the maximally likely or best compromise quantile by comparing simulated *TVaR* or *AEP* with the theoretical bound for the best match but that is the topic of a different research.

http://dx.doi.org/10.5772/intechopen.76265

2.2. Simulation deviation

Typical correlation coefficient ρ from yearly model update range from 0.27 to 0.96, the *CV* range is 0.003–316 as we calculated in **Table 1**. With these parameters we used Mathematica Manipulate function to explore the probability of the ratio of two simulated annual-mean-loss be within the range of 0.5–1.5, assuming the annual-mean-loss obeys the Normal Distribution. We find that the probability is small when ρ is close to 0, and is decreasing when *CV* is increasing, but is stabilized after $CV \geq 7$. For an example $\rho = 0.822434$, for almost half of the contracts, the simulated annual-mean-loss being within 50% to each other has probability of 0.459, that is, with probability of 0.551 we will see two simulation have simulated annual-mean-loss increased or decreased by more than 50% (**Figures 1** and **2**). These factors should be considered for model risk management or individual contract evaluation.

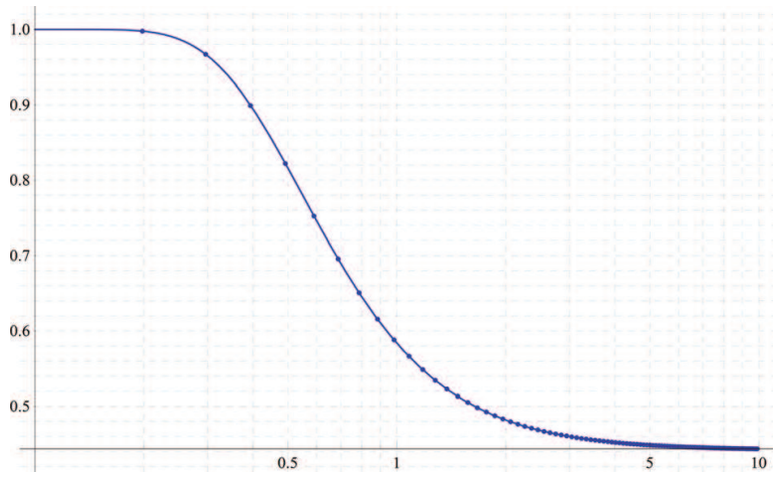


Figure 1. Probability that two simulated mean are within 50% of each other. The horizontal axis is the *CV* and the vertical axis is the probability. (a) Using Hinkley formula. Formulas used in (a) is complex and very different which may cast doubt for their validity. Our plot is a numerical assurance for their correctness.

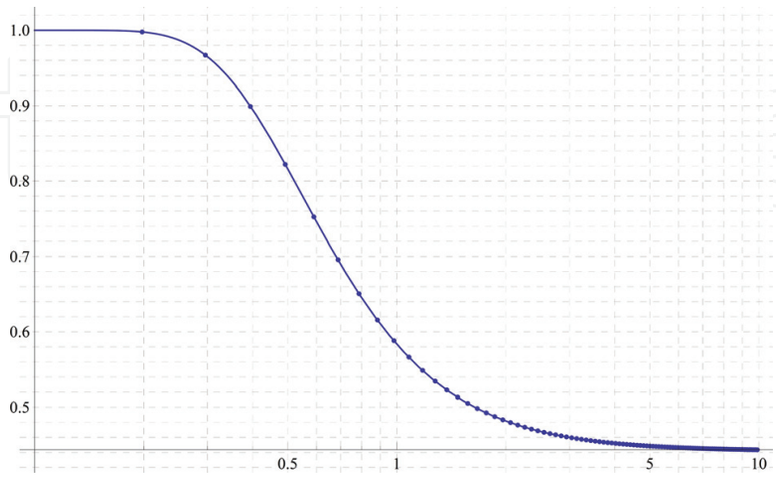


Figure 2. Probability that two simulated mean are within 50% of each other. The horizontal axis is the *CV* and the vertical axis is the probability. (b) Using Marsaglia formula. Formulas used in (b) is complex and very different which may cast doubt for their validity. Our plot is a numerical assurance for their correctness.

The Mathematica code for the plot in **Figures 1** and **2** is in Appendix A. The two plots are identical even though their formulas are quite different and we do not know whether they can be analytically proved to be equivalent: our plots are numerical validation of both of their formulas.

3. Discussion

The lower bound for *CV* of reinsurance contract annual loss is established. The largest of those bound are also proved to be the upper bound for all *CV*. Applying this range information to ratio distribution, we can get theoretical value of the probability that different simulations will have simulated mean annual loss with deviation from each other less than a given percentage, under the Normal Distribution assumption of the mean annual loss. We think this assumption can be removed by using more suitable distributions, with numerical methods, but may still give the probability not too different. Typical example case numerical study confirmed this, and showed that the "Normal approximation" gives probability only a few percent (2–5%) less than using more suitable distributions that do not have explicit formula for the probability.

As the starting point and the application of the *CV* range, the ratio distribution and the model risk quantification results we get may be only rudimentarily correct due to other factors, such as the distribution modeling, the dependence modeling, and additional parameters dependence than just the *CV* and ρ, but our *CV* inequality itself is mathematically sound.

The less general upper bound $\sqrt{n-1}$ where all probabilities are equal is obtained by Katsnelson and Kotz in the literature [4, 5].

Using the same Hölder's inequality and calculus technique which may not have a simple elementary inequality approach counterpart, we can prove a more complex formula:

Theorem 2. *For reinsurance contract simulated annual losses f, the standard deviation σ_f with respect to the mean m_f is bound below by:*

$$\sigma_f \geq \sqrt{\left(m_f - m\right)^2 \mu(m) + \left(M - m_f\right)^2 \mu(M) + \frac{\left[\left(m_f - m\right)\mu(m) - \left(M - m_f\right)\mu(M)\right]^2}{1 - \mu(m) - \mu(M)}} \tag{40}$$

when $\mu(m) + \mu(M) < 1$, where σ_f is the standard deviation, m_f is the mean, m is the minimum annual loss, M is the maximum annual loss, $\mu(m)$ denote the measure of the numbers of minimum losses years, $\mu(M)$ denote the measure of the numbers of maximum losses years.

Proof:

In the inequality Eq. (2), we divide Ω into three subset and let the nonnegative function *g* be constant in each of the three sets:

$$g|_{\{f=m\}} = a, g\big|_{\{m<f<M\}} = b, g\big|_{\{f=M\}} = c \tag{41}$$

Then:

$$\int g^2 d\mu = a^2\mu(m) + b^2\big(1-\mu(m)-\mu(M)\big) + c^2\mu(M) \tag{42}$$

$$\int |f-m_f| g d\mu = \int_{\{f=m\}} |f-m_f| a d\mu + \int_{\{m<f<M\}} |f-m_f| b d\mu + \int_{\{f=M\}} |f-m_f| c d\mu \tag{43}$$

$$\geq \big(m_f - m\big)a\mu(m) + \int_{\{m<f<M\}} f b d\mu - \int_{\{m<f<M\}} m_f b d\mu + \big(M-m_f\big)c\mu(M) \tag{44}$$

$$= \big(m_f - m\big)a\mu(m) + m_f b - mb\mu(m) - Mb\mu(M) - m_f b\big(1-\mu(m)-\mu(M)\big) + \big(M-m_f\big)c\mu(M) \tag{45}$$

$$= \big(m_f - m\big)(a+b)\mu(m) + \big(M-m_f\big)(c-b)\mu(M) \tag{46}$$

due to $|f-m_f| \geq f-m_f$.

We get:

$$\sigma_f \geq \frac{\big(m_f - m\big)(a+b)\mu(m) + \big(M-m_f\big)(c-b)\mu(M)}{\sqrt{\big(a^2-b^2\big)\mu(m) + \big(c^2-b^2\big)\mu(M) + b^2}} \tag{47}$$

$$= \frac{\big(m_f - m\big)\left(\frac{a}{b}\right)\mu(m) + \big(M-m_f\big)\left(\frac{c}{b}\right)\mu(M) + \big(m_f - m\big)\mu(m) - \big(M-m_f\big)\mu(M)}{\sqrt{\left(\frac{a}{b}\right)^2\mu(m) + \left(\frac{c}{b}\right)^2\mu(M) + 1 - \mu(m) - \mu(M)}} \tag{48}$$

suppose $b > 0$.

Using the negative form of the inequality $|f-m_f| \geq m_f - f$, we also get a dual form inequality:

$$\sigma_f \geq \frac{\big(m_f - m\big)(a-b)\mu(m) + \big(M-m_f\big)(c+b)\mu(M)}{\sqrt{\big(a^2-b^2\big)\mu(m) + \big(c^2-b^2\big)\mu(M) + b^2}} \tag{49}$$

$$= \frac{\big(m_f - m\big)\left(\frac{a}{b}\right)\mu(m) + \big(M-m_f\big)\left(\frac{c}{b}\right)\mu(M) - \big(m_f - m\big)\mu(m) + \big(M-m_f\big)\mu(M)}{\sqrt{\left(\frac{a}{b}\right)^2\mu(m) + \left(\frac{c}{b}\right)^2\mu(M) + 1 - \mu(m) - \mu(M)}} \tag{50}$$

suppose $b > 0$.

Define:

$$t = \frac{a}{b}, A = \big(m_f - m\big)\mu(m), B = \big(M-m_f\big)\left(\frac{c}{b}\right)\mu(M) + \big(m_f - m\big)\mu(m) - \big(M-m_f\big)\mu(M) \tag{51}$$

$$C = \mu(m), D = \left(\frac{c}{b}\right)^2\mu(M) + 1 - \mu(m) - \mu(M) \tag{52}$$

$$F(t) = \frac{At + B}{\sqrt{Ct^2 + D}} \tag{53}$$

Then

$$A \geq 0, C \geq 0, D > 0 \tag{54}$$

The derivative

$$F'(t) = \frac{AD - BCt}{\left(D + Ct^2\right)^{\frac{3}{2}}} \tag{55}$$

If $B \leq 0$, then $F'(t) > 0$, $F(t)$ take the maximum $\frac{A}{\sqrt{C}}$ at ∞. If $B > 0$, then $F(t)$ increase on $(0, \frac{AD}{BC})$ and decrease on $(\frac{AD}{BC}, \infty)$, attain the maximum $\sqrt{\frac{A^2}{C} + \frac{B^2}{D}}$ at $\frac{AD}{BC}$.

Apply the same argument to

$$\frac{B}{\sqrt{D}} = \frac{\left(M - m_f\right)\left(\frac{c}{b}\right)\mu(M) + \left(m_f - m\right)\mu(m) - \left(M - m_f\right)\mu(M)}{\sqrt{\left(\frac{c}{b}\right)^2\mu(M) + 1 - \mu(m) - \mu(M)}} \tag{56}$$

with $(M - m_f)\mu(M) > 0$ since $\mu(m) + \mu(M) < 1$, we have if $(m_f - m)\mu(m) - (M - m_f)\mu(M) > 0$, then $\frac{B}{\sqrt{D}}$ attain the maximum $\sqrt{\left(M - m_f\right)^2\mu(M) + \frac{\left[\left(m_f - m\right)\mu(m) - \left(M - m_f\right)\mu(M)\right]^2}{1 - \mu(m) - \mu(M)}}$ at $\frac{\left(M - m_f\right)\left(1 - \mu(m) - \mu(M)\right)}{\left(m_f - m\right)\mu(m) - \left(M - m_f\right)\mu(M)}$.

If $(m_f - m)\mu(m) - (M - m_f)\mu(M) = 0$, then $\frac{B}{\sqrt{D}}$ is monotonically increasing with respect to $\frac{c}{b}$ and attain the maximum $\left(M - m_f\right)\sqrt{\mu(M)}$ at ∞.

If $(m_f - m)\mu(m) - (M - m_f)\mu(M) < 0$, then use the inequality Eq. (50), we can follow the same steps to arrive at the same form of maximum formula. We thus proved the maximal

$$\sqrt{\frac{A^2}{C} + \frac{B^2}{D}} = \sqrt{\left(m_f - m\right)^2\mu(m) + \left(M - m_f\right)^2\mu(M) + \frac{\left[\left(m_f - m\right)\mu(m) - \left(M - m_f\right)\mu(M)\right]^2}{1 - \mu(m) - \mu(M)}} \tag{57}$$

with the specific choice of $\frac{c}{b}$ and $\frac{a}{b}$. □

We can also prove by calculus that:

Theorem 3. *In the terminology of Theorem* 2, *if* $m = 0$,

$$\sqrt{\left(m_f - m\right)^2\mu(m) + \left(M - m_f\right)^2\mu(M) + \frac{\left[\left(m_f - m\right)\mu(m) - \left(M - m_f\right)\mu(M)\right]^2}{1 - \mu(m) - \mu(M)}} \geq m_f\sqrt{\frac{\mu(m)}{1 - \mu(m)}} \tag{58}$$

Proof:

Define

$$F(t) = \left(m_f - m\right)^2 \mu(m) + \left(M - m_f\right)^2 t + \frac{\left[\left(m_f - m\right)\mu(m) - \left(M - m_f\right)t\right]^2}{1 - \mu(m) - t} - m_f^2 \frac{\mu(m)}{1 - \mu(m)} \tag{59}$$

for $t \geq 0$.

Then $F(t)$ is continuous at 0 and

$$F(0) = \left(m_f - m\right)^2 \mu(m) + \frac{\left[\left(m_f - m\right)\mu(m)\right]^2}{1 - \mu(m)} - m_f^2 \frac{\mu(m)}{1 - \mu(m)} \tag{60}$$

$$= \frac{\left(m_f - m\right)^2 - m_f^2}{1 - \mu(m)} \mu(m) = 0 \tag{61}$$

The derivative of $F(t)$ is

$$F'(t) = \frac{\left(m_f - M + (M - m)\mu(m)\right)^2}{\left(-1 + t + \mu(m)\right)^2} \tag{62}$$

which is always nonnegative, so $F(t) \geq 0$ for any $t \geq 0$. □

Theorem 3 can be combined with the following form of the Hölder's inequality:

$$\left(\int\limits_{\{m<f<M\}} |f - m_f| 1 \mathrm{d}\mu \right)^2 \leq \int\limits_{\{m<f<M\}} |f - m_f|^2 d\mu \int\limits_{\{m<f<M\}} 1^2 d\mu \tag{63}$$

to give an alternative proof of Theorem 2 and then Eq. (16) (or directly for Eq. (20) by using the set $\{f > m\}$).

So there is a complex but better lower bounds Eq. (40), and empirical study shows that when $\mu(m) > 0.86$, both bounds are close to the true σ_f to within 86% with the simple form Eq. (8) 3–4% lower than the complex form Eq. (40). Even though the complex form Eq. (40) is generally valid for any discrete random variable, it may not be as easily applicable as the simple form Eq. (8) when we need a fast first approximation, and hence of less practical interest.

With numerical simulation, we can get σ_f and CV directly, so these formulas seems not to be useful for the numerical results. But since each simulation may arrive at a different value, known a priori their approximate value will be a check for any possible simulation process problem. Our inequalities also reveal that the CV is intrinsically related to important value distribution characteristics of the annual loss random variable. This essentialness of CV is also confirmed by other studies, such as the correlation and cluster analysis of these random variables.

4. Conclusions

Lower bound for reinsurance contract annual loss standard deviation involving zero losses years counts are obtained, which imply a general upper bound for annual loss *TVaR* or *AEP* with no mention of zero losses years. Alternative forms of these bounds give inequalities for probability of attaching and exceeding. These bounds can explain the difficulties or instabilities observed in numerical simulations, show the major reason of the limitation of the simulation is high *CV* and give clue to alternative solutions.

Acknowledgements

This research is supported by Validus Research Inc.

Conflict of interest

The authors declare no conflict of interest.

Thanks

The author thanks Nancy Wang for checking against a C++ application that validated the practical usefulness of our inequality.

Appendix-A

```
hinkley[x_?NumericQ,c_,p_]:=1/Sqrt[2 Pi]/c (x + 1) (1 − p)/(x^2–2 p x + 1)^1.5 Exp[−1/2/c^2 (x − 1)^2/(x^2–2 p x + 1)](CDF[NormalDistribution[0,1],1/c Sqrt[(1 − p)/(1 + p)] (x + 1)/Sqrt [x^2–2 p x + 1]] − CDF[NormalDistribution[0,1],−1/c Sqrt[(1 − p)/(1 + p)] (x + 1)/Sqrt[x^2–2 p x + 1]]) + 1/Pi Sqrt[1 − p^2]/(x^2–2 p x + 1) Exp[−1/c^2/(1 + p)];

H[x_,c_,p_]:=Integrate[hinkley[y,c,p],{y,−Infinity,x}];

LogLinearPlot[Evaluate[H[1.5,x,0.822434] − H[0.5,x,0.822434]],{x,0.1,10},PlotRange− > All, GridLines− > Full,GridLinesStyle− > Directive[Gray,Dashed],Mesh− > Automatic, ImageSize− > Full,Frame− > on];

Clear[q,marsaglia,M,MA]

Off[NIntegrate::inumr]

q[t_,p_,c_]:=q[t,p,c] = With[{},1.0/c (1.0 + (1.0 − p) t/Sqrt[1.0 − p^2])/Sqrt[1.0 + t^2]]
```

```
marsaglia[t_,p_,c_]:=marsaglia[t,p,c] = With[{},Exp[−1.0/(1.0 + p)/cˆ2]/Pi/(1.0 + tˆ2)(1.0 + q[t,p,c]
Exp[q[t,p,c]ˆ2/2.0] Evaluate[Integrate[Exp[−yˆ2/2.0],{y,0.0,q[t,p,c]}]])]

M[v_,p_,c_]:=M[v,p,c] = With[{},CDF[NormalDistribution[0,1],1/c (1 − p)/Sqrt[1 − pˆ2]] + CDF[

NormalDistribution[0,1],1/c] − 2 CDF[NormalDistribution[0,1],1/c (1 − p)/Sqrt[1 − pˆ2]] CDF[.

NormalDistribution[0,1],1/c] + Evaluate[NIntegrate[marsaglia[u,p,c],{u,0.0,v}]]]

MA[v_,p_,c_]:=MA[v,p,c] = With[{},M[(v − p)/Sqrt[1.0 − pˆ2],p,c]];

DistributeDefinitions [q,marsaglia,M,MA];

LogLinearPlot[Evaluate[MA[1.5,0.822434,x] − MA[0.5,0.822434,x]],{x,0.1,10},PlotRange− > All,
GridLines− > Full,GridLinesStyle− > Directive[Gray,Dashed],Mesh− > Automatic,
ImageSize− > Full,Frame− > on]
```

Author details

Frank Xuyan Wang

Address all correspondence to: frank.wang@validusresearch.com

Validus Research Inc., Waterloo, ON, Canada

References

[1] Lloyd's Market Association. Catastrophe Modelling Guidance for Non-Catastrophe Modellers [Internet]. 2013. Available from: http://www.lmalloyds.com/lma/finance/Cat_Modelling_Guidance.aspx [Accessed: March 5, 2018]

[2] Hinkley DV. On the ratio of two correlated normal random variables. Biometrika 1969; **56**(3):635-639. DOI: 10.2307/2334671. Available from: http://www.jstor.org/stable/2334671 [Accessed: May 12, 2016]

[3] Marsaglia G. Ratios of normal variables. Journal of Statistical Software. 2006;**16**:1-10. DOI: 10.18637/jss.v016.i04. Available from: https://www.jstatsoft.org/article/view/v016i04/v16i04.pdf [Accessed: May 05, 2016]

[4] Cox NJ. Speaking Stata: The limits of sample skewness and kurtosis. Stata Journal. 2010;**10**:492-495. Available from: http://www.stata-journal.com/sjpdf.html?articlenum=st0204 [Accessed: July 25, 2016]

[5] Katsnelson J, Kotz S. On the upper limits of some measures of variability. Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B. 1957;**8**(1):103-107. DOI: 10.1007/BF02260299. Available from: http://link.springer.com/article/10.1007%2FBF02260299 [Accessed: July 25, 2016]

IntechOpen

IntechOpen