ST3246 Statistical Model for Actuarial Science
Term Paper

Value-at-Risk (VaR) and Its Application in Actuarial Science

Author: Tianyu Wu

National University of Singapore
April 12, 2022

Contents

1 Introduction

Risk management is important in many areas of business, as are insurers, which hold a portfolio of insurance policies that could result in varying degrees of claims [1]. To better help insurers with risk control, many different risk measures have been used to calculate and evaluate the overall risk exposure [2], and they are divides into two categories: premium-based and capital-based in Tse's book [1].

But for business owners and executives, it is more intuitive for them to understand the current risks faced by the company through one or a few simple numbers. This claim has contributed to the development of the increasing popularity of the concept of Value-at-Risk (VaR) since the 90s [3]. As for now, no matter the financial industries but also financial regulators are making wide use of VaR when measuring the risk of insurers, banks, and other financial institutions [4]. As a capital-based risk measure, VaR plays an important role for both financial investors, practitioners, and regulators [1].

In this paper, we will systematically introduce the definition of Value-at-Risk (VaR), how we can calculate its value with different methods from the insurers' perspectives, and evaluate its pros and cons.

2 Notation

Table 1 shows the notations we will mainly use in this paper.

Name	Symbol
Value-at-Risk	VaR
Confidence Level	α
Holding Period	Δt
Loss Variable	$X_{\Delta t}$

Table 1: Notations

3 Value-at-Risk (VaR)

3.1 Definition

Value-at-Risk (VaR) of a loss variable shows the minimum value of the distribution such that the probability of a loss greater than that value would not exceed a given confidence level [1].

For example, suppose a portfolio of insurance policies has a one-year 5% VaR of \$1 million, it means that this portfolio has a 95% chance of paying less than or equal to \$1 million for the claims over one year.

More intuitively speaking, according to Figure 1, the red line represents the pre-defined confidence level (0.05 in this example), which is equivalent to saying that the blue area accounts for 5% of the whole distribution. By definition, VaR, in this case, is the intersection value between the red line and x-axis, which is approximately 11,645.

Adapted from the VaR expression commonly used in the investment returns [5], we claim that VaR of measuring the loss variable within a holding period Δt is expressed to

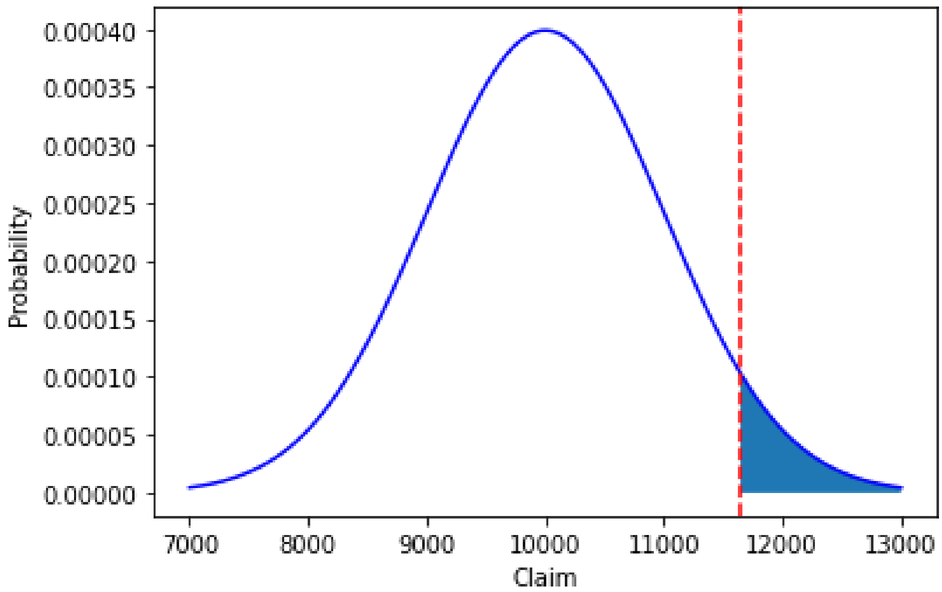

Figure 1: VaR visualization

satisfy the following equation:

$$Prob(X_{\Delta t} \geq VaR) = 1 - \alpha$$

where the confidence level α is usually selected to be close to 1, for example, 0.9, 0.95, or 0.99.

3.2 Calculation

The estimation of the probability distribution of the loss variable in a certain holding period $X_{\Delta t}$ is the most essential problem in calculating the VaR value. In the following paragraphs, I will introduce two typical approaches, Historical Simulation Method and Variance-Covariance Method to calculate VaR, and provide one example for readers to digest easily.

3.2.1 Historical Simulation Method (HSM)

Simply speaking, Historical Simulation Method (HSM) assumes that future changes in risk factor variables are exactly equivalent to the past. Following this belief, HSM takes prior data over a defined period to extract a certain amount of risk factors, and further uses the outcomes to make future estimations.

In general, suppose that the loss at time t is X_t, then

$$X_t = X(f_1(t), f_2(t), \cdots, f_n(t))$$

where X_t is influenced by n risk factors.

If we perform first-order difference on the historical loss variables, then the changes in past T time in each risk factor variable are

$$\Delta f_i(-t) = f_i(-t+1) - f_i(-t), i = 1, 2, \cdots, n; t = 1, 2, \cdots, T$$

According to the assumption of HSM, we have the future risk factors satisfying

$$f_i(t) = f_i(0) + \Delta f_i(-t), i = 1, 2, \cdots, n; t = 1, 2, \cdots, T$$

Thus we can obtain the estimation of future loss distribution by

$$X_t = X(f_1(t), f_2(t), \cdots, f_n(t)), t = 1, 2, \cdots, T$$

With a pre-defined confidence level, this new distribution estimation allows us to calculate the value of VaR.

3.2.2 Variance-Covariance method

The logic behind Variance-Covariance method is similar to the HSM we introduced above. However, here we assume that the losses, comprising a certain amount of risk factors, follow a specific distribution. In real life, it is commonly assumed to be a normal distribution for convenience, and people will rely on the properties of the normal distribution to make estimations [3]. In other words, VaR can be estimated from a known probability density function $f(x)$, then with the given confidence level α, it is not difficult to derive the following equation with the integral format:

$$1 - \alpha = Prob(X_{\Delta t} \geq VaR) = \int_{VaR}^{\infty} f(x)\, dx$$

Here are in general two main steps when applying the Variance-Covariance method:

Step 1:

Analyze and estimate the parameter values of the risk factor loss distribution through historical data, such as variance, mean, correlation coefficient (if a portfolio), etc.;

Step 2:

For one insurance policy, determine the VaR value of each risk factor according to the given confidence level and the assumed probability density function, and for a portfolio of insurance policies, further rely on the correlation coefficient between each risk element to determine the VaR value of the entire portfolio.

3.2.3 Example

In this section, we create a simple application scenario of how VaR can be used and calculated from the insurers' perspective.

We generate 100 random variables $X_i \sim \mathcal{N}(10,000,\, 1,000^2)$, where each variable represents the claim value ($) within one day from the previous 100 days (Appendix C), and we apply the two methods above to calculate the VaR value under different confidence levels, 0.99, 0.95, 0.9. We also include the theoretical value derived from the normal distribution, and found out that compared to Variance-Covariance method, Historical Simulation Method might be more susceptible to some outliers, and together with some endogenous limitations of interpolation function, it somewhat gives estimates with a larger error. Table 2 shows the derived results.

Code implementation is specified in Appendix A.

α	99	95	90
Historical Simulation Method	12692.8	11711.5	11390.5
Variance-Covariance Method	12339.8	11631.5	11253.9
Thoretical Value from Normal Distribution	12326.3	11644.9	11281.6

Table 2: VaR Estimation: Comparison of Historical Simulation Method, Variance-Covariance Method, and Theoretical Value

3.3 Evaluation

Coherence is one of the most commonly used theories to evaluate the pros and cons of risk measures [1]. In this section, we will use the four axioms of measures (Appendix B) to evaluate whether the risk measure VaR satisfies the property of coherence.

3.3.1 Translational Invariance

For a loss variable X and any nonnegative constant a, it is not difficult to derive

$$VaR(X + a) = VaR(X) + a$$

Proof: According to the definition of VaR, consider the loss variable $Y = X + a$, then

$$1 - \alpha = Prob(Y_{\Delta t} \geq VaR(Y)) = Prob(X_{\Delta t} \geq VaR(Y) - a)$$

But for the definition of VaR of loss variable X, it satisfies

$$1 - \alpha = Prob(X_{\Delta t} \geq VaR(X))$$

Therefore,

$$Prob(X_{\Delta t} \geq VaR(Y) - a) = Prob(X_{\Delta t} \geq VaR(X))$$

and it implies that $VaR(Y) - a = VaR(X)$ according to the uniqueness of cumulative density function, and thus it is natural to derive the result that $VaR(X + a) = VaR(X) + a$. □

This means that the risk measure VaR satisfies the requirement of Translational Invariance.

3.3.2 Positive Homogeneity

For a loss variable X and any nonnegative constant a, it is not difficult to derive

$$VaR(aX) = aVaR(X)$$

Proof: According to the definition of VaR, consider the loss variable $Y = aX$. then

$$1 - \alpha = Prob(Y_{\Delta t} \geq VaR(Y)) = Prob(X_{\Delta t} \geq \frac{1}{a}VaR(Y))$$

But for the definition of VaR of loss variable X, it suffices to claim

$$1 - \alpha = Prob(X_{\Delta t} \geq VaR(X))$$

Therefore,

$$Prob(X_{\Delta t} \geq \frac{1}{a}VaR(Y)) = Prob(X_{\Delta t} \geq VaR(X))$$

and it implies that $\frac{1}{a}VaR(Y) = VaR(X)$ according to the uniqueness of cumulative density function, and thus it is naturally to derive that $VaR(aX) = aVaR(X)$. $\square$

This means that the risk measure VaR satisfies the requirement of Positive Homogeneity.

3.3.3 Monotonicity

For loss variables X and Y such that $X \leq Y$ among all the states, it is easy to see

$$VaR(X) \leq VaR(Y)$$

For the simplest case, if $X = Y$, then naturally $VaR(X) = VaR(Y)$. And if $X < Y$, then if we let $Y = X+a, a > 0$, it would naturally direct to the conclusion of Section 3.3.1, where

$$VaR(Y) = VaR(X + a) = VaR(X) + a > VaR(X)$$

as desired. $\square$

This means that the risk measure VaR satisfies the requirement of Monotonicity.

3.3.4 Subadditivity

VaR does not satisfy the requirement of subadditivity, which causes VaR not to be a coherent risk measure. In other words, we can not guarantee all loss variables X and Y, such that

$$VaR(X + Y) \leq VaR(X) + VaR(Y)$$

Counter Example:
Consider the following two portfolios of insurance policies:
A: There is a 98% probability of paying \$50 claim, and a 2% probability of paying \$70 claim.
B: There is a 96% probability of paying \$40 claim, and a 4% probability of paying \$90 claim.

Therefore, for the equal-weighted portfolio $A + B$, it is easy to calculate that there is a 94.08% probability of paying \$90 claim, a 1.92% probability of paying \$110 claim, a 3.92% probability of paying \$140 claim, and a 0.08% probability of paying \$160 claim.

According to the definition of VaR, we can tell that under 95% confidence level

$$VaR(A) = 50, VaR(B) = 40, VaR(A + B) = 110$$

which implies that

$$VaR(X + Y) > VaR(X) + VaR(Y) \quad \square$$

This result is against the principle of subadditivity, which implies that VaR does not satisfy this requirement of Subadditivity.

4 Conclusion and Future Work

The core of risk management is the quantitative analysis and assessment of risk. As the size and liquidity of financial markets have increased over time, the risk measurement techniques have become more esoteric. However, currently, it is still not possible to find a single risk measurement method that is suitable for all situations globally. The VaR method is currently one of the most popular risk measures worldwide since it answers a critical question that many executives or business owners are concerned about: "What is the worst situation" simply with one number. However, it also has a lot of drawbacks. For example, VaR is not a coherent risk measure and this approach makes it difficult to capture the expectation value of risks beyond VaR because of its thick-tailed nature. Some researchers then come up with another revised risk measure, Conditional Value-at-Risk ($CVaR$), also known as the expected shortfall, by taking the amount of tail risk into consideration beyond the VaR cut-point [1].

In the future, we expect to make a comparative empirical study between VaR and $CVaR$ methods in the application scenario of actuarial science, to better understand how they work in either one or a portfolio of insurance policies.

References

[1] Yiu Kuen Tse. *Nonlife actuarial models : theory, methods and evaluation*. Cambridge University Press, 2009.

[2] Terje Aven. Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253:1–13, 08 2016.

[3] Peter Adamko, Erika Spuchľáková, and Katarína Valášková. The history and ideas behind var. *Procedia Economics and Finance*, 24:18–24, 2015.

[4] Jose A. Lopez. Regulatory evaluation of value-at-risk models. *SSRN Electronic Journal*, 1997.

[5] Darrell Du and Jun Pan. An overview of value at risk, 1997.

[6] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 07 1999.

A Appendix A: Python Implementation

Listing 1: Python Code Implementation

```
# import packages
import numpy as np
import pandas as pd

# set the expected mean and variance to generate random loss variables
mu = 10000
sigma_1 = 1000
s1 = np.random.normal(mu, sigma_1, 100)

# get the dataframe
df_1 = pd.DataFrame(s1, columns = ['claim'])
temp1 = df_1['claim'].sort_values(ascending=False)

# HSM
## find the percentile from historical data to get VaR
p = np.percentile(temp1, (99,95,90), interpolation='midpoint')

# Variance-covariance method
## Calculate the sample mean and variance from the sample data
from scipy.stats import norm
u_1 = df_1.claim.mean()
var_1 = df_1.claim.var()
std_1 = df_1.claim.std()

# Variance-covariance method assumes that the data follows a normal
    distribution, thus we use the properties of normal distributions to
    calculate VaR
Z_01 = norm.ppf(0.99)
# (R*-u)/std = Z_01
print(Z_01 * std_1 + u_1)

Z_05 = norm.ppf(0.95)
# (R*-u)/std = Z_01
print(Z_05 * std_1 + u_1)

Z_10 = norm.ppf(0.9)
# (R*-u)/std = Z_01
print(Z_10 * std_1 + u_1)

# Theoretical value of the normal distribution (10,000, 1,000,000) under
    different confidence levels
norm.ppf(0.99, loc=10000, scale=1000)
norm.ppf(0.95, loc=10000, scale=1000)
norm.ppf(0.9, loc=10000, scale=1000)
```

B Appendix B: Axioms of coherent risk measures

A risk measure $\rho(\cdot)$ that satisfies the following four axioms is said to be **coherent** [1] [6].

B.1 Axiom 1: Translational Invariance

For any loss variable X and any nonnegative constant a,

$$\rho(X + a) = \rho(X) + a$$

This shows that if there is a fixed amount increase on the loss variable X, then the corresponding risk will increase by the same amount.

B.2 Axiom 2: Positive Homogeneity

For any loss variable X and any nonnegative constant a,

$$\rho(aX) = a\rho(X)$$

This ensures that the change of monetary units of risks would not have a difference to the risk measure.

B.3 Axiom 3: Monotonicity

For any loss variables X and Y such that $X \leq Y$ among all the states,

$$\rho(X) \leq \rho(Y)$$

This ensures that the one risk measure cannot be more than the other, if the loss of the former one is no more than the latter one, under all the states.

B.4 Axiom 4: Subadditivity

For any loss variables X and Y,

$$\rho(X + Y) \leq \rho(X) + \rho(Y)$$

This shows that combining different insurance policies would not make the company riskier and it also implies that the insurer is not able to reduce its risk by dividing its business into smaller blocks.

C Appendix C: Data

count	claim
0	10679.41806327092
1	10739.47215875537
2	7761.754686393027
3	9205.540236602646
4	10323.726348702214
5	9744.267191356881
6	9474.198168688761
7	10186.787146466913
8	9204.982283050465
9	10882.505567622202
10	10773.366989476392
11	11475.304198452117
12	11344.470876428099
13	11123.605360215517
14	7892.363738487131
15	11040.158482200584
16	9310.390754805629
17	8798.875392182985
18	9170.82421425786
19	9550.057560912257
20	11946.277894855159
21	8930.498232175056
22	9893.389755007567
23	9140.264261795199
24	9029.220927152845
25	9622.517408585885
26	9397.375180644838
27	10561.183401871003
28	9032.522700000383
29	9886.144163193667
30	11592.47400687748
31	11020.01674449707
32	9549.63979715555
33	10241.84301166382
34	10223.027732678638
35	9521.940276307787
36	11436.52923279852
37	10505.532045663707
38	9959.588820034149
39	10216.398800727937

count	claim
40	9238.191643058104
41	10615.540130178948
42	10267.233707957997
43	9049.714423297364
44	9797.48039371873
45	10431.93507853403
46	8710.741055822236
47	10279.638546692757
48	8966.593107756402
49	9737.413584848056
50	8822.17871821643
51	9512.768718376947
52	9788.935822542278
53	9665.621777999488
54	9411.595061684056
55	7134.007707867774
56	9463.802308080067
57	9835.04271147001
58	10727.775781749697
59	9618.653767040785
60	9375.921928418686
61	9233.549564352246
62	9887.572080386102
63	13371.1008464959
64	10322.478218680102
65	8710.655285594896
66	8557.083434087397
67	10725.601327715218
68	11852.371407315848
69	8211.84622932242
70	10010.76592019593
71	11588.055565125585
72	11830.540137930991
73	8957.889395275943
74	7773.326790406256
75	11107.326067246562
76	10403.145627606851
77	9464.818712466278
78	9845.896639991319
79	10870.27411702727

count	claim
80	9706.241793381245
81	8974.531546028938
82	10084.027288831725
83	9469.133350983495
84	10119.38177997874
85	8623.753190733232
86	10806.53458271859
87	8532.508427337609
88	9643.284312883654
89	9543.9918614486
90	12014.524829164908
91	8698.601977840734
92	9593.868538604343
93	10183.057728699567
94	10618.623575899472
95	10354.355531739842
96	11455.562855665916
97	9611.04500719501
98	9816.928539400418
99	10784.776910248209