Subadditivity and Parameter Uncertainty of VaR and Solvency Capital Requirement (SCR) in Tail Region of a Non-life Insurance Portfolio

Jan Dhaene¹
Ahmad Salahnejhad Ghalehjooghi²

Received: 04. March. 2015
Accepted: 21. Apr. 2015

Abstract
Although VaR is important due to its widespread usage to obtain overall Solvency Capital Requirement (SCR) in the standard model of Solvency II directives, it is not subadditive. Without subadditivity, the summation of SCRs of different lines of business, which is usually used by risk managers, may underestimate overall SCR for an insurance company. This research examines the subadditivity property of VaR for fat-tailed insurance losses in a dependent structure. The foundation of the paper is based on Danielson et al (2013), a study on subadditivity of VaR in the tail region of asset return data. We applied the same idea by using Generalized Pareto Distribution (GPD) to model the fat-tailed insurance losses and capture their dependence structure by the Gumbel-Hougaard copula through the tail of the joint distribution. Using these instruments, we proposed a simulation method to examine subadditivity of VaR and SCR. By empirical methods, we found that, similar to the fat-tailed asset returns, insurance losses are also more subadditive in tail region. We found that only going deep into the tail, will not guarantee monotonically more subadditivity, where “Variation of dependence” and “shape parameter” through the tail of the distribution are other important factors that Danielson et al didn’t take into account. More specially, when the correlation measure in different thresholds is changed, subadditivity of VaR deviates to increase monotonically in the tail. Furthermore, we observed that the uncertainty of VaR estimation is not always monotonically increasing through the tail; it may increase in the first thresholds of the right tail, it decreases in higher thresholds.

Keywords: Copula, Excess of loss Contract, Solvency Capital Requirement (SCR), Stop-loss Contract, Subadditivity, Tail Dependence, VaR.

1. Professor, KU Leuven. (Jan.Dhaene@kuleuven.be)
2. PhD Candidate of Actuarial Science, Maastricht University, Corresponding Author. (salahnejhad@maastrichtuniversity.nl)
1. Introduction

Why is the aggregate behavior of different lines of business an important factor in the survival of insurance companies? The efforts to answer this question result in vast research comparing the solvency (or shortfall) risk in the case of each individual line of business and the case of the aggregate as an overall risk. In a non-life insurance company with several lines of business, risk managers in each department can measure the solvency risk and prepare for capital requirements of their own line. The total solvency capital in practice, for simplicity, will usually be calculated as the summation of the capital for each line. However, this may not be sufficient due to unknown relationship between different lines of business and the possibility of producing an extra source of risk hidden under the joint movements and comonotonicity of different risks and their relationship in the aggregate portfolio of the company. In other words, risk managers don’t have enough information on how the individual portfolio of risks behaves together in a joint distribution (world). Hence, the main issue is to examine whether aggregated portfolio will produce more risk than summation of individual risk portfolios. Such an extra risk may be interpreted as the lack of the “Subadditivity” property of the risk measure that we use to calibrate the solvency (shortfall) risk. The Subadditivity property of two risks $X$ and $Y$ can be formulized as follows:

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

Without subadditivity, the summation of SCR of the all line of business will not cover the total SCR needed for the whole insolvency risk of the company.

Artzner et al (1999) introduced a class of risk measures called “Coherent” that have four essential properties of Subadditivity, Monotonicity, Homogeneity and Translation invariance. To support the coherency of a risk measure, subadditivity captures the diversification effect as merging two risks should not create any extra risk. Except for some special cases such as normality below the mean, $VaR$ fails subadditivity as well as coherency (Artzner et al, 1999) and alarms for the problem in calculation of the so-called total $SCR$ in
Solventcy II. Moreover, VaR does not represent any information about the behavior of distribution beyond the related percentile.

1.1. Relevance

Insurers always gather premiums before they pay for future claims. Then, to protect policyholders, the regulator will request the insurer to provide a Solvency Capital Requirement (SCR) in case premiums and investment returns won’t be sufficient to cover the future claims. Such “Insolvency” may be the result of extreme or catastrophe claims or market risks. To provide such a protection, European Union issued the Solvency II directives for European insurance companies. In Solvency II, “the SCR is the capital required to ensure that the (re)insurance company will be able to meet its obligations over the next 12 months with a probability of at least 99.5%”. Having SCR, the possibility of bankruptcy for an insurance company should decrease to less than once in a two hundred years period.

SCR in Solvency II is based on the Value at Risk (VaR) that measures the riskiness of a risk portfolio via 99.5% quantile of the risk distribution. As VaR is subadditive in a multivariate normal dependence structure, Solvency II uses a multivariate normal as a standard model for the dependence structure of different lines of business. Hence, assuming the covariance/correlation matrix risk managers can simply add VaR of each line of business and cross correlation terms to measure the overall SCR. However, non-life insurance risks normally don’t follow a normal distribution and their dependence may deviate considerably from the assumption of the standard model in Solvency II. As a result, due to lack of subadditivity, we must use internal models to fulfill risk management procedures to measure the overall SCR.

On the other hand, the uncertainty level in the VaR/SCR estimation is highly important and can affect the life of an insurance company. We consider a case of SCR for a reinsurance company which is active in non-life business, covering “Excess of loss” or “Stop-loss” contracts. The risk manager may receive some left-truncated claim data which has just information about the right tail of the original loss distribution. Considering extreme/catastrophic losses as the most important factor related to insolvency, Extreme Value Theory (EVT) can be justified to analyze the distribution of loss. However, due to
lack of frequent observations in higher layers of reinsurance coverage, we expect more uncertainty. This uncertainty may come from the model selection criteria or parameter uncertainty of the fitted loss distribution. Hence, even if we achieve subadditivity, an unreasonably high uncertainty will lead to under/over-estimation of SCR that can endanger the existence of the firm.

1.2. Previous Researches

Study of the insolvency risk is connected to three main subjects consisting of: Extreme Value Theory, Subadditivity and Tail Dependence Structure, and Uncertainty in Tail Region.

Before 1990s many researchers worked on special properties of extreme value theory and heavy tailed distribution. The Fundamental results to build this theory was obtained by Fisher & Tippett (1928), when they modeled the limiting behavior of normalized extremals and found a non-degenerate distribution for them called Generalized Extreme Value Distribution. Belkema & de Haan (1974) and Pickands (1975) offered one of the most useful theorems to get benefits of the extreme values theory in the world of the insurance loss distributions. The Theorem represents that under the condition of Fisher-Tippett theorem, if $X_1, X_2, \ldots$ are i.i.d random variables by $M_n = \max(X_1, X_2, \ldots X_n)$ for the first $n$ observations and

$$\lim_{n \to \infty} \Pr \left\{ \frac{M_n - b_n}{a_n} \leq x \right\} = \lim_{n \to \infty} F^n(a_nx + b_n) = G(x)$$

and assuming the coverage over a predetermined threshold of loss, when this threshold tends to the right tail of the distribution, the limiting distribution of excess random variables will be Generalized Pareto Distribution (GPD). This theorem plays an important role in modeling the excess of loss reinsurance layers and the behavior of extreme losses in the right tail.

The GPD, as an extreme value distribution, captures the heavy tailed nature of individual losses very well in non-life insurance business. Also Embrechts et al (1997) and McNeil (1997) examined the fitness of GPD as a parametric distribution for modeling losses above a threshold using some standard graphical and parametric methods, and proposed it as a better choice compared to lognormal. McNeil (1997) mentioned the results obtained by Davison (1984) and Davison &
Smith (1990) to develop the result of the Pickands-Belkema-deHaan theorem to ground-up the loss exceedances (the excess plus the threshold \( u \)). We will use different properties of GPD confronting the real situation, where we may need to change the threshold to study the behavior of special statistics such as \( VaR \) or even (re)insurer may accept various risks with different attachment points in their portfolio.

If we assume \( S = \sum_{i=1}^{n} X_i \) as the whole risk of a (re)insurance company, subadditivity (\( Var(\sum_{i=1}^{n} X_i) \leq \sum_{i=1}^{n} Var(X_i) \)) represents the summation of \( VaR \)s as an upper bound for the \( VaR \) of \( S \). Although in general \( VaR \) deviates from subadditivity property, in some especial cases or different data structures it may give positive response to subadditivity. At first many researchers believed that only in case of diversifiable risks such a bound can happen, and in case of non-diversifiable risks such as insurance losses, it will fail. Embrecht et al. (2003) showed that for a sequence of comonotonic random variables \( X_i \) s, \( VaR \) is additive, and \( \sum_{i=1}^{n} VaR(X_i) \) can be assumed as a comonotonic bound. Although in each line of business for example stop-loss and excess of loss coverage can be assumed as comonotonic risks in relation to original loss/risk/random variable, not all lines of business will necessarily behave together as comonotonic risks.

We would like to measure the whole solvency risk of a company which is related to its several risks, and as subadditivity of \( VaR \) also examines the behavior of several risks together, we should study them in a multivariate world to capture their dependency structure. To study the joint behavior of random variables and dependence structures in multivariate models, we use Joe (1996) and Nelsen (2006). Some risk specialists argued that dependence structure and the method used to measure correlation of risks can affect the diversification effect of \( VaR \) on aggregate risk portfolio. Burgi et al (2008) examined diversification benefits of aggregate risks regarding different methods of modeling the dependence structure by different forms of copulas. They found out that functional form of the copula has a serious effect on obtaining diversification gain, and can be a significant factor in achieving subadditivity. They also mentioned that for heavy tailed distribution (in our case insurance losses), Clayton or Gumbel copula can easily capture tail dependence by selecting \( \alpha \) parameter in copula. See also Jackie (2006) for more clarification specifically on modeling dependency for different insurance lines of business with copula.
Embrecht et al (2009) examined some general characteristics of underlying risk distribution that can affect subadditivity or superadditivity of quantile based risk measures. They argued that existence/non-existence of finite moment, one/two sidedness and symmetry/asymmetry of risks’ distributions are effective on subadditivity of VaR.

Ferri et al (2011) examined dependence and comonotonicity of different lines of business by using real data of three non-life insurance covers, and argued that even comonotonicity is not the only case to achieve the subadditivity of VaR and does not represent the worst possible scenario. They worked with the whole domain of loss distribution rather than just the right tail, and argued that even when risks are independent, but not comonotonic, VaR deviates from subadditivity. He examined the usefulness of Tail VaR instead of VaR to capture subadditivity, but mentioned the possible underestimation of SCR due to the error in model estimation imposed by TVaR. He also compared two different viewpoints of risk measurement for VaR and TVaR. With value at risk we just show the number of losses that may happen beyond the quantile, but with expected shortfall or TVaR we capture the severity and bigness of losses beyond the quantile.

On the other hand, Danielsson et al (2005) re-examined subadditivity of VaR in case of extreme value theory for heavy tailed distributions. Danielsson et al (2013) in a revised paper examined the subadditivity of asset returns and showed that VaR is subadditive in tail region, if multivariate distributions of returns have regular variation (which is a mathematical property of fat-tailed distributions) regardless of their dependency structure. They also discussed that to achieve subadditivity, the tail index for both returns must be equal and bigger than 1, (which is not either easy to achieve for different insurance loss distributions). They also argued that different tail indices can result just in weaker form of subadditivity. They argued that this result can be deviated if coarseness of empirical distribution affects the fatness of the tail and proposed a semi-parametric extreme value technique to solve this problem. They examined subadditivity for one tail of multivariate return distribution and argued that they can achieve the same for the other tail, which cannot be fully the same case for insurance loss distribution!
According to Danielsson et al (2013), obtaining the subadditivity in tail region is conditioned on going deep enough into the tail to apply the Feller’s convolution theorem. But in practice, going deep into the tail decreases the number of observations in estimation and imposes an expensive cost of uncertainty to buy subadditivity.

In 1997 Embrecht et al, and in 1997 McNeil et al. mentioned the magnified effect of parameter uncertainty in EVT. In 2000 Ana Mata, studied the effects of parameter uncertainty on calculation of insurance premiums, where price of insurance risks in excess of loss layers can be obtained better by EVT and using GPD of loss data. She incorporated three types of uncertainty that may affect the estimation as Model uncertainty, Parameter uncertainty for loss distribution and Parameter uncertainty for frequency distribution. She concluded:

“For high layers where there are very few data points, the premiums are very variable and they are very sensitive to the choice of threshold. We noticed that for higher thresholds the premiums are subject to higher standard error …”

She argued that due to the deficit in number of data points in parameter estimation for fitted GPD, any statistics estimated by these parameters are subject to high standard error. This may apply to statistics like VaR and SCR that we want to estimate by this method. McNeil (1997) compared the effect of threshold choice on estimation of quantile, and obtained different quantile estimations by changing the threshold in the same model and fitted distribution. He concluded that inference from statistics in tail region is quite sensitive to threshold choice in such a way that high threshold will result in parameter uncertainty, and low threshold will result in losing theoretical justification of the model. In 2007, Borowicz and Norman investigated the effect of parameter uncertainty in extreme event frequency-severity model. He used Bayesian approach to incorporate parameter uncertainty to fit a frequency-severity for modeling large events, and assessed capital requirements for VaR and TVaR. He illustrated existence of extra capital requirement and (re)insurance premium subject to parameter uncertainty. Therefore this can be a nice template to estimate the related VaR and capital requirements with incorporation of the parameter uncertainty in estimation of VaR for individual and aggregate losses to check their subadditivity property.
2. Model and Methodology

Using extreme value theory \((EVT)\) for fat (heavy) tailed distributions, we may observe quite different behavior in statistical measures such as correlation or convolution of random variables in tail region compared to the whole distribution.

We will check the existence of subadditivity of \(VaR\) in the tail region of a heavy tailed loss distribution using extreme value theory. We will also measure the related uncertainty level in estimation of related \(VaR\). To measure subadditivity, we calculate the ratio of “subadditivity violation” for \(VaR/SCR\) through different thresholds deep into the tail of multivariate insurance losses. We examine different factors that can be effective on subadditivity of \(VaR\) such as level of probability \((p)\), magnitude of correlation, structure and variation of dependence in tail region and type of loss random variables: Stop-loss \((SL)\) and Excess of loss \((EL)\). Moreover, according to scarcity of observations, we are interested in the price we should pay for subadditivity of \(VaR\) in heavy tail region in terms of uncertainty that we may confront in estimation.

The main part of the study is conducted by different empirical methods such as simulation, data generation and bootstrapping. The scope of study is non-life insurance. We will use severity of claims/losses and will ignore loss frequency and related uncertainty that it may cause. To make sense to the usefulness of subadditivity for heavy tailed distributions, the study focuses on the higher layers of losses in tail region using Stop-loss and Excess of loss random variables. We used two different series of loss data as two risks covered in third party vehicle liability insurance policy to give a real sense to dependent risks that a risk manager may confront in analysis.

2.1. Data Description

We will analyze two different insurance risks that are jointly covered in a unique insurance policy in one Iranian Insurance company. Covered risks are as follows:

\(X: \) Insured’s liability to compensate loss of or damage to third party properties such as car, building or any other infrastructure due to an accident caused by the insured’s car.
Y: Insured’s liability to compensate bodily injury to third party due to an accident caused by the insured’s car.

The base of bodily injury loss is Islamic Wergild or Blood Money called “Dieh” for at least one person, which is determined by the Iranian Judiciary every year.

Each loss file may contain summation of reported losses in risk X, risk Y or both of them related to a one year policy. Data set is huge and consists of 349,700 records for the policies that have had at least one loss in one of the covered risks. So it’s usual to have records with no observation in a coverage (that does not mean we can record loss amount as zero!) and also there would be some files with losses in both risks. Losses have happened in more than five years, from March 2007 to June 2012.

The value of losses in both risks is in Rials, the Iranian currency, and is adjusted to inflation. The observations of material damage (risk X) are adjusted according to annual services’ inflation (change in price index of services) and updated to their future value in the previous year. The observations of bodily injury are updated based on annual changes of base wergild value during the period. Thus, for example the annual percentage changes in the wergild value has been calculated for each year and then injury losses of each in a specific year are accumulated by the accumulation factor to the previous year. There is also the recorded date of loss for each line of business, in such a way that we are able to measure daily losses from policies that have had loss on a specific date.

We will use ground-up loss data for both risks beyond presumed thresholds rather than original data. In case of third party material damage (Risk X), we use observations above $d_X = u_{x1} = 5,000,000$, and in case of third party bodily injury we use observations above $d_Y = u_{y1} = 40,000,000$. The main form of data we use to do most of inferences is Stop-loss $X_{+u}$ and excess of loss $(X - u)_+$ data to study tail of loss distribution with different thresholds. For more simplicity we will divide all loses by 1,000,000 and will express them in scale of 1 million Rials.

1. Farvardin 1386 until Khordad 1391 in Persian calendar
2.1.1. Fitting method

Lots of researchers examined the usefulness of Generalized Pareto Distribution (GPD) in explaining original loss, and approved its preference to other parametric choices such as lognormal. Especially GPD is very flexible to fit to both ground-up and excess of loss data when we need plenty of changes in thresholds in tail region (see Embrecht et al 2005).

Based on Belkema-de Haan theorem, GPD can be fit to both ground-up loss and excess of loss distributions with the same shape and scale parameter, but with different thresholds. However, it is not sufficient to reflect our research methodology as our method requires showing process of probable subadditivity violations through distribution of losses from middle points to extreme points. Hence, some of the thresholds will not be high enough to apply Belkema-de Haan theorem to approximate stop-loss distribution by Excess of loss GPD distribution. Also, in case of higher thresholds where the frequency is relatively low, we might get biases in shape and scale parameters because of change in shape of distribution. Therefore, we will fit GPD to both stop-loss and excess of loss distributions of material damage and bodily injury risks in each threshold separately to update the change of parameters and improve the advantage of having the best possible accuracy in estimation.

2.2. Modeling by Generalized Pareto Distribution

The random variable $X$ follows Generalized Pareto Distribution (GPD) with parameters $u, \xi$ and $\sigma$, if its cumulative distribution function is

$$G_{u,\xi,\sigma}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-u)}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{(x-u)}{\sigma}\right) & \text{if } \xi = 0 \end{cases}$$

where $\xi$ is the shape parameter and $\sigma > 0$ is the scale parameter. Generally, as the shape parameter decreases the GPD gets fatter in the right tail. In case of $\xi = 0$, GPD turns to a simple exponential distribution with parameter $\sigma$. $\xi > 0$ represents a usual Pareto distribution with a shape parameter $\alpha = \frac{1}{\xi}$ and $\xi < 0$ parameterizes a
type II Pareto distribution with a “Super Fat-tailed” distribution with huge and long tail in right extremes.

Regarding extreme value distribution, Mata (2000) mentioned some useful properties of the Generalized Pareto distribution given by Embrechts et al (1997), which are useful in insurance related applications as below:

1) If \( X \sim GPD(\xi, \sigma) \), then excess of loss over the threshold \( u \), \( X - u \) has also \( GPD(\xi, \sigma, u) \).

2) Mean excess function over a special threshold \( u \),

\[
E[X - u/X > u] = \frac{\sigma + \xi u}{1 - \xi} \quad \text{for} \quad \xi < 1
\]

is a linear function of related threshold.

3) If we change the threshold, the excess of loss random variable still has GPD. i.e. “... if \( X \) has a \( GPD(\xi, \sigma) \), the probability that \( X \) exceeds \( u + v \) given that it exceeded \( u \) is also a probability in the generalized Pareto family.”

4) If the distribution of the excess of loss random variable is \( GPD \), it is possible to estimate the distribution of the original loss variable \( X \) in the area \( X > u \), and that \( X \) will have a \( GPD \) distribution with the same shape parameter \( \xi \) of the excess random variables and different scale location parameter.

Suppose \( F_u(x) = \Pr(X - u \leq x/X > u) = \frac{F(x+u) - F(u)}{1 - F(u)} \) as the distribution function of conditional excess of loss, and as \( GPD \) distributions with the same shape and scale parameters, but with and without threshold. Case (4) above states that by (Belkema & de Haan 1974 and Pickands 1975) theorem, for a sufficiently higher threshold \( u \), \( G_{\xi,\sigma}(x) \) can be appropriately fit to excess of loss distribution. In addition, without changing the parameters, just by assuming the threshold \( u \), \( G_{\xi,u,\sigma}(x) \) may approximate ground up loss distribution.

Stability of \( GPD \) as well as its flexibility against the change of threshold and transformation of ground-up and excess of loss is quite consistent with our aim to study subadditivity of \( VaR \) in tail region of
loss distribution. We will use GPD as the main parametric distribution to fit different data.

2.3. Measuring Dependence Structure

To capture the dependence structure among marginal risks, we assume two types of dependence; 1) Dependence on whole range of risk distribution, 2) Tail dependence. Typically there are three correlation coefficients to measure association of two risks.

**Pearson:** measures only dependency of risks with elliptical distributions such as normal marginal in Gaussian copula, and can only capture the linear dependence defined as below:

\[ \rho_{\text{linear}}(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \]

For non-elliptically distributed risks or other types of dependence such as non-linear on tail dependence, Pearson correlation is not a useful measure (Priest, 2003).

**Spearman:** as a nonparametric measure, relaxes the distribution assumption of risk to measure dependency, and measures correlation between the ranks of observations:

\[ \rho_{\text{Spearman}}^{XY} = 12 \cdot E[(U_X - 0.5)(U_Y - 0.5)] \]

Nelsen (1999) discussed that \( \rho_{\text{Spearman}}^{XY} \) is stable if we change the scale of marginals as it is invariant under any strictly increasing linear/non-linear transformation of \( X \) and \( Y \).

**Kendall’s Tau:** Also measures dependence in a nonparametric way, by comparing probability of concordance and discordance for any pair of loss observations:

\[ \tau_{XY} = \Pr[(X_i - X_j)(Y_i - Y_j) > 0] - \Pr[(X_i - X_j)(Y_i - Y_j) < 0]. \]

Similar to \( \rho_{\text{Spearman}}^{XY} \), \( \tau_{XY} \) is also invariant under strictly increasing linear or non-linear transformation, and is not dependent on marginal distributions, either.

Typically non-life insurance losses are not elliptical, and have heavy right tail while due to truncation and censoring required in different (re)insurance contracts, ground up losses should be transformed. Embrechts et al. (2001) referred to this point and advised
Nelsen (1999), Kendall and Gibbons studied some theoretical and practical relationships between $\rho_{XY}^S$ and $\tau_{XY}$. As a special result, if $|\rho_{XY}^S|$ and $|\tau_{XY}|$ are not too close to 1, then $\rho_{XY}^S$ will be approximately 50% more than $\tau_{XY}$. This means that $\rho_{XY}^S$ may overestimate the dependence. Moreover, if we use copulas to simulate dependence structure, there is a closed form relationship between Kendall’s tau and the Archimedean copula parameter. Therefore, we prefer to use $\tau_{XY}$ as the main measure of association to study dependence of underlying losses.

2.4. Copula Modeling

To simulate the marginal distribution of each risk based on its association or dependence structure, the proper way is to use a copula to capture the related association among marginal losses. Based on Sklar’s theorem, the copula is defined as a multivariate distribution function of dependent marginal uniform distributions which is unique where for any $u = (u_1, ..., u_n) \in [0,1]^n$ and generalized inverse function of $F_i$:

$$ C(u_1, ..., u_n) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)). $$

Function $C$ must be in (0, 1) and increasing. Furthermore, $C_i$ margins must satisfy $C_i(u) = C(1, ..., 1, u, 1, ..., 1)$ for all $u \in [0,1]$.

Copulas can capture dependency of different random variables in the form of multivariate uniform distribution regardless of their marginal distribution and any monotone transformation on them (see Schweizer & Wolf, 1981).

Amongst all, as Archimedean copulas allow for explicit formulas with a unique parameter for high dimension of dependent distributions they include, they are widely used in application. An Archimedean copula can be defined with a decreasing convex generator function $\psi(.)$ as represented below:

$$ C(u_1, ..., u_n) = \psi^{-1}[\psi(u_1), ..., \psi(u_n)] $$

Some famous forms of bivariate Archimedean copulas are as follows:
Subadditivity and Parameter Uncertainty ...

Clayton copula:
\[
\psi(t) = t^{-1} - 1 ; \quad C(u_1, u_2) \\
= (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha} ; \quad \alpha > 0
\]

Gumbel-Hougaard copula:
\[
\psi(t) = (-\ln t)^\alpha ; \quad C(u_1, u_2) \\
= \exp\{-(\ln u_1)^\alpha + (\ln u_2)^\alpha\}^{1/\alpha} ; \quad \alpha \geq 1
\]

Frank copula:
\[
\psi(t) = -\ln \frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1} ; \quad C(u_1, u_2) \\
= -\frac{1}{\alpha} \ln\left\{\frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}{e^{-\alpha} - 1}\right\} ; \quad \alpha \geq 1
\]

In all of the above Archimedean copulas, there is a closed form relationship between one of the correlation coefficients and association parameter \(\alpha\), and we may easily fit these copulas by estimating correlation coefficients. For example the association parameter \(\alpha = \frac{1}{1-\tau}\).

Among above copulas, Gumbel-Hougaard has useful properties which are consistent with extreme value of our data structure. It can reflect the upper tail dependence and positive association of marginal loss distributions. It is also typically useful for fat-tailed marginal distributions that are not symmetric.

Several methods have been examined in literature to recognize the preferred copula to fit insurance loss data such as Goodness of fit test, CDF test, \(K_c(t)\) test and binominal test (Jackie, 2006). As a relatively reliable method, Frees and Valdez (1998), summarized \(K_c(t)\) test suggested by Genest & Rivest (1993) to recognize the form of generator function in Archimedean copulas, and concluded that Gumbel-Hougaard is the preferred copula compared to the other two. Jackie (2006) also proposed Gumbel-Hougaard as a better copula to apply to dependence structure of non-life insurance losses comparing with Frank and Clayton. As our data represent a fat-tailed distribution with extreme values positively skewed, and has upper tail dependence,
we will use Gumbel-Hougaard copula to fit and simulate bivariate distributions with marginal GPD.

2.5. Subadditivity of VaR in Tail Region

As we discussed in literature, Danielson et al (2013) proposed a subadditivity property of VaR for special class of fat-tailed distribution of asset returns.

**Proposition:** “Suppose that $X$ and $Y$ are two asset returns with jointly regularly varying† non-degenerate tails with tail index $\xi > 0$. Then VaR is a subadditive sufficiently deep into the tail region.”

They also proved that for different tail indices and/or their degenerate version in above proposition with $\xi_X, \xi_Y < 1$, VaR is still weakly subadditive and

$$\lim_{p \to 0} \sup \frac{\text{VaR}_p(X + Y)}{\text{VaR}_p(X) + \text{VaR}_p(Y)} \leq 1$$

We are going to check numerically whether fitted GPD for different types of insurance loss random variables belongs to this proposition to achieve subadditivity of VaR in tail region or not. In non-life insurance, it’s usual to have different losses. As we obtained in the parameter estimation section, all fitted distributions of stop-loss and excess of loss material damage ($X$) and bodily injury ($Y$) risks had estimated degenerately shape parameters (tail index). GPD also satisfies regular variation as below:

$$\lim_{t \to \infty} \frac{1 - G_{u,\xi,\sigma}(tx)}{1 - G_{u,\xi,\sigma}(t)} = \lim_{t \to \infty} \frac{1 - [1 - (1 + \frac{\xi(tx - u)}{\sigma})^{-1}]^{-1}}{1 - [1 - (1 + \frac{\xi(t - u)}{\sigma})^{-1}]^{-1}} =$$

Under the transformation $\frac{1}{\xi} = \alpha$, if $\xi < 0$, then $\alpha > 0$ and we have

$$\lim_{t \to \infty} \frac{1 - G_{u,\xi,\sigma}(tx)}{1 - G_{u,\xi,\sigma}(t)} = \lim_{t \to \infty} \frac{(1 - \frac{tx - u}{\alpha \sigma})^{\alpha}}{(1 - \frac{t - u}{\alpha \sigma})^{\alpha}} = \lim_{t \to \infty} \frac{(\alpha \sigma - tx + u)^{\alpha}}{(\alpha \sigma - t + u)^{\alpha}}$$

† $F(x)$ is regularly varying if $\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^\alpha$, $\alpha \in \mathbb{R}$, $x > 0$
Subadditivity and Parameter Uncertainty …

\[
\lim_{t \to \infty} \frac{-\alpha x (\alpha \sigma - t x + u)^{\alpha-1}}{-\alpha (\alpha \sigma - t + u)^{\alpha-1}} = \lim_{t \to \infty} \frac{-\alpha x^2 (\alpha \sigma - t x + u)^{\alpha-2}}{-\alpha (\alpha \sigma - t + u)^{\alpha-2}} = \ldots
\]

\[= x^\alpha = x^{-\xi}.\]

Where in the last line we have used L’Hopital’s rule \(\alpha\) times. So we can say that marginal fitted GPD distribution is regularly varying. By the same method we may prove that Gumbel-Hougaard copula is also a regularly varying copula.

As we have degenerated (negative and/or less than 1) estimated shape parameter, based on weakly subadditivity definition, we expect an increase in subadditivity if we reduce \(p\) (let say from 0.1 to 0.05 and 0.005). We will examine this by choosing different levels of \(p\) in simulation.

2.6. Monte Carlo study of VaR Subadditivity

In this section, we will provide the Monte Carlo simulation procedure to calculate \(VaR(X), VaR(Y)\) and \(VaR(X + Y)\) in different threshold levels to study violations of subadditivity through the tail of bivariate loss distribution. To perform the Monte Carlo simulation there are couple of things to be done as follows:

1) Fitting appropriate distribution (GPD) to marginal losses (risks), estimating parameters and measuring their correlation/dependence for each threshold pair.

2) Fitting a proper copula to risks, estimating measure of association and producing dependent bivariate and uniform distributions.

3) Applying inverse fitted distribution function to produce simulated losses of both risks \(X\) and \(Y\) with size \(n\) as well as \(X + Y\).

4) Repeating 1-3, \(N\) times as the size of simulation and counting number of times that \(VaR(X + Y) < VaR(X) + VaR(Y)\) or not.

We work with losses beyond the base threshold as a new original source of data. We examine subadditivity of simulated \(VaR\) in base threshold of different risks such as stop-loss and excess of loss, and then compare it with higher layers by increasing thresholds in right tail.
2.6.1. Simulation Method

To simulate VaR for each threshold level \( k = 1, \ldots, 11 \) and examine its subadditivity, we need to confirm the following cases for each round of simulation:

- Related threshold level \((u_{Xk}, u_{Yk})\), estimated shape parameters \((\xi_{Xk}, \xi_{Yk})\), Scale parameter \((\sigma_{Xk}, \sigma_{Yk})\).
- Set level of \( V_{aR_p} (p) \), size of sample in each round of simulation \((n)\), and size of simulation \((N)\).
- Estimated Kendall’s tau for each threshold \( k \), \( \tau_k \).

Among all possible simulation methods, based on the nature and limitation of our loss data, which is full of jumps and unbalanced observations, we proposed the following procedure:

1) Choosing \((u_{X1} = 5, u_{Y1} = 40)\) as the base threshold for damage and injury risks, \((X, Y)\). Thus we use Stop-loss random variables \( X_{+5} \) and \( Y_{+40} \) as base distribution in right tail.

2) Making a vector of thresholds (let’s say vector dim is \((T \times 1)\)) for material damage \((X)\) and bodily injury \((Y)\) risks starting from their base threshold as \( U_X = (u_{x1} = d_x = 5, u_{x2}, \ldots, u_{xk}, \ldots, u_{xT}) \), \( U_Y = (u_{y1} = d_y = 30, u_{y2}, \ldots, u_{yk}, \ldots, u_{yT}) \). For each threshold level \( k \) of \( X \) and \( Y \), we have \( n_{Xk} \) and \( n_{Yk} \) number of observations, respectively.

3) Fitting \( GPD \) distribution to data of both risks \( X \) and \( Y \) in each threshold level \( k \), and estimating shape parameter \((\xi_{Xk}, \xi_{Yk})\), Scale Parameter \((\sigma_{Xk}, \sigma_{Yk})\). We have also location parameter / threshold \((u_{Xk}, u_{Yk})\).

4) For related data in each threshold level, using simulated Kendall’s \( \tau_k \) that we have got in previous sections, to calculate the parameter \( \alpha \) in Gumbel-Huggard copula, based on the formula of Tau and theta \((\alpha_k = \frac{1}{1-\tau_k})\).

5) For each threshold level, simulating Gumbel-Huggard copula with \( \alpha_k \) and generating a sample with size \( n \) for \( X \) and \( Y \) in a form of \((x_1, \ldots, x_n)\) and \((y_1, \ldots, y_n)\). Thus size of simulated sample for all
thresholds is the same. Then calculating $VaR_p(X)$, $E(X)$, $VaR_p(Y)$ and $E(Y)$ for generated samples.

6) After having simulated $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, we may easily form $(x_1 + y_1, \ldots, x_n + y_n)$ as a simulated sample of $X + Y$, and calculate $VaR_p(X + Y)$ and $E(X + Y)$ for generated sample.

7) Doing (5) and (6) for all threshold levels of data N times and calculating each time required statistics $VaR_p$ and $E(.)$ for $X$, $Y$ and $X + Y$.

8) In each threshold level of data, among N triples of $VaR_p(X)$, $VaR_p(Y)$ and $VaR_p(X + Y)$ counting those of subadditive ones and seeing how percentage of subadditivity can vary in each level.

In this method:

- We update GPD to fit to each threshold level with new parameters that may bring more accurate fit for each threshold level to estimate and simulate related copula.
- We use different correlation measures in each special threshold level to simulate copula to allow for varying dependence structures through the tail.
- The size of sampling will be unique (n) in all thresholds.

The last case above can be improved by an average of available sample size in original data set for each of risks $X$ and $Y$ to simulate each sample of threshold $k$ equal to $n_k = \frac{n_{xk} + n_{yk}}{2}$ which allows for different sizes in different thresholds, although this doesn’t play an important role in simulation. However, in case of bootstrapping for uncertainty measurement, sample size in different thresholds will be always varying and decreasing through the higher layers of loss distribution.

2.7. Uncertainty Measurement

Based on the theory of maximum likelihood, maximum likelihood estimation of parameters’ vector (px1) follows the p-variate normal distribution asymptotically, under the certain condition of regularity.
McNeil (1997) showed that if $\xi > -0.5$, $GPD$ also satisfies the regularity condition of maximum likelihood estimation (In our case $p = 2$). He expressed that as below:

$$n^{1/2}(\hat{\xi}_n \sigma_n^2) \xrightarrow{d} N[(\xi, \sigma), (\frac{(1 + \xi)^2 \sigma(1 + \xi)}{\sigma(1 + \xi)} \frac{2\sigma^2(1 + \xi)}{\sigma(1 + \xi)} \frac{(1 + \xi)^2 \sigma(1 + \xi)}{\sigma(1 + \xi)} \frac{2\sigma^2(1 + \xi)}{\sigma(1 + \xi)} \frac{(1 + \xi)^2 \sigma(1 + \xi)}{\sigma(1 + \xi)} \frac{2\sigma^2(1 + \xi)}{\sigma(1 + \xi)}]$$

Although we are not interested in uncertainty of above parameters, as $VaR_p$ and $SCR_p$ are functions of these parameters, we need to make inference about them. We may measure standard error of $VaR_p$ and $SCR_p$ in 2 methods:

1) Asymptotic Normality

2) Bootstrapping

In asymptotic normality method, we may recognize asymptotic normal distribution by substitution of estimated parameters for each threshold level of loss data. Then, we may produce $B$ pairs of parameters from this normal distribution and use them to generate enough data point from $GPD$ and calculate $VaR_p$ and $SCR_p$ for each series. The problem is that when we assume $(\xi, \sigma) \sim Normal$, we may obtain both positive and negative parameters in simulation which may be out of defined $GPD$ range of parameters. For example, for negative $\hat{\xi}$, we need to truncate simulated $GPD$ data at $-\sigma/\xi$ in the right tail, and this may lead to some asymmetry in $VaR$ estimation. Also, if $\sigma < 0$, then we cannot produce any $GPD$ simulation to calculate $VaR$ or $SCR$.

The other method is “Bootstrapping” that is built on the actual data, and the result of simulation will be always consistent with them. Bootstrapping is based on the sampling by replacement from any set of data with the same size, but probably with repeated observations in the sample. If we do it enough times, we may obtain any statistics (especially standard error of $VaR$ or $SCR$) we need to estimate from the original sample. As this method does not produce any inconsistent result, we will use it to measure uncertainty of $VaR_p$ and $SCR_p$ for each threshold level, and to see how it goes through the tail of distribution.
2.7.1. Bootstrapping Method

In case of $VaR_p$, we have three statistics to measure their uncertainty, $VaR_p(X)$, $VaR_p(Y)$ and $VaR_p(X + Y)$, thus we must simulate standard error of $VaR$ for each of them. We do this in each threshold level.

For $VaR_p(X), VaR_p(Y)$, size of bootstrapping in each threshold is the sample size of data in that threshold level ($n_{x_k}$ for $k$th threshold of $X$ and $n_{y_k}$ for $k$th threshold of $Y$) in original sample. In case of $VaR_p(X + Y)$, we will use a simulated sample with size of 300,000 that we produced in correlation measurement to estimate Kendall’s $\tau$ for each threshold. In that sample, we obtained enough pairs of data in each threshold pair $(u_{x_k}, u_{y_k})$ which was $n_{(x,y)_k}$, and we will do bootstrapping with this size of $n_{(x,y)_k}$ for $X + Y$ in each threshold pair.

In bootstrapping the number of sampling (or simulation) is the same for all groups and thresholds, but the size of sampling will be different as mentioned above. We should use original data for sampling by replacement.

We constructed the bootstrapping process as follows:

1) In each threshold level $k$, we have $n_{x_k}, n_{y_k}$ and $n_{(x,y)_k}$ for each loss random variable $X_{ik}, Y_{ik}$ and $(X + Y)_{ik}$. We should do sampling with replacement with corresponding size for each threshold level of data. We will have each sample with replacement in the form of: $(x^*_1, ..., x^*_{n_{x_k}})$, $(y^*_1, ..., y^*_{n_{y_k}})$ and $((x + y)^*_1, ..., (x + y)^*_{n_{(x,y)_k}})$.

2) In the threshold level $k$, we assign the Kaplan-Meier Product Limit probability of $s_i/n_{x_k}$, $s_i/n_{y_k}$ and $s_i/n_{(x,y)_k}$ to each observation of $X_{ik}, Y_{ik}$ and $(X + Y)_{ik}$ where $s_i$ is the number of repeated times for observation $i$. (We use non equal probabilities for the observations that are repeated more than once.)

3) We do sampling with replacement for each threshold level $B$ times. In each sample of $X^*_k, X^*_k$ and $(X + Y)^*_k$, we can calculate
VaR_p and \( \mu \) to form the \( SCR_p = VaR_p - \mu \). Thus we have \( B \) number of \( VaR_p \) and \( SCR_p \).

4) For each random variable, the standard error of simulated \( VaR_p \) and \( SCR_p \) in \( B \) samples we took, can be a measure of uncertainty as below:

\[
Se_B(VaR_p) = \sqrt{\frac{\sum_{i=1}^{B}(VaR_i - \bar{VaR})^2}{B-1}}
\]

We will perform all bootstraps by \( B = 450 \) repeating, which is a relatively good size of bootstrapping replication.

### 3. Analysis of Material Damage/Bodily Injury Loss Data

#### 3.1. Exploratory Data Analysis

We start with exploratory analysis to make some preliminary sense of the form and basic characteristics of the loss data. The analysis is performed for three preliminary forms of loss data, Ground-up loss, Stop-loss in base thresholds and Aggregate Daily Loss, for both material damage (\( X \)) and bodily injury (\( Y \)) risks. In aggregate daily loss, we added total losses that happened together in every single day. This can be useful to track dependency of two risks during time, if there is any time depending factor affecting the risks.

Figure (1) exhibits a histogram for both material damage and bodily injury losses with a \( GPD \) fit for each histogram based on a maximum likelihood estimation of \( GPD \) parameters. The data is skewed to the right tail and has pretty asymmetric distribution reporting a typical form of the non-life loss distribution. We don’t observe any clustering in huge losses in right tail of distribution which supports the i.i.d assumption of observations to establish maximum likelihood estimation of parameters. Also there are frequent extreme losses in the tail, which supports the idea to use extreme value techniques to analyze data.
3.1.1. Q-Q Plots

To compare shape of distribution in different layers, we divided right tail of loss distribution into 11 threshold levels beginning from a base threshold which is 5 Mn Rials for material damage loss and 40 Mn Rials for bodily injury loss. Figure (2) and (3) represent Q-Q plots against the exponential and generalized Pareto distribution to examine the shape of distribution in lower and higher levels of the right tail.
Figure 2: Q-Q plot for Stop-loss random variable beyond the base threshold \((u_1)\) for (a) Damage risk \((X)\) beyond \((u_{x1}=5)\) against exponential, (b) Injury risk \((Y)\) beyond \((u_{y1}=40)\) against exponential, (c) Damage risks \((X)\) beyond \((u_{x1}=5)\) against GPD and (d) Injury risk beyond \((u_{y1}=40)\) against GPD.

Figure (2) illustrates the Q-Q plot for material damage and bodily injury, and Stop-loss (SL) distribution for base threshold, against the exponential and fitted GPD. We can recognize the possibility of having negative shape parameter for both risks in tail region which are under exponential in tail region. But compared with fitted GPD, we may see that in left tail, they are well fitted while in right tail they still have deviation from fitted.

In Figure (3), we provided Q-Q plot for stop-loss of material damage and bodily injury risks beyond 8\(^{th}\) and 11\(^{th}\) thresholds. We observed that for both 8\(^{th}\) and 11\(^{th}\) threshold levels we obtained better fitting of GPD than base threshold for stop-loss observations. The figures show that observed losses are still a little bit more skewed than GPD, but their tails are less fat than related GPD. Generally saying, although they are not perfectly fitted yet by maximum likelihood GPD, we observe fewer data in tail deviated from the GPD, and we may expect that going more deep into the tail the empirical distribution in higher thresholds level deviates less from GPD.
3.1.2. Mean Excess Function

Assuming $GPD$ for loss data, mean excess function and threshold $u$, have a linear relationship with slope of shape parameter $\left( E[X - u/X > u] = \frac{\sigma + \xi u}{1-\xi} \right.$ for $\xi < 1$). Therefore, if beyond some thresholds
the sample mean excess function is downward (upward) sloping we may conclude the existence of a negative (positive) shape parameter.

Figure 4: Sample mean excess function for excess of loss risk for (a) damage loss beyond the base threshold ($u_{x1}=5$), (b) injury loss beyond the base threshold ($u_{y1}=40$), (c) damage loss beyond the last threshold ($u_{x11}=47$), (d) injury loss beyond the last threshold ($u_{y11}=1990.5$) - Mn Rials

Figure (4) shows the sample mean excess function for Excess of loss of material damage and bodily injury losses beyond the base and the last threshold. As we may see for base threshold level, different layers of the excess of loss data represent sample mean excesses with different slopes. In some layers they are downward, and in some others we observe upward sloping function against the thresholds.

The multiple step form of mean excess function is due to existence of huge frequency of repeated measures in data as most of claim settlers tend to round the amount of reported or paid loss. As shape parameter varies through different thresholds (here it varies between
base and last threshold), We may infer from the sample mean excess function in figure (4) that we should update estimation of parameters when we are going to fit $GPD$ distribution for loss data in each threshold level.

### 3.1.3. Setting the Thresholds’ level

We set several thresholds $u_k$, $k = 1, ..., T$ starting from a base point which is basically in the middle of original loss distribution rather than too much extreme or deep into the tail and will be extended along the right tail of distribution until a point that is relatively extreme and there are just relatively few observations beyond that. So basically our threshold set will cover more than half of the density of both loss distributions.

The reason to set such a threshold set may be as follow:

1. We can study difference between the behavior of lower threshold and the so called “sufficiently higher thresholds” (as mentioned in Belkema & de Haan 1974 and Pickands 1975 theorem).

2. To have better view into variation of dependency between two risks through the tail and how it may affect the theoretical foundation of subadditivity.

We set threshold levels for both risks $X$ and $Y$. Thus, in bivariate setting we should study tail of both distribution simultaneously. Therefore we have pairs of thresholds $(u_{Xk}, u_{Yk})$ and going deeper into the tail refers that we will study conditional random variables larger than related thresholds in both risks rather than just one of them. Later we will see how this requirement will affect a huge drop out in sample size of bivariate loss to study dependence structure and also simulation of $X + Y$. 


The suggested threshold set for the data set is provided in table (2) as below:

**Table 2: Threshold Set for Both Damage and Injury Risks**

<table>
<thead>
<tr>
<th>k</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>X: u_{sk}</td>
<td>5</td>
<td>9.2</td>
<td>13.4</td>
<td>17.6</td>
<td>21.8</td>
<td>26</td>
<td>30.2</td>
<td>34.4</td>
<td>38.6</td>
<td>42.8</td>
<td>47</td>
</tr>
<tr>
<td>Y: u_{yk}</td>
<td>40</td>
<td>235.1</td>
<td>430.1</td>
<td>625.2</td>
<td>820.2</td>
<td>1015.3</td>
<td>1210.3</td>
<td>1405.4</td>
<td>1600.4</td>
<td>1795.5</td>
<td>1990.5</td>
</tr>
</tbody>
</table>

Table (3) also represents concise information about position and coverage of thresholds in overall loss distribution. The threshold sets cover 67.2% and 43.8% of probability distribution of Damage and Injury risks respectively. There are also enough observations beyond the last threshold of both risks and as we are interested to estimate reasonable \( \text{VaR}_{99.5\%} \) it is important to have more than 100 data points to make to estimate 99.5% quantile.

**Table 3: Thresholds quantiles, coverage ratios of threshold range and number of data beyond the last threshold**

<table>
<thead>
<tr>
<th>Type of Risk</th>
<th>Base Threshold Amount</th>
<th>Base Threshold quantile</th>
<th>Last threshold Amount</th>
<th>Last threshold quantile</th>
<th>increment</th>
<th>Density coverage</th>
<th>Number of data beyond last threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damage</td>
<td>5</td>
<td>32.7%</td>
<td>47</td>
<td>99.9%</td>
<td>4.2</td>
<td>67.2%</td>
<td>186</td>
</tr>
<tr>
<td>Injury</td>
<td>40</td>
<td>55.3%</td>
<td>1990.5</td>
<td>99.1%</td>
<td>195.05</td>
<td>43.8%</td>
<td>193</td>
</tr>
</tbody>
</table>

3.1.4. Special Considerations in measuring association in non-life business

Jackie (2006) argued that usually claims liability data are insufficient for estimation of association measures such as Kendall’s \( \tau_{X,Y} \) to capture the inherent dependency between different lines of business and mentioned that correlation assessment will be judgmental and can become more feasible only when more data are collected in the future or when an insurer has a long history and keeps a good track of relevant claims data records. The nature of data we use in this research is not consistent with usual data to measure correlation.

To measure correlation between two random variables we should have pairs of observations, whereas in our case, most of the pairs are incomplete as we have not observed losses for both risks (damage and
Subadditivity and Parameter Uncertainty ...

body injury) in each policy of loss file. Also, we cannot assume amount of loss to be zero for those pairs without losses in one of the lines of business. If we do so, the result will be an artificial negative correlation between the lines of business as for most of them we have a positive value versus zero. Reasonably we expect positive correlation between two insurance risks rather than negative one because we don’t expect that if losses in one line of business increase, in general, the other line will have less amount of loss. To deal with this problem we may use only those records that have complete pairs of observation in both risks. Although this method is straightforward and easy, we lose lots of observations that may be effective on the measure of association between two risks.

As another method we may keep track of time along with losses in both lines of business instead of loss files. In a daily basis we should think of amount of losses reported every day in each line and finally just measure correlation between them to see how they behave together. In this method all observations are participated in final measure of association. In table (4) the estimated Kendall’s tau is reported for both methods and different forms of data such as ground up loss, stop-loss and excess of loss coverage.

Table 4: Estimation of Kendall’s tau and Spearman correlation coefficients for complete pair losses and aggregate daily losses provided for different forms of data such as ground up loss, stop-loss and excess of loss coverage.

<table>
<thead>
<tr>
<th>Method</th>
<th>Complete pairs</th>
<th>Daily Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of observations</td>
<td>Kendall’s Tau</td>
</tr>
<tr>
<td>Original Loss</td>
<td>9107</td>
<td>0.085**</td>
</tr>
<tr>
<td>Stop-loss X_+, Y_+</td>
<td>6592</td>
<td>0.081**</td>
</tr>
<tr>
<td>Excess of loss (X – d_x)<em>+, (Y – d_y)</em>+</td>
<td>6592</td>
<td>0.081**</td>
</tr>
</tbody>
</table>

** Correlation estimations are significant at the level of 1%.

We can see that Kendall’s τ is invariant for any comonotonic transformation such as stop-loss and excess of loss reformation of loss.
3.1.5. Measuring Dependency through the Tail

In multivariate structure we always need to measure tail dependence as it may differ significantly from dependency structure of overall distribution. In case of insurance loss we are interested just in measuring dependence in upper tail. We need to simulate a proper copula to construct bivariate structure of loss distributions in each threshold pair through the tail. To achieve this, we need to measure dependency of losses beyond each threshold pair.

As our data set is unbalanced, we have just 9107 policies with joint losses which is relatively scarce comparing marginal loss observations. As lots of pairs are different in magnitude, when we increase both thresholds measuring correlation is impossible after 4th threshold level as there is no data left. We may assume the same correlation coefficient of base threshold for higher ones, but this will mislead copula fit if there would be a possible dependency variation of losses beyond different thresholds.

To solve the problem we may simulate empirical loss distribution using Monte Carlo simulation by copula to obtain enough dependent data in higher thresholds. We prefer empirical distribution to parametric one to avoid assigning the bias we may produce by inaccurate fit for higher layers by parametric fit of base layer where they may be different. To do this we use estimation of correlation coefficients for base threshold and empirical distribution functions of both damage ($X$) and injury ($Y$) losses in procedure below:

1) We estimate Kendall’s, Spearman and Pearson correlation coefficients for joint observations we have in our original data ($\hat{\rho} = 0.09$).

2) We estimate marginal empirical distributions by Kaplan-Meier Product Limit method using all observations beyond base thresholds ($u_{X1}, u_{Y1}$) we have in each loss rather than just joint ones.

3) Calculate empirical inverse distribution function in form of a stair function.

4) Use bivariate t-copula with $\nu = 5$ and $\hat{\rho} = 0.09$ to simulate a bivariate uniform distribution using with 300,000 data point.
5) Apply empirical inverse CDF to produce a simulated sample of 300,000 dependent data points by minimum value of base thresholds.

As we simulated enough data points in each risk, we quit to smooth the inverse empirical CDF. We have enough simulated data to measure correlation in higher layers beyond each joint threshold level which is estimated in table (5) as follows:

<table>
<thead>
<tr>
<th>Threshold</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td># Observations</td>
<td>300,000</td>
<td>42,661</td>
<td>6,949</td>
<td>7,007</td>
<td>1,957</td>
<td>787</td>
<td>443</td>
<td>277</td>
<td>167</td>
<td>109</td>
<td>72</td>
</tr>
<tr>
<td>Kendall's</td>
<td>0.091</td>
<td>0.121</td>
<td>0.144</td>
<td>0.14</td>
<td>0.142</td>
<td>0.145</td>
<td>0.186</td>
<td>0.182</td>
<td>0.144</td>
<td>0.211</td>
<td>0.164</td>
</tr>
<tr>
<td>Spearman</td>
<td>0.132</td>
<td>0.178</td>
<td>0.21</td>
<td>0.201</td>
<td>0.208</td>
<td>0.212</td>
<td>0.265</td>
<td>0.266</td>
<td>0.221</td>
<td>0.31</td>
<td>0.243</td>
</tr>
<tr>
<td>Pearson</td>
<td>0.154</td>
<td>0.277</td>
<td>0.296</td>
<td>0.29</td>
<td>0.276</td>
<td>0.283</td>
<td>0.312</td>
<td>0.339</td>
<td>0.285</td>
<td>0.299</td>
<td>0.21</td>
</tr>
</tbody>
</table>

As we see the estimated Kendall’s τ is not unique for all joint loss layers and it varies between (0.09, 0.21) through the tail of bivariate empirical distribution. Later we will use each \( \hat{\tau} \) to estimate copula for each threshold level in simulation of Value at Risk.

4. Results

We provide numerical results about subadditivity of VaR and how it can be affected with different factors. We will show if GPD is stable by changing the type of loss random variable from Stop-loss to excess of loss and changing the thresholds.

4.1. Parameter Estimation

Figures (5) & (6) illustrates the estimated GPD parameters for different thresholds for Stop-loss (SL) and Excess of loss (EL) loss random variables for both Damage (X) and Injury (Y) risks.

In Figure (5), for both stop-loss Damage and Injury risks, we observe generally negative shape parameters \( \xi_k < 1 \) that shows they follows super fat-tailed distributions and as we expect, the deeper into the tail the fatter the tail we capture. For excess of loss case, the shape parameter is positive but still the same super fat-tailed property of distribution as \( \xi_k < 1 \). We observe that GPD turns to more stability on
shape, when we transform stop-loss random variable to excess of loss, while there is significant difference between shape parameters of different thresholds. This justifies the idea of not to fit a general GPD distribution to all thresholds. Also we see that the excess of loss shows more fluctuations than stop-loss.

Figure 5: Variation of Estimated shape parameter versus threshold levels for both Stop-loss (SL) and Excess of loss (EL) distributions for damage (X) and Injury (Y) risks

Figure (6) represents the similar information about the scale parameter in different threshold of the tail. For both stop-loss and excess of loss random variables we observe a gradual increase in scale parameter of GPD fit, when we go through the higher thresholds deeper in the tail. On the other hand we observe generally huge increase for the estimated scale parameter in higher thresholds (deeper in tail), but growth in scale parameter of stop-loss random variable is
much steeper than what we observed for excess of loss random variable.

![Graph showing variation of estimated scale parameter through the tail for both stop-loss (SL) and excess of loss (EL) distributions for damage (X) and injury (Y) risks.]

**Figure 6:** Variation of estimated scale parameter through the tail for both stop-loss (SL) and excess of loss (EL) distributions for damage (X) and injury (Y) risks

### 4.2. Subadditivity Measurement

We apply the simulation method with respect to the following parameters and criteria:

- \( p = 0.005, \ p = 0.05 \)
- Stop-loss (SL) & excess of loss (EL) random variables
- Three levels of correlation: Estimated \( \hat{t}_k \) and overall correlations: \( \tau = 0, \tau = 0.5 \)

In case of overall Kendall’s \( \tau \) we will use the same amount for all threshold levels of SL/EL loss random variable in copula simulation whereas in case of varying estimated \( \hat{t}_k \) through the tail it will be
changed in each threshold. We do this to capture overall and through-the-tail effect of correlation on subadditivity, simultaneously.

For all cases the number of simulation is \( N = 50,000 \), and each simulation has been performed by sample size \( n = 2,000 \). Thus, each simulation is performed for risks \( X \) and \( Y \) (while \( X + Y \) will be calculated simply by adding them together for each data point) regarding the following components:

1) Threshold level: \( (u_{xk}, u_{yk}) \), \( k = 1, 2, ..., 11 \)
2) Estimated Shape parameter: \( (\hat{\xi}_{xk}, \hat{\xi}_{yk}) \), \( k = 1, 2, ..., 11 \)
3) Estimated Scale parameter: \( (\hat{\sigma}_{xk}, \hat{\sigma}_{yk}) \), \( k = 1, 2, ..., 11 \)
4) Kendall’s correlation coefficient: \( \hat{\tau}_k \), \( k = 1, 2, ..., 11 \)
5) Level of \( p \),
6) \( N=50,000 \), \( n=2,000 \)

In each table we provide number of subadditivity violations out of \( N = 50,000 \), for both \( VaR_p \) and \( SCR_p \) as well as subadditivity ration which is equal to proportion of subadditive triples of \( (VaR_p(X), VaR_p(Y), VaR_p(X + Y)) \) out of \( N \).

In Table (6) we provided result of computations about one roll of simulation procedure for a set of components above with the special \( p = 0.005 \), for stop-loss random variable. First of all, we observe that number of subadditivity violations and ratio is the same for both \( VaR \) and \( SCR \). This implies that Subadditivity (or superadditivity) of \( VaR \) preserves the subadditivity (or superadditivity) of \( SCR \) and we can make inference in case of \( SCR \) by \( VaR \).

Generally going through higher threshold levels (deeper into the tail), we observe more subadditivity ratio (less violations) that is consistent with proposition given in Danielsson et al. (2013), whereas in our case there is no need to equal tail index as shape parameter of two risks is always different.
**Table 6: Subadditivity of VaR & SCR at 99.5% level, through the right tail of Stop-loss (SL) damage & Injury risks in third party vehicle liability insurance policy for an Iranian Insurance Company.** Columns 1-11, represents threshold level index \( k \), from base threshold to last threshold for both risks \( X \) and \( Y \). Number of subadditivity violations and ratios is provided in three overall level of correlation consist of \( \tau = 0 \), estimated tau (\( \hat{\tau}_k \)) and \( \tau = 0.5 \) in Rows 3-6, 8-11 and 13-16, respectively. Row 7 represents the estimation of \( \tau \) for each threshold level and cells below each \( \hat{\tau}_k \) contains number of respective violations and subadditivity ratio, for VaR and SCR.

Loss Random Variable: SL, Probability Level: \( p = 0.005 \), Simulation Size: \( N = 50,000 \), Sample Size: \( n = 2,000 \).

<table>
<thead>
<tr>
<th>Threshold level ( k )</th>
<th>Tau=0</th>
</tr>
</thead>
<tbody>
<tr>
<td>VaR</td>
<td>SCR</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Estimated Tau ( \hat{\tau}_k )</td>
<td>0.091</td>
</tr>
</tbody>
</table>

| Tau=0.5 | 
|-------------------------|--------|
| VaR | SCR |
| 1 | 528 | 623 | 767 | 641 | 454 | 264 | 360 | 156 | 24 | 68 | 8 |
| 2 | 529 | 623 | 768 | 641 | 455 | 264 | 360 | 156 | 24 | 68 | 8 |
| Estimated Tau \( \hat{\tau}_k \) | 0.091 | 0.121 | 0.144 | 0.14 | 0.142 | 0.145 | 0.186 | 0.182 | 0.144 | 0.211 | 0.164 |

**4.2.1. Effect of Correlation (Tail Dependence)**

We observe significant difference in subadditivity ratio of VaR and SCR by overall and through the tail correlation. In this case, we observe that generally higher dependence; measured by Kendall’s...
correlation, endangers subadditivity of \( \text{VaR} \) and \( \text{SCR} \) and leads to relatively huge subadditivity violations. Independent risks preserves almost fully subadditive \( \text{VaR} \) and \( \text{SCR} \). In our case a 50% correlation decreases subadditivity ratio approximately by 8-20 percent (4,000-10,000 subadditivity violations out of 50,000) that is considerable to our compared to reasonable confidence we may need in Solvency II directives. Figure (7) represents the overall subadditivity violations of \( \text{VaR} \) versus different correlation levels.

![Figure 7: Subadditivity Violations through the tail of joint loss distribution for three overall correlation level; \( \tau=0 \), estimated \( \tau \) and \( \tau=0.5 \). \( N=50,000, n=2,000, p=0.005 \).](image)

As explained we observe significantly higher violations for \( \tau = 0.5 \) comparing \( \tau = 0 \) and \( \hat{\tau}_k \). Also when we go through the tail of joint loss distribution of material damage (\( X \)) and bodily injury (\( Y \)) risks, we observe that tail dependence measured by Kendall’s tau correlation in different thresholds can affect subadditivity of \( \text{VaR} \) and \( \text{SCR} \). Looking at rows 8-11 in table (6) through threshold levels we observe relatively more violations related to higher correlation in tail (tail dependence).
Figure (8) also represents the effect of through the tail correlation (tail dependence) in different threshold levels on subadditivity violations of VaR and we more violations for higher correlations. However, going deeper into the joint tail, negative effect of correlation gets neutralized by positive effect of higher threshold levels on subadditivity of VaR. Especially we observe that in initial threshold levels high correlation, force more violations whereas in higher relatively last thresholds, the positive effect of fatter tails dominate the negative effect of correlation.

![Figure 8: Subadditivity violations & estimated correlation through the tail of joint loss distribution for stop-loss (SL) random variable. N=50,000, n=2,000, p=0.005.](image)

We have provided other versions of table (8), tables 9-11, for Excess of loss random variable (EL) and probability level of $p = 0.05$.

4.3. Uncertainty Measurement

We applied bootstrapping method to measure the standard error of VaR through the tail as a measure of uncertainty. We applied the method for all of $VaR_p(X)$, $VaR_p(Y)$ and $VaR_p(X + Y)$. Table (12) represents the result of uncertainty measurement through the tail of $X$, $Y$ and $X + Y$ along the 11 threshold levels.
Table 12: Uncertainty of $\text{VaR}_p(X)$, $\text{VaR}_p(Y)$ and $\text{VaR}_p(X+Y)$ through the tail of Stop-loss (SL) random variable. Columns 2 and 3 represent the threshold level of each risk $X$ and $Y$ respectively. In case of $X+Y$ threshold level is the pair of numbers in column 2 & 3. Column 3-5 represents the sample size for each threshold level. Columns 7-9 provides the standard error of bootstrapped $\text{VaR}_p(X)$, $\text{VaR}_p(Y)$ and $\text{VaR}_p(X+Y)$.

<table>
<thead>
<tr>
<th>k</th>
<th>Threshold Level</th>
<th>Sample Size</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>$u_{x,k}$</td>
<td>$u_{y,k}$</td>
<td>$n_{x,k}$</td>
</tr>
<tr>
<td>1</td>
<td>5.0</td>
<td>40.0</td>
<td>143596</td>
</tr>
<tr>
<td>2</td>
<td>9.2</td>
<td>235.1</td>
<td>72099</td>
</tr>
<tr>
<td>3</td>
<td>13.4</td>
<td>430.1</td>
<td>40357</td>
</tr>
<tr>
<td>4</td>
<td>17.6</td>
<td>625.2</td>
<td>16246</td>
</tr>
<tr>
<td>5</td>
<td>21.8</td>
<td>820.2</td>
<td>4800</td>
</tr>
<tr>
<td>6</td>
<td>26.0</td>
<td>1015.3</td>
<td>2387</td>
</tr>
<tr>
<td>7</td>
<td>30.2</td>
<td>1210.3</td>
<td>1347</td>
</tr>
<tr>
<td>8</td>
<td>34.4</td>
<td>1405.4</td>
<td>813</td>
</tr>
<tr>
<td>9</td>
<td>38.6</td>
<td>1600.4</td>
<td>413</td>
</tr>
<tr>
<td>10</td>
<td>42.8</td>
<td>1795.5</td>
<td>269</td>
</tr>
<tr>
<td>11</td>
<td>47.0</td>
<td>1990.5</td>
<td>186</td>
</tr>
</tbody>
</table>

* Based on Historical Simulation of original data

Figure (9) provides a graphical view of the variation of uncertainty of $\text{VaR}_p(X)$, $\text{VaR}_p(Y)$ and $\text{VaR}_p(X+Y)$ along different threshold levels when we go deep into the tail of loss distribution. In case of material damage risk $X$, uncertainty of $\text{VaR}_p$ increases steadily through the tail but at last thresholds it stops to increase. In case of $\text{VaR}_p(Y)$ and $\text{VaR}_p(X+Y)$ we observe that in initial thresholds uncertainty increases considerably. But when we go much deeper into the tail in higher threshold levels, although we expect more uncertainty due to the fewer observations, we observe that uncertainty starts to decrease. The reasoning for this phenomenon may be expressed as when we have fewer data in sample in last thresholds, VaR as a value at the end of the tail of distribution can be selected from more limited range than middle thresholds and this result in reducing the uncertainty level of VaR in the last thresholds of loss distribution.

Of course in first thresholds or stop-loss random variable we have large size of observation and VaR can be selected from wider range of.
values and simultaneously taking into account the drops in the size of observations may produce more dispersion in VaR.

Figure 9: Variation of VaR uncertainty through the tail of loss distribution for three random variables (X), (Y) and (X+Y).

5. Conclusion

Although generally VaR is not subadditive, in higher layers of fat-tailed distributions we can achieve theoretical and empirical evidences to achieve subadditivity property of it. Regarding the nature of nonlife insurance losses with rightly skewed and fat tail distributions they vary regularly and we may model them by Generalized Pareto Distribution (GPD). We examined subadditivity of VaR and SCR for two real dependent insurance losses provided by the insurer in a unique policy, using 11 threshold levels in right tail of distribution. Although, loss observation of two risks were unbalanced we captured correlation measure through the tail by simulation.

We used an innovative simulation procedure for joint loss distribution combining renewal GPD fitting to each threshold level.
Based on the numerical methods we used to examine the subadditivity of VaR and SCR for damage and injury insurance loss random variables in a real world example, we achieved subadditivity of VaR deep into the tail for one set of insurance loss data, which is consistent with proposition provided by Danielson et al (2012). We also observed relative decrease in subadditivity violations of VaR in higher layers of joint loss distribution.

Based on the comparison of subadditivity violations in different levels of correlation, we observed that there is no deviation between subadditivity ratio of VaR and SCR. Thus Subadditivity of VaR implies subadditivity of $SCR = VaR - E(X)$.

Overall dependency affects subadditivity of VaR so as independent risks have fully subadditive VaRs whereas highly dependent risks have considerable deviation of subadditivity even in tail region.

Based on the numerical results, we observed that sometimes when we go deeper into the tail, we have less subadditivity ratio (more subadditivity violations) which is in contrast with proposition of Danielsson et al (2012). Therefore, we infer that only going deep into the tail will not guarantee to ensure more subadditivity of VaR but also some other factors can affect it.

We conclude that one of the important factors that Danielson et al didn’t take into account going into the tail is probable “variation of dependence measure through the tail of distribution” that can change the subadditivity increasing trend of VaR through the tail. As we showed in a realistic example, the shape of the joint loss distribution is subject to change through the tail by variation of tail index and dependence structure. For example, by changing thresholds to go deep into the tail, tail dependence is not constant and may vary as we estimated significantly different Kendall’s tau for each threshold level. Then when correlation varies, subadditivity deviates from the tendency to increase through the tail. The more correlation we have, the more violation of subadditivity will be seen.

We also considered varying shape parameter through the tail and observed that magnitude (absolute value) of shape parameter is an effective factor on violations of subadditivity of VaR and produce more violations of VaR subadditivity.
The final unexpected result was calibration of uncertainty level of \( VaR \) in different thresholds. We measured standard error of \( VaR \) through the tail of marginal \((X)\) and \((Y)\) as well as joint loss \((X,Y)\) distributions by bootstrapping. We obtained the result of bootstrapping such that uncertainty is not always monotonically increasing through the tail. In our case, it was increasing from base threshold until middle thresholds and then it started to decrease and it reduced through last thresholds. We understood intuitively as when we go deeper into the tail, the frequency of losses decreases. As we are interested in 99.5% quantile, the effect of scarce observations in deeper tail region forces \( VaR \) to be selected among more limited number of extreme points. This may decrease the variability of \( VaR \) estimation in the last thresholds.

References


## Appendixes

### Subadditivity Measurement

Table 9: Subadditivity of VaR & SCR at 99.5% level, through the right tail of Excess of loss (EL) damage & Injury risks in third party vehicle liability insurance policy for an Iranian Insurance Company. Columns 1-11, represents threshold level index k, from base threshold to last threshold for both risks X and Y. Number of subadditivity violations and ratios is provided in three overall level of correlation consist of $\tau = 0$, estimated $\hat{\tau}_k$ and $\tau = 0.5$ in Rows 3-6, 8-11 and 13-16, respectively. Row 7 represents the estimation of $\hat{\tau}_k$ for each threshold level and cells below each $\hat{\tau}_k$ contains number of respective violations and subadditivity ratio, for VaR and SCR.

Loss Random Variable: EL, Probability Level: $p = 0.005$, Simulation Size: $N = 50,000$. Sample Size: $n = 2,000$.

<table>
<thead>
<tr>
<th>Threshold level, $(\tau)$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>$#$ of Violations</td>
<td>Va R</td>
<td>4</td>
<td>6</td>
<td>20</td>
<td>300</td>
<td>48</td>
<td>31</td>
<td>23</td>
<td>42</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>SCR</td>
<td>4</td>
<td>6</td>
<td>20</td>
<td>300</td>
<td>48</td>
<td>31</td>
<td>23</td>
<td>42</td>
<td>14</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>$(\text{Estimated/Tau} \cdot \hat{\tau}_k)$</td>
<td>0.091</td>
<td>0.121</td>
<td>0.144</td>
<td>0.14</td>
<td>0.142</td>
<td>0.145</td>
<td>0.186</td>
<td>0.182</td>
<td>0.144</td>
<td>0.211</td>
<td>0.164</td>
</tr>
<tr>
<td>$#$ of Violations</td>
<td>Va R</td>
<td>689</td>
<td>1013</td>
<td>1636</td>
<td>3856</td>
<td>2202</td>
<td>1945</td>
<td>2615</td>
<td>2710</td>
<td>1340</td>
<td>2476</td>
</tr>
<tr>
<td>SCR</td>
<td>691</td>
<td>1013</td>
<td>1636</td>
<td>3857</td>
<td>2204</td>
<td>1945</td>
<td>2617</td>
<td>2714</td>
<td>1340</td>
<td>2481</td>
<td>1266</td>
</tr>
<tr>
<td>Subadditivity Ratio</td>
<td>Va R</td>
<td>98.6%</td>
<td>98.0%</td>
<td>96.7%</td>
<td>92.3%</td>
<td>95.6%</td>
<td>96.1%</td>
<td>94.8%</td>
<td>94.6%</td>
<td>97.3%</td>
<td>95.0%</td>
</tr>
<tr>
<td>SCR</td>
<td>98.6%</td>
<td>98.0%</td>
<td>96.7%</td>
<td>92.3%</td>
<td>95.6%</td>
<td>96.1%</td>
<td>94.8%</td>
<td>94.6%</td>
<td>97.3%</td>
<td>95.0%</td>
<td>97.5%</td>
</tr>
<tr>
<td>$#$ of Violations</td>
<td>Va R</td>
<td>10864</td>
<td>10498</td>
<td>10999</td>
<td>13133</td>
<td>11651</td>
<td>11230</td>
<td>11146</td>
<td>11534</td>
<td>10458</td>
<td>10177</td>
</tr>
<tr>
<td>SCR</td>
<td>1084</td>
<td>10515</td>
<td>11019</td>
<td>13152</td>
<td>11671</td>
<td>11265</td>
<td>11177</td>
<td>11567</td>
<td>10483</td>
<td>10220</td>
<td>9807</td>
</tr>
<tr>
<td>Subadditivity Ratio</td>
<td>Va R</td>
<td>78.8%</td>
<td>79.0%</td>
<td>78.0%</td>
<td>73.7%</td>
<td>76.7%</td>
<td>77.5%</td>
<td>77.7%</td>
<td>76.9%</td>
<td>79.1%</td>
<td>79.6%</td>
</tr>
<tr>
<td>SCR</td>
<td>78.7%</td>
<td>79.0%</td>
<td>78.0%</td>
<td>73.7%</td>
<td>76.7%</td>
<td>77.5%</td>
<td>77.6%</td>
<td>76.9%</td>
<td>79.0%</td>
<td>79.6%</td>
<td>80.4%</td>
</tr>
</tbody>
</table>