

Original Research Paper

Estimating the Claim Severity Distribution using Variable Neighborhood Search

Kunjira Kingphai and Samruam Chongcharoen

Graduate School of Applied Statistic, National Institute of Development Administration, Bangkok, Thailand

Article history

Received: 22-12-2015

Revised: 09-02-2016

Accepted: 24-10-2016

Corresponding Author:
Kunjira Kingphai
Graduate School of Applied
Statistic, National Institute of
Development Administration,
Bangkok, Thailand
Email: kunjira.k@grads.nida.ac.th

Abstract: In this study, Variable Neighborhood Search (VNS) is utilized to estimate the parameters of actual motor insurance claims data set and compared them obtained by the Moment Estimation Methods (MOM) and Maximum likelihood Estimation Method (MLE) which are known as a conscientious method. Then, the Kolmogorov-Smirnov test (K-S) is used to show how well the selected distribution fits the actual claims. From the results, we found that the lognormal distribution which their parameters were estimated from VNS technique fits the actual motor claims data set better than the other two techniques with significant level 0.01.

Keywords: Variable Neighborhood Search (VNS), Claim Severity Distribution, Maximum Likelihood Estimation Method (MLE), Moment Estimation Method (MOM)

Introduction

Modeling the distribution of claim data is an important task of general insurance companies regularly faced up with. Then, a responsibility about finding an appropriately statistical distribution by using an available claim data and testing how well this statistical distribution fits the data is necessary for the opportunities in predicting future claim severity (Tosaporn and Sattayatham, 2012). One of the key steps in modeling the distribution of this data is to estimate parameters within the models. In this study, we will use a meta-heuristic to estimate the parameters of claim severity distribution. The meta-heuristic is method that used to find a good solution in difficult and complex problem that cannot solve by using the exact method otherwise, this may take a long time to fix its. However, the response from the meta-heuristic method cannot guarantee that will be the best answer as from the exact solution. To find the right answers by meta-heuristic, there have many purposed algorithm such as Simulated Annealing, Tabu Search, Ant Colony Optimization, Particle Swarm Optimization, Variable Neighborhood Search (VNS), etc. Soontorn *et al.* (2013) have evaluated the severity of fire accidents by using a Randomized Neighborhood Search (RNS) method to estimate the Weibull parameters. Babak *et al.* (2011) have maximized the likelihood function of three-parameter Weibull distribution by a new methodology

that they have proposed. That is a combination between a Variable Neighborhood Search (VNS) and a Simulated Annealing (SA). Like as Nima *et al.* (2012), a new hybrid meta-heuristic have been proposed which compounds of a Variable Neighborhood Search (VNS) together with an Iterated Local Search (ILS) algorithm. This procedure is applied to maximize the likelihood function of Weibull of Burr III distribution which composes of four parameters. In this study, we will use meta-heuristic that developed in the form of the Variable Neighborhood Search (VNS) for parameter estimation because of this process is not complicated and it is not as prevalent in the statistics for used it to estimate a parameter and some references such as Ming-Jun and Liu (2015; Christopher and Schmid, 2009; Yang, 2015) which the works are related to this paper and you can view them a generalization or specialization in some aspect. Moreover, some of them have built the heuristical foundation for MOM.

The remaining of the paper is arranged as follows. First, the background of Variable Neighborhood Search: VNS is briefly introduced then we used a real data set to fit the model and estimate parameters with MOM, MLE and VNS. In the next section, the estimates of parameters are shown and how well of the distribution fits are improved in Checking Model Fit section. Finally, we conclude the present work accordingly some further research for those interest in parameter estimation using the VNS method.

The Background of Variable Neighborhood Search: VNS

Nenad and Hansen (1997) have introduced a novel technique named Variable Neighborhood Search to explore remote neighborhoods of the present responsible solution and move to a novel one in case this new one is an improvement. This technique is based on Local Search method. It is designed to be used in solving the problem in several ways. Whether the problems are on a continuous or discontinuous and can also be used for solving linear program, nonlinear program, integer program, etc.

In programming a VNS algorithm, we will consider the following:

- N_s ($s = 1, \dots, s_{\max}$), a set of all conceivable neighborhood structures and with the $N_s(\theta)$ set of solutions in s^{th} neighborhood of θ . While, Pierre *et al.* (2008) claims that “most local search heuristics use only one neighborhood structure, i.e., $s_{\max} = 1$ ”
- Initial solution θ which may be a random number from various distribution or from others method to generate
- The target function
- The stopping criteria which is up to researcher discretion may be the highest round of iteration between two improvements or the longest CPU time
- A local optimizer which is used in the Local Search step frequently relies on the target function properties. The familiar choices of the local optimizer used for local search are Fletcher-Power (FP), Fletcher-Reeves (FR), Hooke-Jeeves (HJ), Nelder-Mead (NM), Rosenbrock (RO) and Steepest Descent (SD) (Nenad *et al.*, 2008). Here we used the Iterated Local Search algorithm (ILS) (Nima *et al.*, 2012)

Its step is given in VNS Algorithm as follow:

Threshold: Assigning the set of neighborhood structures N_s ; $s = 1, \dots, s_{\max}$; an initial solution θ ; a stopping criteria then,

Stress the following until the stopping criteria is reached: Set $s = 1$;

Iterate until $s = s_{\max}$

a) **Initiation:** Produce a point θ' at arbitrary from the neighborhood s^{th} of $\theta(\theta' \in N_s(\theta))$;

b) **Local Search:** Employ a local search method with θ' as a prior solution; indicate with θ'' the acquired local optimum;

c) **Neighborhood Changes:** If this local optimum is progressive than the present responsible solution, move there ($\theta = \theta''$) such as $f(\theta)$ is the target function to be

magnified if $f(\theta'') > f(\theta)$ set $\theta = \theta''$ and keep on searching with $N_1(s = 1)$; if not, define $s = s + 1$;

VNS is a meta-heuristic method that easy to use and fast to find the most appropriate solution under various problems (Bessadok *et al.*, 2009). Especially of this study, in the case of a parameter estimates cannot be found by statistical methods.

Modeling on a Real Data Set with Expected Distribution

In this study, a real motor claim data set which was collected between 2005 and 2009 from a public non-life insurance company in Thailand is used. A plot of empirical distribution function of the actual claim data set was compared with the cumulative distribution function of the reference skewed right distribution (Philip, 2007; Stuart *et al.*, 2012) because in non-life insurance, claims data are very skewed (Martin, 2012). Eventually, the Lognormal distribution and the Gamma distribution were relatively closed to have the best fit. That can be observed from Fig. 1 and 2.

The probability density function (pdf) of the Lognormal distribution is defined by:

$$f(x, \mu, \sigma^2) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2}; x > 0 \quad (1)$$

where, $-\infty < \mu < \infty$ and $\sigma > 0$ are the location parameter and scale parameter respectively and the pdf of the Gamma distribution is defined by:

$$f(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}; x > 0 \quad (2)$$

where, $\alpha > 0$ and $\beta > 0$ are the shape parameter and scale parameter respectively.

Parameter Estimation

Moment Estimation Method (MOM)

From the probability density function in Equation 1, we can find the parameter estimator from MOM by find k^{th} moment of sample and population and we obtain:

$$\hat{\mu} = -\frac{\ln\left(\sum_{i=1}^n x_i^2\right)}{2} + 2 \ln\left(\sum_{i=1}^n x_i\right) - \frac{3}{2} \ln(n) \quad (3)$$

$$\hat{\sigma}^2 = \ln\left(\sum_{i=1}^n x_i^2\right) - 2 \ln\left(\sum_{i=1}^n x_i\right) + \ln(n) \quad (4)$$

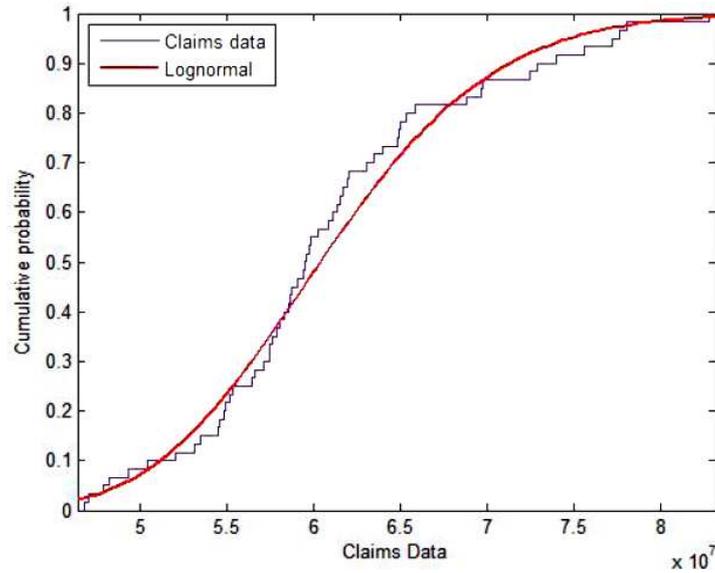


Fig. 1. Empirical CDF of data set against CDF of the Lognormal distribution

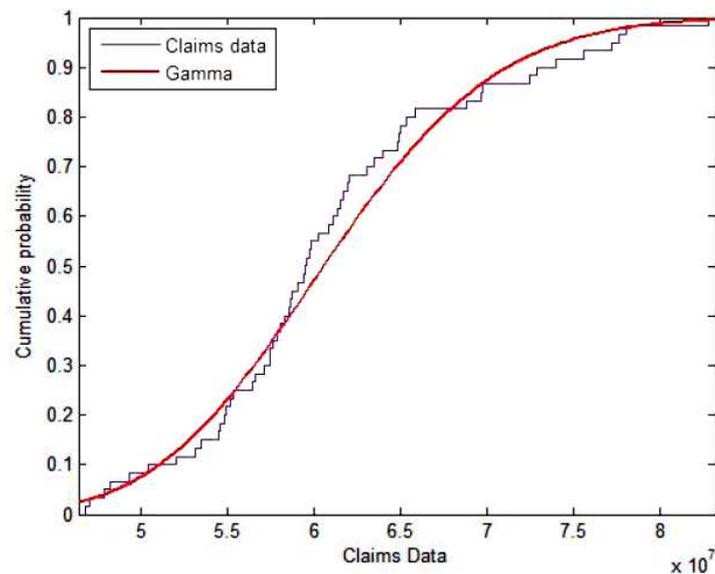


Fig. 2. Empirical CDF of data set against CDF of the Gamma distribution

When x is a random variable from the probability density function as in Equation 2 of the Gamma distribution, we can find the moment estimator as follows:

$$\hat{\beta} = \frac{\sum_{i=1}^n \left(x_i - \left(\frac{1}{n} \sum_{i=1}^n x_i \right) \right)}{\sum_{i=1}^n x_i} \quad (5)$$

$$\hat{\alpha} = \frac{\left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2}{\frac{1}{n} \sum_{i=1}^n \left(x_i - \left(\frac{1}{n} \sum_{i=1}^n x_i \right) \right)^2} \quad (6)$$

Maximum Likelihood Estimation Method (MLE)

Suppose $x = (x_1, x_2, \dots, x_n)$ is a random sample of the Lognormal distribution and the likelihood function is describe as Equation 7:

$$L(x; \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{x_i \sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \left(\frac{\ln x_i - \mu}{\sigma} \right)^2} \quad (7)$$

We can obtain MLE $\hat{\mu}$ and $\hat{\sigma}$ for μ and σ by partial derivative of $\ln L(x; \mu, \sigma^2)$ with respect to μ and σ respectively which set to zero gives. Then, the maximum likelihood estimates for μ and σ are:

$$\hat{\mu} = \frac{\sum_{i=1}^n \ln x_i}{n} \quad (8)$$

and:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (\ln x_i - \hat{\mu})^2}{n}} \quad (9)$$

The Hessian is negative-definite, indicating a strict local maximum.

Similarly, for the Gamma distribution, the likelihood function is obtained as Equation 10:

$$L(x; \alpha, \beta) = \prod_{i=1}^n \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-\frac{x_i}{\beta}} \quad (10)$$

We can get MLE of β and α as $\hat{\beta}$ and $\hat{\alpha}$ by partial derivative of $\ln L(x; \alpha, \beta)$ with respect to β and α respectively which set to zero gives. So, the maximum likelihood estimates of β and α can be acquired as follows:

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \quad (11)$$

$$\ln(\alpha) - \psi(\alpha) - \ln\left(\frac{\bar{x}}{\hat{\alpha}}\right) = 0 \quad (12)$$

where, $\psi(\alpha) = -\frac{d}{d\alpha} \log \Gamma(\alpha)$ and $\bar{x} = \left(\prod_{i=1}^n x_i\right)^{1/n}$.

Since (12) cannot be solved in a closed form then, the value $\hat{\alpha}$ will be obtained by Newton-Raphson iteration method from Equation 13 as follow:

$$\hat{\alpha}_{k+1} - \hat{\alpha}_k = \frac{\ln(\hat{\alpha}_k) - \psi(\hat{\alpha}_k) - \ln(\bar{x} / \hat{\alpha}_k)}{1 / \hat{\alpha}_k - \psi'(\hat{\alpha}_k)} \quad (13)$$

This method will be performed until $|\hat{\alpha}_{k+1} - \hat{\alpha}_k| < \varepsilon$ where ε is some defined error endurance and $\varepsilon > 0$.

Variable Neighborhood Search Method (VNS)

The main reason of using the VNS algorithm to find parameter estimates is that although the maximum likelihood estimators has many desirable properties but in practice the likelihood functions are hard to maximize and there do not have a closed form. So, we will present the VNS algorithm used to find parameter estimator.

In the VNS algorithm planning, we have meditated details as follows:

- Choosing the initial solution θ from estimates of MOM and MLE, then calculating the likelihood function, $L(\theta)$, of this initial, where is a vector of parameters
- In this study, the VNS is performed under a preferred property of the MLE method. This search is expected to maximize the likelihood function of the Lognormal distribution and the Gamma distribution given in Equation 7 and 10 respectively. So, the target function is likelihood function.
- Obtaining the neighborhood of VNS method (N_k) which is created by Equation 14 from Babak *et al.* (2011):

$$\theta' = \theta + udr \quad (14)$$

where u as a random numbers is generated randomly using a uniform distribution in the intervals $[0,1]$, d will be taken whether 1 or -1 which is a run of the new neighborhood and r is the span of neighborhood initiator of the VNS algorithm, here it is defined to be one while the target function improves. In spite of the target function does not enhance, its value is expanded by 0.001. In this study, we have only one the neighborhood structured so, $k_{\max} = 1$

- The stopping criteria is 10,000th iteration or $|\theta'' - \theta| < 0.001$
- Determine a random number v and s which are generated randomly for the ILS algorithm in Local search step to generate the neighboring value
- In the local search step of the VNS algorithm, we choose the Iterated Local Search (ILS) which its algorithm can be split into three stages: The first stage is to create the initial solution which the initial solution will be improved in the second stage and the final as choosing the best answer
- Letting n denotes the number of iteration of searching
- The following steps are reiterated until the stopping condition is reached
 - **Initiation:** Producing parameter θ' as a neighboring value for $\theta' = (\hat{\mu}, \hat{\sigma})$ or $\theta = (\hat{\alpha}, \hat{\beta})$ at random from Equation 13 and compute $L(\theta')$
 - **Local Search:** Use the ILS algorithm
 - Generate v and s using a uniform distribution that take between 0 and 2 and 0 and 1 respectively
 - Generate the neighboring value θ'' for θ' from $\theta'' = \theta' + s$ if $0 < v \leq 1$ or $\theta'' = \theta' - s$ if $1 < v \leq 2$ and compute $L(\theta'')$
 - Evaluate the parameter estimates: If $L(\theta'') > L(\theta')$ then the local best solution is θ'' otherwise, back to 7.2.1. End local search
 - **Neighborhood Changes:** If $L(\theta'') > L(\theta)$, change there ($\theta \leftarrow \theta''$) and keep on searching with $n = n + 1$ and $r = 1$ if not, define $r = r + 0.001$

To apply VNS procedure, we compute the estimated parameter of MOM and MLE from the actual claim data in Table 1.

Then, we obtained estimated value of parameters of the Lognormal distribution by VNS as in Table 2 and for the Gamma distribution by VNS as in Table 3 as follows.

Results

We fitted the Lognormal and the Gamma distribution that we chose as in section of Modeling on a Real Data Set with Expected Distribution to the actual motor claim data set which was mentioned. Moment Estimation Method, Maximum likelihood Estimation Method and Variable Neighborhood Search Method are applied to

estimate parameters. The estimates of parameters are shown in Table 1 to 3.

Checking Model Fit

For showing how well distributions fit our data set, the Kolmogorov-Smirnov test (K-S test) is applied. The K-S test statistic is defined by $KS = \max \{KS^+, KS^-\}$. Where:

$$KS^+ = \max_{1 \leq i \leq n} |S_n(x_i) - F_0(x_i)| = \max_{1 \leq i \leq n} \left| \frac{i}{n} - F_0(x_i) \right|$$

$$KS^- = \max_{1 \leq i \leq n} |F_0(x_i) - S_n(x_{i-1})| = \max_{1 \leq i \leq n} \left| F_0(x_i) - \frac{i-1}{n} \right|$$

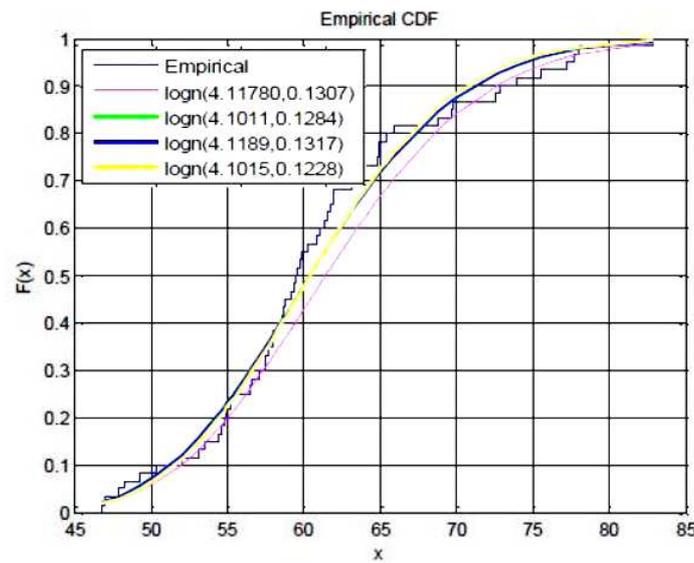


Fig. 3. Empirical CDF against CDF of the Lognormal distribution with parameter estimates from four sources

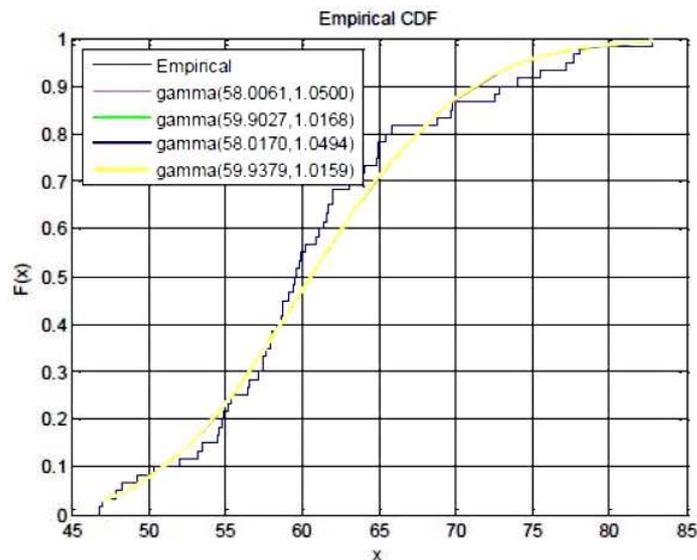


Fig. 4. Empirical CDF against CDF of the Gamma distribution with parameter estimates from four sources

Table 1. The estimated parameters from MOM and MLE

Lognormal	Estimates		Gamma	Estimates	
	$\hat{\mu}$	$\hat{\sigma}$		$\hat{\alpha}$	$\hat{\beta}$
MOM	4.1180	0.1307	MOM	58.0062	1.0501
MLE	4.1010	0.1284	MLE	59.9027	1.0168

Table 2. The estimated value of parameters of the Lognormal distribution from VNS

Lognormal	Estimates		Number of iteration	Run time (s)
	$\hat{\mu}$	$\hat{\sigma}$		
Initial from MOM	4.1189	0.1317	1906	2.2379
Initial from MLE	4.1015	0.1228	3541	4.3121

Table 3. The estimated value of parameters of the Gamma distribution from VNS

Gamma	Estimates		Number of iteration	Run time (s)
	$\hat{\mu}$	$\hat{\sigma}$		
Initial from MOM	58.0170	1.0494	2848	2.6966
Initial from MLE	59.9379	1.0159	622	0.6997

Table 4. Kolmogorov-Smirnov test result

Method	K-S Test							
	Lognormal				Gamma			
	Estimates				Estimates			
	$\hat{\mu}$	$\hat{\sigma}$	KS value	p-value	$\hat{\alpha}$	$\hat{\beta}$	KS value	p-value
MOM	4.1180	0.1307	0.1544	0.1025	58.0062	1.0501	0.1113	0.4164
MLE	4.1010	0.1284	0.1019	0.5311	59.9027	1.0168	0.1107	0.4250
VNS								
Initial from MOM	4.1189	0.1317	0.1403	0.1717	58.0170	1.0494	0.1099	0.4327
Initial from MLE	4.1015	0.1228	0.0996	0.5574	59.9379	1.0159	0.1096	0.4354

The Fig. 3 and 4 show how the estimated fit to the actual data as follows.

Discussion

From the results of this real motor claim data set, they show that the Lognormal model which this parameters were estimated by VNS method and initial estimates from MLE has the smallest D value according to K-S test as in Table 4. So, for this data set we can conclude that this model is the best fit and VNS gives a more accurate estimation of parameter than do MOM and MLE.

In this study, VNS algorithm, a uniform random number u which used in the neighborhood of VNS structure (N_k) is a good choice but, we may choose other distributions to generate a random number u because it may be much better performance. Moreover, as a Local Search from this research has been used the Iterated Local Search in future are may be bring others meta-

heuristics such as Ant Colony Optimization, Simulated Annealing, etc. in to VNS algorithm.

Conclusion

When distributions that we interest are complicate and have many parameters in model, the algorithm of estimation is not a comfortable work. For instant, the likelihood function which is created for the Gamma distribution in the Maximum likelihood Estimation Method is hard to maximize and do not has a closed form. So, the proposed methodology: VNS which is computationally efficient and easily implemented can solve this problem and have the performances in term of both the estimation accuracy and less time to calculate.

Acknowledgement

We are thankful to an informant for their motor claim data set support and for providing necessary guidance concerning projects implementation.

Author's Contributions

Samruam Chongcharoen: Participating in all steps in conducting research and contribution to the writing of the manuscript.

Kunjira Kingphai: Computer programming, analyzing and interpreting the data, reporting and evaluating research

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

- Babak, A., S.T.A. Niakia, M.A. Khalifeb and Y. Faizeb, 2011. A hybrid variable neighborhood search and simulated annealing algorithm to estimate the three parameters of the Weibull distribution. *Expert Syst. Applic.*, 38: 700-708. DOI: 10.1016/j.eswa.2010.07.022
- Bessadok, A., P. Hansen and A. Rebai, 2009. EM algorithm and variable neighborhood search for fitting finite mixture model parameters. *Proceedings of the International Multiconference on Computer Science and Information Technology*, Oct. 12-14, IEEE Xplore Press, pp: 725-733. DOI: 10.1109/IMCSIT.2009.5352758
- Christopher, W.L. and F.A. Schmid, 2009. Claim severities, claim relativities and age: Evidence from SOA group health data. *SSRN Electr. J.* DOI: 10.2139/ssrn.1412243
- Martin, E., 2012. Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models? *Insurance: Math. Econom.*, 51: 239-248. DOI: 10.1016/j.insmatheco.2014.08.004
- Ming-Jun, L. and Y. Liu, 2015. The probabilistic estimates on the largest and smallest q-singular values of random matrices. *Math. Comput.*, 84: 1775-1794. DOI: 10.1090/S0025-5718-2014-02895-0
- Nenad, M. and P. Hansen, 1997. Variable neighborhood search. *Comput. Operat. Res.*, 24: 1098-1100. DOI: 10.1016/S0305-0548(97)00031-2
- Nenad, M., M. Dražićb, V. Kovačević-Vujčićc and M. Čangalović, 2008. General variable neighborhood search for the continuous optimization. *Eur. J. Operat. Res.*, 191: 753-770. DOI:10.1016/j.ejor.2006.12.064
- Nima, Z., B. Abbasi, S.T.A. Niaki and M. Abdi, 2012. Estimating the four parameter of the burr III distribution using a hybrid method of variable neighborhood search and iterated local search algorithms. *Applied Math. Comput.*, 218: 9664-9675. DOI: 10.1016/j.amc.2012.03.003
- Pierre, H., N. Mladenovic and J.A.M. Perez, 2008. Variable neighbourhood search: Methods and applications. *4OR. Q. J. Operat. Res.*, 6: 319-360. DOI: 10.1007/s10288-008-0089-1
- Philip, J.B., 2007. *Statistical and Probabilistic Method in Actuarial Science*. 1st Edn., Interdisciplinary Statistics, Chapman and Hall/CRC, The United State of America. ISBN-10: 1584886951.
- Soontorn, B., A. Sattayatham and P. Sattayatham, 2013. Estimation of weibull parameter using a randomized neighborhood search for the severity of fire accidents. *J. Math. Stat.*, 9: 12-17. DOI: 10.3844/jmssp.2013.12.17
- Stuart, A.K., H.H. Panjer and G.E. Willmot, 2012. *Loss Models from Data to Decisions*. 4th Ed. Wiley Series in Probability and Statistics, John Wiley and Sons, The United State of America. ISBN: 9780470187814.
- Tosaporn, T. and P. Sattayatham, 2012. Fitting of finite mixture distribution to motor insurance claims. *J. Math. Stat.*, 8: 49-56. DOI: 10.3844/jmssp.2012.49.56.
- Yang, L. 2015. The probabilistic estimates of the largest strictly convex p-singular value of pregaussian random matrices. *J. Math. Stat.*, 11: 7-15. DOI: 10.3844/jmssp.2015.7.15