THE IMPORTANCE OF FAT-TAILED AND SKEWED DISTRIBUTIONS IN MODELING VALUE-AT-RISK

KALIN KUYRUKLU VE ÇARPIK DAĞILIMLARIN RİSKE MARUZ DEĞER MODELLEMESİNDE ÖNEMİ

EMRAH ALTUN

PROF. DR. HÜSEYİN TATLIDİL
Supervisor

Submitted to Graduate School of Science and Engineering of Hacettepe University
as a Partial Fulfillment to the Requirements
for the Award of the Degree of Doctor of Philosophy
in Statistics

2018

This work named "**THE IMPORTANCE OF FAT-TAILED AND SKEWED DISTRIBUTIONS IN MODELING VALUE-AT-RISK**" by EMRAH ALTUN has been approved as a thesis for the Degree of **DOCTOR OF PHILOSOHPY IN STATISTICS** by the below mentioned Examining Committee Members.

Prof. Dr. Durdu KARASOY
Head ....

Prof. Dr. Hüseyin TATLIDİL
Supervisor ....

Assoc. Prof. Gamze ÖZEL KADILAR
Member ....

Assoc. Prof. Bülent ALTUNKAYNAK
Member ....

Assoc. Prof. Dr. Rukiye DAĞALP
Member ....

This thesis has been approved as a thesis for the Degree of **DOCTOR OF PHILOSOHPY IN STATISTICS** by the Board of Directors of the Institute for Graduate School of Science and Engineering.

Prof. Dr. Menemşe GÜMÜŞDERELİOĞLU
Director of the Institute of
Graduate School of Science and Engineering

YAYINLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

- ☐ **Tezimin/Raporumun tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.**
  (Bu seçenekle teziniz arama motorlarında indekslenebilecek, daha sonra tezinizin erişim statüsünün değiştirilmesini talep etseniz ve kütüphane bu talebinizi yerine getirse bile, tezinin arama motorlarının önbelleklerinde kalmaya devam edebilecektir.)

- ☒ **Tezimin/Raporumun 06/03/2020 tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.**
  (Bu sürenin sonunda uzatma için başvuruda bulunmadığım taktirde, tezimin/raporumun tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı ve ya tamamının fotokopisi alınabilir)

- ☐ **Tezimin/Raporumun .... tarihine kadar erişime açılmasını istemiyorum, ancak kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisinin alınmasını onaylıyorum.**

- ☐ **Serbest Seçenek/Yazarın Seçimi**

06 / 03 / 2018

Emrah ALTUN

ETHICS

In this thesis study, prepared in accordance with the spelling rules of Institute of Graduate Studies in Science of Hacettepe University,

I declare that

- all the information and documents have been obtained in the base of the academic rules
- all audio-visual and written information and results have been presented according to the rules of scientific ethics
- in case of using others Works, related studies have been cited in accordance with the scientific standards
- all cited studies have been fully referenced
- I did not do any distortion in the data set
- and any part of this thesis has not been presented as another thesis study at this or any other university.

06/03/2018

EMRAH ALTUN

ÖZET

KALIN KUYRUKLU VE ÇARPIK DAĞILIMLARIN RİSKE MARUZ DEĞER MODELLEMESİNDE ÖNEMİ

Emrah ALTUN
Doktora, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Hüseyin TATLIDİL
Mart 2018, 130 sayfa

Finansal risk yönetimi, finansal kuruluşların yatırımlarını garanti altına almaları ve meydana gelebilecek olası kayıpları karşılayabilmeleri açısından oldukça önem arz eden bir konudur. Son yıllarda, beklenmedik ve sıra dışı olaylara bağlı olarak meydana gelen aşırı fiyat dalgalanmaları finansal risk yönetim sistemlerinin gözden geçirilmesi gerçeğini ortaya koymuştur. Finansal anlamda risk, ilgili yatırımın belirli bir zaman periyodundaki maksimum kayıp miktarı olarak tanımlanabilir. Riske Maruz Değer (RMD), yatırımın riskini değerlendirmek ve ölçmek için kullanılan bir risk ölçüsüdür. Birçok RMD modeli, finansal getirilerin normal dağıldığı varsayımı altında önerilmiştir. Finansal getiri serilerinin karakteristik özellikleri incelendiğinde, normal dağılıma göre daha kalın kuyruklu yapıya sahip oldukları ve genellikle sola çarpık oldukları görülmektedir. Diğer önemli bir özellikleri ise normal dağılıma göre basıklık değerleri daha yüksektir. Bu nedenle, normal dağılım varsayımına dayanan RMD modelleri ile elde edilen risk öngörü değerleri, gerçek piyasa riskinin altında kalmaktadır. Birçok çalışmada, kalın kuyruklu yapıyı ve çarpıklığı modelleyebilecek dağılımlardan yararlanılarak RMD modellerinin finansal risk öngörü başarıları arttırılmaya çalışılmıştır.
Tez çalışmasında, Genelleştirilmiş Otoregresif Koşullu Değişen Varyans modelleri önerilen

yeni dağılımlar altında tanımlanmış ve RMD öngörü değerleri elde edilmiştir. Modellerin örneklem dışı öngörü performansları geriye dönük testler ve kayıp fonksiyonları ile değerlendirilmiştir. BİST-100, Nikkei-225, S&P-500 ve Nasdaq-100 indeksleri üzerine yapılan beş farklı uygulama ile modellerinin öngörü performansları karşılaştırılmış ve önerilen modellerin diğer modellere karşı olan üstünlükleri değerlendirilmiştir.

Anahtar Kelimeler: GARCH modelleri; Riske Maruz Değer; Geriye Dönük Test; Kayıp Fonksiyonları; Simülasyon; Finansal Risk.

ABSTRACT

THE IMPORTANCE OF FAT-TAILED AND SKEWED DISTRIBUTIONS IN MODELING VALUE-AT-RISK

Emrah ALTUN
Doctor of Philosophy, Department of Statistics
Supervisor: Prof. Dr. Hüseyin TATLIDİL
March 2018, 130 pages

Most of the Value-at-Risk models assume that financial returns are normally distributed, despite the fact that they are commonly known to be left skewed, fat-tailed and excess kurtosis. Forecasting Value-at-Risk with misspecified model leads to the underestimation or overestimation of the true Value-at-Risk. This study proposes new conditional models to forecast the daily Value-at-Risk by employing the new fat-tailed and skewed distributions to GARCH models. Empirical results show that the fat-tailed and skewed distributions provide superior fit to the conditional distribution of the log-returns among others. Backtesting methodology and loss functions are used to compare the out-of-sample performance of Value-at-Risk models. We conclude that the effects of skewness and fat-tails are more important than only the effect of the fat-tails on accuracy of Value-at-Risk forecasts.

Keywords: GARCH models; Value-at-Risk; Backtesting; Loss functions; Simulation; Financial Risk.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Prof. Dr. Hüseyin Tatlıdil, for his unflagging support, encouragement and guidance throughout my thesis study. He is the best supervisor I have ever seen, and most importantly, he is more than a supervisor. I have to say that he is the my biggest chance in my academic career.

Next, I would like to thank Assoc. Prof. Gamze Özel Kadılar for her valuable suggestions and support throughout the my postgraduate education. I have learnt so much from her.

Also, I would like to thank my parents and my wife, Gökçen Altun, for supporting me in every way.

Finally, I am very grateful to my thesis committee members for their constructive comments/suggestions.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

VaR	Value-at-Risk
FHS	Filtered Historical Simulation
GARCH	Generalized Autoregressive Conditional Heterocedasticity
EVT	Extreme Value Theory
GED	Generalized Error Distribution
SGED	Skewed Generalized Error Distribution
GT	Generalized-T
SGT	Skewed Generalized-T
HT	Heavy-Tailed
HS	Historical Simulation
EWMA	Exponentially Weighted Moving Average
ES	Expected Shortfall
LM	Lagrange Multiplier
JB	Jarque-Berra
ARCH	Autoregressive Conditional Heterocedasticity
EGARCH	Exponential GARCH
apARCH	Asymmetric Power ARCH
IGARCH	Integrated GARCH
IGARCH	Integrated GARCH
SN	Skew-Normal
ST	Skew-T
AST	Asymmetric Student-t
GHST	Generalized Hyperbolic Skewed Student-t
NIG	Normal Inverse Gaussian
ASGT	Alpha Skew Generalized-T
ASGN	Alpha Skew Generalized Normal
SCASGN	Symmetric Component ASGN

GAST	Generalized Alpha Skew-T
MLEs	Maximum Likelihood Estimates
BM	Block-Maxima
GEV	Generalized Extreme Value
POT	Peaks-over-Threshold
GPD	generalized Pareto distribution
QMLE	Quasi Maximum Likelihood Estimation
ME	Mean Excess
LR	Likelihood Ratio
QLF	Quadratic Loss Function
RLF	Regulator's Loss Function
UL	Unexpected Loss
FLF	Firm's Loss Function

1. INTRODUCTION

The Basel Committee on Banking Supervision (BCBS) is a committee of banking supervisory authorities. The main role of BCBS is to provide recommendations on banking regulations about credit, market and operational risks. The BCBS develops the fundamentals of risk management system for financial institutions. The BCBCS enforces the financial institutions to have an enough capital to meet the obligations and absorbs the unexpected and unpredictable losses. It is clear that financial risk management is an essential task for the financial institutions. Financial risk management plays an critical role for financial institutions to guarantee the investments against to unpredictable events. In the last decade, whole world has been faced with the extra ordinary events. Therefore, it is obvious that more reliable financial risk management systems are needed to be developed.

The simple method to measure the risk is the evaluation of losses when the prices of the portfolio assets fall down. The risk for portfolio represents the maximum loss of an investor over a defined time period and specified probability level. The Value-at-Risk (VaR) is a common risk measure to quantify the risk of investment. The VaR is generally used by banks to determine the potential losses in their portfolios. The BCBS makes the usage of internal financial risk models for VaR forecasts possible for financial institutions. Hence, VaR is a essential market risk management tool for financial institutions. Therefore, researchers and academics show great interest to develop new approaches to forecast VaR.

Approaches to VaR can be investigated in three categories. These are parametric approaches, semi-parametric approaches and non-parametric approaches. The Riskmetrics^{TM} model, introduced by Morgan [1], can be viewed as the first parametric approach to VaR. Morgan [1] assumed that financial returns are normally distributed despite the fact they are commonly skewed and far away from the normal distribution. Therefore, Riskmetrics^{TM} model causes underestimated VaR forecasts.

The contributions to parametric VaR models have been done with two different approaches by researchers. First approach is to introduce more accurate volatility model to capture volatility dynamics of financial returns. The second approach is to search more flexible distributions to model the financial returns more accurately. Here, the goal of this study is match up with the second one.

The Filtered Historical Simulation (FHS), introduced by [2] and [3], is the most used non-parametric VaR model. The FHS model can be viewed as a hybrid model combining the historical simulation and generalized autoregressive conditional heterocedasticity (GARCH) models. The most used semi-parametric VaR models consist of conditional and unconditional methods based on the Extreme Value Theory (EVT). These two models are examined comprehensively throughout the study.

In this thesis study, the most important and commonly used VaR models are illustrated. The relative advantages and weaknesses of these models are discussed. Three alternative flexible distributions are proposed and applied to GARCH models. The role of innovation distribution for FHS model is discussed.

The rest of the study is organized as follows: The statistical properties of VaR measure are presented in Section 2. The characteristic properties of financial return series are examined in Section 3. The VaR models are given in Section 4 and this section contains the contributions of the study to financial risk forecasting. The backtesting methodology to compare the performance of VaR models is presented in Section 5. Empirical results are given in Section 6. To demonstrate the usefulness of proposed VaR models against to existing models, five real data applications are provided. Finally, Section 7 contains the conclusion of the study.

1.1. Related Studies

In the last decade, there has been great interest from academics, researchers and financial institutions to develop more extensive financial risk models. The main reason of this interest is based on the financial uncertainty, especially increased in recent years. The more sophisticated risk models are needed to increase the accuracy of financial risk forecasts.

Despite the fact that asset returns have non-normal characteristics, the normal distribution assumption is still widely used in financial risk models. It is widely documented that the normal distribution assumption leads to underestimation or overestimation for true market risk. Therefore, skewed and fat-tailed distributions should be used to increase the accuracy of VaR models. Here, the studies related to VaR are examined in three approaches: Parametric, semi-parametric and non-parametric approaches.

Parametric approaches

The studies on parametric approach are based on the distributional assumption on innovation process of volatility models or related to different volatility models under the same distributional assumption for innovation process. The performances of GARCH type models specified under skewed and fat-tailed distributions were discussed in many studies. The commonly used competitive distributions are Student-t, mixture normal, generalized error distribution (GED), skewed generalized error distribution (SGED), generalized-T (GT) and skewed generalized-T (SGT) distributions.

Angelidis et al. [4] discussed the GARCH models in forecasting daily VaR by means of backtesting results. It is an important research article and cited over 400 times by researchers. Angelidis et al. [4] considered only three distributions for innovation process of GARCH models. They stated the following results: (i) the mean process does not have a significant role for VaR forecast; (ii) The ARCH model yields the satisfactory VaR forecasts when the innovation distribution is normal or generalized error distribution; (iii) the leptokurtic distributions produce more accurate VaR forecasts than normal distributions; (iv) the used sample size does not have an effect on the performances of the models. These results confirmed that the flexible distributions are needed to improve the forecasting ability of GARCH type models. It is obvious that usage of more complex volatility models instead of GARCH model are not necessary.

The GED is a good choice to model excess kurtosis. However, this distribution is not able to model the skewness. Therefore, Lee et al. [5] considered the SGED as an innovation process of GARCH model. Lee et al. [5] compared the performance of

GARCH model specified under SGED with normal distribution and concluded that SGED produces more accurate VaR forecasts than normal distribution for both low and high quantiles. It is an expected result due to the fact that normal distribution assumption is an unrealistic assumption for financial returns.

For the first time, the SGT distribution was applied to GARCH models by Bali and Theodossiou [6]. Bali and Theodossiou [6] concluded that GARCH model with SGT innovation distribution produce more realistic VaR forecasts than GARCH-normal model. The results of this paper are similar to results given in Lee et al. [5]. A similar research was conducted by Lee and Su [7]. Lee and Su [7] concluded that SGT distribution provides better VaR forecasts than normal and Student-t distributions.

The other important studies on parametric VaR models can be summarized as follows: Venkataraman [8] and Zangari [9] used the mixture normal distribution in forecasting VaR. Giot and Laurent [10] examined the performance of skew-T distribution in modeling daily VaR. Giot and Laurent [10] concluded that forecasting VaR with skew-T distribution yields more satisfactory results than other competitive models. Brooks and Persand [11] emphasized that the VaR models which is not able to model asymmetric effects for both volatility process and conditional return distribution produce underestimated VaR forecasts.

Hung et al. [12] compared normal, Student-t and heavy-tailed (HT) distributions, proposed by Politis [13], in terms of VaR accuracy. Hung et al. [12] concluded that GARCH-HT model provides the most accuracate VaR forecasts for energy commodities. Dendramis et al. [14] suggested to use of parametric volatility models with skewed distributions to increase the accuracy of VaR forecasts.

Braione and Scholtes [15] concluded that the skew-T distribution provides better VaR forecasts than normal, Student-t and multivariate exponential power distributions. Recently, Lyu et al. [16] examined the performance of GARCH model under eight innovation distributions in terms of accuracy of VaR forecasts. Lyu et al. [16] concluded that financial institutions take into consideration more flexible distributions in their internal risk management system to increase the accuracy of their internal risk system.

Semi-parametric approaches

One of the most important semi-parametric VaR approach is EVT. EVT is widely used in modeling the tail distribution of financial returns instead of whole modeling of the financial return distribution. EVT has increased its popularity in field of financial risk management after the pioneer work of McNeil and Frey [17]. Note that their study was cited over 1600 times by researchers.

McNeil and Frey [17] proposed two-stage model combining EVT with GARCH volatility model. In literature, this model is also known as GARCH-EVT model. Gilli and Kellezi [18], Chan and Gray [19], Onour [20] and Singh et al. [21] evaluated the performance of GARCH-EVT in modeling daily VaR and compared the GARCH-EVT model with well-known competitive models. Chan and Gray [19] evaluated the performance of conditional EVT approach in forecasting the daily VaR in electricity market. Karmakar [22] investigated the forecasting performance of GARCH- EVT model in Indian stock market. Furió and Climent [23] compared the GARCH-EVT model with GARCH models specified under normal and Student-t distributions. Furió and Climent [23] concluded that the GARCH-EVT produce better VaR forecasts than other models. Harmantzis et al. [24] concluded that EVT model outperforms to non fat-tailed models. One of the important researches of this field was carried out by Altun and Tatlidil [25]. Altun and Tatlidil [25] compared the GARCH-EVT model with GARCH models specified under four innovation distributions: including normal, Student-t, GED and SGED. Altun and Tatlidil [25] emphasized that conditional EVT approach provides the most realistic VaR forecasts among others.

Non-parametric approaches

The Historical Simulation (HS) is the most traditional non-parametric VaR model. The HS model assumes that the price movements repeat itself. Thus, the future distribution of asset could be well-defined by the current returns.

The HS model has several advantages. For instance, HS model could be implemented easily since it is a non-parametric model. Therefore, it does not depend on the distributional assumption. However, the HS model has also several shortcomings. The HS model ignores the time-varying volatility dynamics. In order to remove the lack of HS model, FHS model was introduced ([2, 3]). The FHS model is a mixture of HS and the GARCH models. Barone-Adesi and Giannopoulos [26] demonstrated the usefulness of FHS model agains to the HS model. Kuester et al. [27] compared the performance of various VaR models. Kuester et al. [27] concluded that GARCH-Skew-T, GARCH-EVT approach and FHS model under skewed and fat-tailed distributions produce the satisfactory VaR forecasts. Roy [28] estimated the daily VaR of Indian capital market using the FHS model. Omari [29] compared the FHS, Exponentially Weighted Moving Average (EWMA) and GARCH models in terms of accuracy of VaR forecasts. Omari [29] demonstrated that GJR-GARCH with Student-t innovations and FHS model with GARCH volatility specification performed competitively accurate in estimating VaR forecasts for both standard and more extreme quantiles thereby generally out-performing all the other models under consideration.

1.2. Motivation and Contributions

When examined the recent studies on VaR models, it is clear that there is a great interest to introduce more advanced risk models to increase the accuracy of VaR forecasts. The most important component of parametric VaR models is the true quantile estimation of financial return series. The financial return series have many characteristic properties (see Section 5 for details) such as excess kurtosis and skewness, generally left skewed. Therefore, the developed financial risk models have to take into consideration both excess kurtosis and skewness properties.

The purpose of this thesis study is to introduce new parametric VaR models under the new flexible distributions enable to model both excess kurtosis and skewness observed in financial return series. This study consists of four research articles published or still under consideration in journals. The research articles produced from this study are given below:

1. Altun, E., Tatlidil H., Ozel, G. Conditional ASGT-GARCH approach to Value-

at-Risk, Iranian Journal of Science and Technology, Transactions A: Science, DOI: https://doi.org/10.1007/s40995-018-0484-1.

2. Altun, E., Tatlidil H., Ozel, G., Nadarajah, S. Does the assumption on innovation process play an important role for Filtered Historical Simulation model?, Journal of Risk and Financial Management, DOI: https://doi.org/10.3390/jrfm11010007.

3. Altun, E., Tatlidil H., Ozel, G. Value-at-Risk Estimation with New Skew Extension of Generalized Normal Distribution, (Revision submitted).

4. Altun, E., Tatlidil H., Ozel, G., Nadarajah, S. A New Generalization of Skew-T Distribution with Volatility Models, Journal of Statistical Computation and Simulation, DOI: https://doi.org/10.1080/00949655.2018.1427240.

The contributions of above articles are given as follows: In the first article, GARCH model with Alpha-Skew Generalized-T innovation distribution was introduced. The VaR forecasting performance of GARCH model under Alpha-Skew Generalized-T distribution was compared with GARCH models specified under normal, Student-t, GED, GT and SGT distributions. In the second article, whether the assumption on innovation process plays an important role for FHS model was investigated. In the third article, a new skew extension of generalized normal distribution was proposed with its stochastic volatility model. Firstly, skew extension of generalized normal distribution was proposed. Then, associated GARCH volatility model was introduced by means of proposed distribution. In the fourth article, a new generalization of Skew-T distribution was proposed with its volatility models for both symmetric and asymmetric GARCH models.

1.3. Future Work

The future works of this study can be given as follows: (i) the proposed models will be extended to multivariate case; (ii) the developed computational codes will be published as a R package; (iii) the novel backtesting method, taking into account the cost of excess capital, will be proposed.

2. RISK MEASURES

The VaR is the most popular risk measure to quantify the level of market risk. The VaR measures and quantifies the maximum loss of asset over a defined time horizon under defined confidence level. The VaR is defined as follows

$$VaR_p = F^{-1}(p), \tag{2.1}$$

where F is the cumulative distribution function (cdf). The F^{-1} denotes the inverse of F and p is the quantile value at which VaR is calculated.
The second formula of VaR is given by

$$VaR_t = \hat{\mu}_t + \hat{h}_t F^{-1}(p), \tag{2.2}$$

where $\hat{\mu}_t$ and $\hat{h}_t$ are the forecasts of conditional mean and standard deviation at time t, respectively. Since the mean process is near to zero, it can be excluded. Then, the VaR, with its simple form, can be re-defined as

$$VaR_t = \hat{h}_t F^{-1}(p). \tag{2.3}$$

It can be easily seen from Equation (2.3) that the main components of VaR are the volatility and quantile estimation of financial returns. Therefore, to increase the accuracy of VaR forecasts, the flexible distributions, capture the both skewness and excess kurtosis, have critical role.
The risk measure VaR has following shortcomings: VaR ignores the losses beyond the VaR level. VaR does not hold the one of the properties of coherent risk measure (see, Section 2.1. for details). For this reason, Artzner et al. [30] introduced the alternative measure to VaR, named Expected Shortfall (ES). The ES, an expectation of losses beyond the VaR level, is defined by

$$ES_p = E\left[X_t \middle| X_t \geq VaR_t\left(X\right)\right] = \frac{1}{p}\int_0^p VaR_t\left(X\right)dt. \quad (2.4)$$

where $0 < p < 1$. Figure 1 displays the VaR and ES under the standard normal distribution.

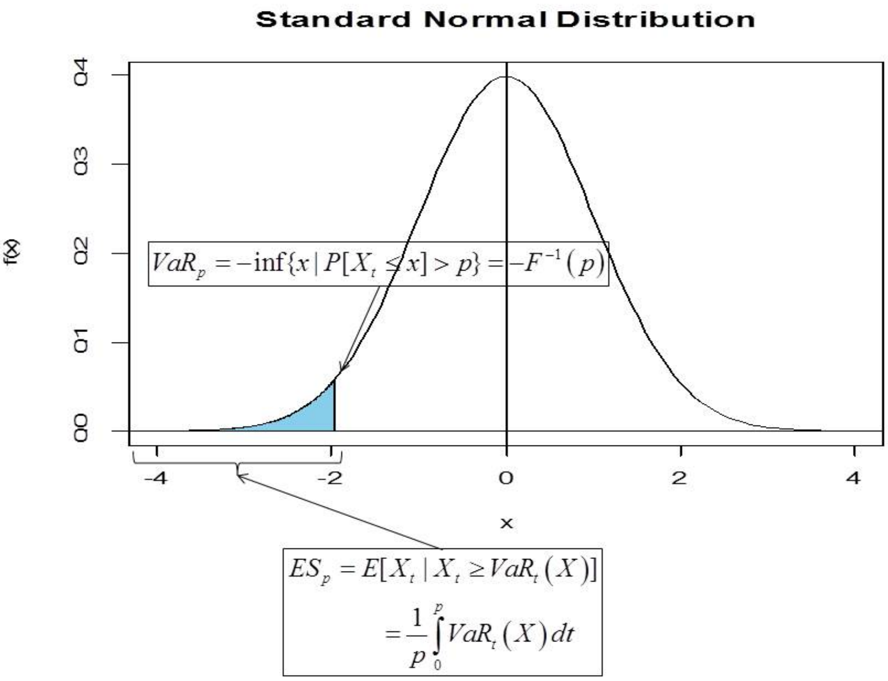


Figure 1: The graphical representation of VaR and ES under standard normal distribution

2.1. Properties of Coherent Risk Measure

Let χ and ω be the set of states and set of all risks, respectively. A risk measure $\gamma\left(Y\right)$ is a mapping function from ω to $\Re$.

The requried axioms of coherent risk measure, introduced by Artzner et al. [30], are given below:

1. *Translation Invariance:* For all $Y \in \omega$ and $k \in \Re$,

$$\gamma\left(Y + k\right) = \gamma\left(Y\right) - k. \quad (2.5)$$

2. *Sub-additivity:* For $Y_1, Y_2 \in \omega$,

$$\gamma (Y_1 + Y_2) \leq \gamma (Y_1) + \gamma (Y_2). \tag{2.6}$$

3. *Positive Homogeneity:* For all $Y \in \omega$ and $\alpha > 0$,

$$\gamma (\alpha \cdot Y) = \alpha \gamma (Y). \tag{2.7}$$

4. *Monotonicity:* For $Y_1, Y_2 \in \omega$ and $Y_1 \leq Y_2$,

$$\gamma (Y_1) \geq \gamma (Y_2). \tag{2.8}$$

The first axiom represents that adding a risk-free amount to an investment, the risk of position decreases by the value of amount. The second axiom reminds the diversification. It means that investor can reduce the risk of investment by composing several assets. The third axiom represents that if the exposure to current position doubles, then the risk of investment increase as double. The fourth axiom represents that if investing Y_2 is more profitable than Y_1, the risk of Y_1 is higher than Y_2.
These four axioms are called as properties of coherent risk measure. These axioms are valid for ES measure; on the other hand, VaR does not hold the second axiom, "sub-additivity" property. Therefore, ES is always coherent risk measure. Hull [31] stated that ES provides the benefits of diversification, but, since the VaR does not hold the second axiom given above, VaR do not provides the benefits of diversification. Yamai and Yoshiba [32] discussed the advantages/disadvantages of the ES over the VaR. The main disadvantages of ES is that when the financial returns are fat-tailed, the forecast error of ES is larger than that of VaR. This means that excess capital value of ES is higher than that of VaR. Thus, if the commercial banks determine their positions with ES measure, they have to hold more capital to meet the unexcepted losses. The other disadvantages of ES is backtesting problem. It should be noted that Basel Committee advises to institutions to replace the 99% VaR with a 97.5% ES. It verifies the results

of Yamai and Yoshiba [32].

3. STYLIZED FACTS OF FINANCIAL RETURN SERIES

The one-period return for an asset is given by

$$R_t = \left(\frac{S_t - S_{t-1}}{S_{t-1}} \right), \tag{3.1}$$

where S_t is the closing price at time t. Then, the simple log-return is given by

$$R_t = \log \left(\frac{S_t}{S_{t-1}} \right). \tag{3.2}$$

The financial returns have some characteristic properties. It is essential knowledge to understand and model the financial returns. Cont [33] provides a comprehensive study for stylized facts of financial returns. Based on the results given in Cont [33], the stylized facts of financial return series can be summarized as below:

3.1. Autocorrelation

The financial returns generally do not display significant serial autocorrelation, especially for high frequency. Ljung-Box Q test can be used for checking the autocorrelation problem. The null hypothesis is $H_0 : \rho_1 = \rho_2 = ... = \rho_p = 0$. The correlation of order j is given by

$$\hat{\rho}_i = \frac{\sum\limits_{t=i+1}^{T} (r_t - \bar{r})(r_{t-i} - \bar{r})}{\sum\limits_{t=i+1}^{T} (r_t - \bar{r})^2}, 0 \le i \le T-1, \tag{3.3}$$

Then, the test statistic of Ljung-Box Q is

$$Q_p = T(T+2) \sum_{i=1}^{p} \frac{\hat{\rho}_i}{T-i}. \tag{3.4}$$

where Q_p is asymptotically distributed as χ_p^2.

3.2. Heteroskedasticity

Heteroskedasticity means that the variance of error term is not constant and it varies across to time. Lagrange Multiplier (LM) test, introduced by Engle [34] and known as ARCH-LM, can be used to test heteroskedasticity problem. The steps of ARCH-LM test is given below.

1. The model is fitted to used data set and extract the residuals,

2. Run the below regression model using the extracted residuals in first step,

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 ..... + \alpha_p u_{t-p}^2 + v_t. \tag{3.5}$$

   where u is the residual extracted from first step and obtain R^2 value of above regression model,

3. Compute $T \times R^2$. $T \times R^2$ is distributed as χ_p^2. The null hypothesis, $H_0 : \alpha_0 = \alpha_1 = ... = \alpha_p$ is that there is no ARCH affect. Note that T represents the number of observations.

3.3. Heavy Tails or Fat Tails

The one of the most important properties of financial returns is fat-tail phenomena. The financial return series exhibits fat-tailed structure. It means that the frequency of extreme losses or gains is bigger than the represented by normal distribution. The unconditional distribution of financial returns displays a power-law or Pareto type tails. Therefore, it excludes the stable laws with infinite variance and the normal distribution.

3.4. Asymmetry or Skewness

The unconditional distribution of financial returns are generally skewed to left. It means that the unpredictable extreme losses have more frequency than gains. Jarque-Berra test can be used to test whether the used time series exhibits excess kurtosis and skewness. Jarque-Berra (JB) statistic is given by

$$JB = T\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right]. \tag{3.6}$$

where S and K are skewness and kurtosis, respectively. JB statistic is asymptotically distributed as χ^2_2.

3.5. Volatility Clustering

It is well-known fact that the large losses or gains tend to be followed by another large losses or gains. The high volatility events are generally clustered in time. This situation is called as volatility clustering.

3.6. Conditional Heavy Tails

The volatility clustering can be solved by modeling the financial returns with GARCH type models. On the other hand, even after modeling the returns with GARCH models, the residuals may have still fat tails.

The stylized facts of financial returns summarize the main motivation of this study. Since financial returns exhibit excess kurtosis and skewness (generally skewed to left), new flexible distributions to model the stylized facts of financial returns with high accuracy are still needed.

4. VALUE-AT-RISK MODELS

The approaches to VaR could be investigated in three categories: (i) fully parametric models approaches based on the stochastic volatility models; (ii) non-parametric approaches based on the HS methods and (iii) EVT approach based on the modeling of the tails of return distribution. These models are investigated in following subsections.

4.1. Parametric Value-at-Risk Models

Volatility modeling is an important topic for econometrics. For the first time, Engle [34] introduced the ARCH model. Then, Bollerslev [35] proposed the generalization of ARCH model, called as GARCH model. ARCH and GARCH models do not enable to capture the asymmetric volatility dynamics in financial return series. For this reason, asymmetric GARCH models were introduced. The most popular asymmetric GARCH model, introduced by [36], is exponential GARCH (EGARCH) model. EGARCH model provides an opportunity to take into account the effects of bad and good news on volatility forecasting. Glosten et. al [37] introduced the GJR-GARCH model. GJR-GARCH model has simplier mathematical form than EGARCH model. Asymmetric power ARCH (apARCH) model, introduced by [38], can be viewed as third popular asymmetric GARCH model. The apARCH model contains the GARCH and GJR-GARCH models as its submodel. This property of apARCH model increases its popularity. These are the most used asymmetric GARCH models. In this study, these models are examined theoretically. According to previous studies given in Section 2, it is clear that the asymmetric volatility models do not have critical impact on accuracy of VaR forecasts and these models do not increase the forecasting accuracy of VaR models. Here, the importance of distributional assumption on innovation process in accuracy of VaR forecasts is investigated as main purpose of the study.
The definition of standard GARCH, GJR-GARCH, EGARCH and apARCH models are given below:

GARCH model

Let r_t denotes the daily log-returns. The benchmark model, GARCH(1,1), is defined as follows:

$$\begin{aligned} &r_t = m_t + e_t, \\ &e_t = \varepsilon_t\, h_t,\ \ \varepsilon_t \sim i.i.d. \\ &h_t^2 = \omega + \gamma_1\, e_{t-1}^2 + \gamma_2 h_{t-1}^2, \end{aligned} \tag{4.1}$$

where m_t and h_t^2 are the conditional mean and variance, respectively. The ε_t is the innovation distribution with zero mean and unit variance. To ensure the positive variance, $\omega > 0, \gamma_1 > 0$ and $\gamma_2 > 0$ and for covariance stationarity $\gamma_1 + \gamma_2 < 1$. When the parameter $\gamma_2 = 0$, the GARCH model reduces to model of Engle [34]. Note that when the $\gamma_1 + \gamma_2 = 1$, this model is called as Integrated GARCH (IGARCH) [39].
The unconditional variance of e_t is given by

$$\begin{aligned} &Var\left(e_t\right) = E\left(e_t^2\right) - \left\{E\left(e_t\right)\right\}^2 \\ &= E\left(e_t^2\right) \\ &= E\left(h_t^2 \varepsilon_t^2\right) \\ &= E\left(h_t^2\right) \\ &= \omega + \gamma_1 E\left(e_{t-1}^2\right) + \gamma_2 h_{t-1}^2 \\ &= \omega + \left(\gamma_1 + \gamma_2\right) h_{t-1}^2, \end{aligned} \tag{4.2}$$

The e_t is stationary process. Therefore, it is easy to see that the unconditional variance of e_t can be given as follows:

$$Var\left(e_t\right) = \frac{\omega}{1 - \gamma_1 - \gamma_2}. \tag{4.3}$$

GJR-GARCH model

The GJR-GARCH(1,1) model is given by

$$h_t^2 = \omega + \gamma_1 e_{t-1}^2 + \gamma_3 I_{t-1} e_{t-1}^2 + \gamma_2 h_{t-1}^2, \tag{4.4}$$

where $\omega > 0, \gamma_1 > 0, \gamma_2 > 0$ and $\gamma_1 + \gamma_3 > 0$. When the innovation distribution is normal, $\gamma_1 + \gamma_2 + \frac{1}{2}\gamma_3 < 1$ for covariance stationary. The parameter γ_3 represents the leverage effect. I_{t-1} is an indicator function and $I_{t-1} = 1$ for $e_{t-1} < 0$, otherwise, $I_{t-1} = 0$. The parameter γ_3 represents the asymmetry effect on volatility. The positive γ_3 parameter indicates that the bad news yields higher volatility than good news. Note that when $\gamma_3 = 0$, GJR-GARCH model reduces to model of Bollerslev [35].
The unconditional variance of e_t for GJR-GARCH models is given by

$$Var\left(e_t\right) = \frac{\omega}{1 - \gamma_1 - \gamma_2 - \kappa\gamma_3}, \tag{4.5}$$

where κ is

$$\kappa = E\left(I_{t-1} e_{t-1}^2\right) \int\limits_{-\infty}^{0} f\left(\varepsilon\right) d\varepsilon. \tag{4.6}$$

It is easy to see that $\kappa = 1/2$ for standard normal distribution.

EGARCH model

The other commonly used asymmetric volatility model, EGARCH(1,1), is given by

$$\ln\left(h_t^2\right) = \omega + \gamma_1 \left(\frac{e_{t-1}}{h_{t-1}}\right) + \gamma_2 \left[\left|\frac{e_{t-1}}{h_{t-1}}\right| - E\left(\left|\frac{e_{t-1}}{h_{t-1}}\right|\right)\right] + \gamma_3 \ln\left(h_{t-1}^2\right), \tag{4.7}$$

where γ_1 represents the sign effect and γ_2 represents the size effect of asymmetric volatility. When the parameter $\gamma_1 < 0$, the bad news has more effect than good news on volatility. In EGARCH model, logarithmic transformation guarantees the non-negative variance. Therefore, no need to be made additional restrictions on model parameters. The expected value given in above equation is given by

$$E\left(\left|\frac{e_{t-1}}{h_{t-1}}\right|\right) = E\left(|\varepsilon_t|\right) = \int\limits_{-\infty}^{\infty} |\varepsilon|\, f\left(\varepsilon\right) d\varepsilon, \tag{4.8}$$

where $f(\varepsilon)$ is the standardized innovation distribution. The unconditional variance of e_t for EGARCH model is given by

$$Var\left(e_t\right) = \frac{\omega}{1-\gamma_3}. \tag{4.9}$$

apARCH model

The apARCH model is given by

$$h_t^\delta = \omega + \gamma_1 (|e_{t-1}| - \gamma_2 e_{t-1})^\delta + \gamma_3 h_{t-1}^\delta, \tag{4.10}$$

where $\delta > 0, \omega > 0, \gamma_1 > 0, \gamma_3 > 0$ and $-1 < \gamma_2 < 1$. The parameter γ_2 is the leverage parameter. The apARCH model of Ding et al. [38] contains the some important models as its sub-models:

- ✓ When the parameter $\gamma_2 = 0$ and $\delta = 2$, apARCH model reduces to standard GARCH.
- ✓ When the parameter $\delta = 2$, apARCH model reduces to GJR-GARCH.

The unconditional variance of e_t for apARCH models is given by

$$Var\left(e_t\right) = \left(\frac{\omega}{1-\gamma_3-\gamma_1\kappa}\right)^{2/\delta}, \tag{4.11}$$

where κ is given by

$$\kappa = E(|\varepsilon| - \gamma_2\varepsilon)^\delta = \int\limits_{-\infty}^{\infty} \left(|\varepsilon| - \gamma_2\varepsilon\right)^\delta f\left(\varepsilon\right) d\varepsilon. \tag{4.12}$$

The distributional assumption on innovation process of volatility models has critical effect on both accuracy of volatility and VaR forecasts. The rest of this section is devoted to present GARCH models with well-known and newly defined flexible distributions.

4.1.1. Normal Distribution

The log-likelihood (ll) function of the GARCH-N model is given by

$$\ell(\boldsymbol{\psi}) = -0.5\left(T\ln 2\pi + \sum_{t=1}^{T}\ln h_t^2 + \sum_{t=1}^{T}\varepsilon_t^2\right). \tag{4.13}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2)$ denotes the parameter vector of the GARCH-N model. The one-day-ahead VaR forecast of GARCH-N model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_p^{-1}(\varepsilon_t)\hat{h}_{t+1}. \tag{4.14}$$

where $F_p{}^{-1}$ is the quantile function (qf) of the normal distribution at the p level. The **qnorm** function of R is used to obtain quantile estimation of standard normal distribution.

4.1.2. Riskmetrics$^{\text{TM}}$ Model

EWMA model is an alternative volatility model and widely used by Riskmetrics$^{\text{TM}}$ company. When the parameters of GARCH model $\gamma_1 + \gamma_2 = 1$, then GARCH model is called as IGARCH model. The idea behind the EWMA model is based on the IGARCH model. EWMA model is widey used by financial institutions because of its simple form and software support. EWMA model, introduced by Morgan [1], is given by,

$$h_t^2 = \lambda h_{t-1}^2 + (1-\lambda)\ r_{t-1}^2, \tag{4.15}$$

where λ is exponential factor. Riskmetrics$^{\text{TM}}$ uses $\lambda = 0.94$ and $\lambda = 0.97$ for daily and monthly returns, respectively, and uses last 75 data points for estimation of the model parameters. The one-day-ahead VaR forecast of Riskmetrics$^{\text{TM}}$ model is given by

$$VaR_{t+1} = F_p^{-1}(\varepsilon_t)\hat{h}_{t+1}. \tag{4.16}$$

where $F_p{}^{-1}$ is qf of the normal distribution at the p level.

4.1.3. Skew-Normal Distribution

The probability density function (pdf) of the skew-normal (SN) distribution, proposed by Azzalini [40], is given by

$$f(x;\lambda) = 2\phi(x)\,\Phi(x\lambda)\,, \mathrm{x} \in \Re, \lambda \in \Re, \tag{4.17}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of standard normal distribution, respectively. The additional parameter λ controls the skewness. Note that the SN distribution is left-skewed for $\lambda < 0$, otherwise, it is right skewed. When $\lambda = 0$, the SN distribution reduces to standard normal distribution. The raw moments of SN distribution are given by

$$E\left(X^{2k+1}\right) = \sqrt{\frac{2}{\pi}}\frac{(2k+1)!}{2^k k!}\sum_{i=0}^{k}(-1)^i\left(\begin{array}{c} k \\ i \end{array}\right)\frac{\delta^{2i+1}}{2i+1}, \tag{4.18}$$

where $k = 0, 1, 2, \ldots, n$ and $\delta = \lambda\big/\sqrt{1+\lambda^2}$. The even moments of SN distribution are equal to standard normal distribution.

The mean and variance of SN distribution is given by

$$E(X) = \sqrt{\frac{2}{\pi}}\delta, \qquad Var(X) = 1 - \frac{2}{\pi}\delta^2. \tag{4.19}$$

Let $\varepsilon = (X-\mu)/\sigma$, the random variable X can be expressed as $X = \varepsilon\sigma + \mu$. Thus, the pdf of standardized SN distribution is given by

$$f(\varepsilon;\lambda) = 2\sigma\phi(\varepsilon\sigma+\mu)\,\Phi([\varepsilon\sigma+\mu]\,\lambda)\,, \tag{4.20}$$

where μ and σ are given in (4.19), respectively. Using the standardized SN distribution, the ll function of GARCH-SN model is given by

$$\ell(\boldsymbol{\psi}) = T\ln(2\sigma) + \sum_{t=1}^{T}\ln(\phi(\varepsilon_t\sigma+\mu)) + \sum_{t=1}^{T}\ln(\Phi([\varepsilon_t\sigma+\mu]\,\lambda)) - \frac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right). \tag{4.21}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \lambda)$ is the parameter vector of GARCH-SN model. The one-day-ahead VaR forecast of GARCH-SN model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_p^{-1}(\varepsilon_t, \lambda)\hat{h}_{t+1}. \tag{4.22}$$

where $F_p^{-1}(\varepsilon_t, \lambda)$ is the qf of SN distribution at p level. The **qsn** function of R is used to obtain quantile estimation of standardized SN distribution.

4.1.4. Student-t Distribution

To relax the assumption on innovations, Bollerslev [35, 41] proposed the GARCH model with the Student-t innovations. According to Bollerslev [35], GARCH model with the Student-t distribution yields more satisfactory results than the normal distribution. The ll function of the GARCH-Student-t (GARCH-T) model is given by

$$\begin{array}{l} \ell(\boldsymbol{\psi}) = T\left[\ln\Gamma\left(\frac{\upsilon+1}{2}\right) - \ln\Gamma\left(\frac{\upsilon}{2}\right) - \frac{1}{2}\ln\left[\pi(\upsilon-2)\right]\right] \\ -\frac{1}{2}\sum\limits_{t=1}^{T}\left[\ln h_t^2 + (1+\upsilon)\ln\left(1+\frac{\varepsilon_t^2}{\upsilon-2}\right)\right] \end{array}. \tag{4.23}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \upsilon)$ denotes the parameter vector. The $\Gamma(\upsilon)$ is the gamma function. The parameter υ is the tail-thickness parameter.
The one-day-ahead VaR forecast of GARCH-T model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_p^{-1}(\varepsilon_t, \upsilon)\hat{h}_{t+1}. \tag{4.24}$$

where $F_p^{-1}(\varepsilon_t, \upsilon)$ is the qf of Student-t distribution at the p level. The **qt** function of R is used to obtain quantile estimation of standardized Student-t distribution.

4.1.5. Skew-T Distribution

The pdf of Skew-T (ST) distribution, introduced by [42], is given by

$$f(x; \lambda, \upsilon) = 2t(x; \upsilon)\, T\left(\sqrt{\frac{1+\upsilon}{x^2+\upsilon}}\lambda x; \upsilon+1\right), x \in \Re, \tag{4.25}$$

where $t(\cdot)$ and $T(\cdot)$ are the pdf and the cdf of Student-t distribution, respectively. Here, parameter λ controls the skewness of the distribution and υ is the tail-thickness parameter. When $\lambda < 0$, ST is skewed to left, otherwise, it is righted skewed. When the parameter $\lambda = 0$, ST distribution reduces to Student-t distribution. The moments of ST distribution are given by

$$E\left(X^k\right) = \frac{\left(\frac{\upsilon}{2}\right)^{\frac{k}{2}}\Gamma\left(\frac{(\upsilon-k)}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)} E\left(Z^k\right), \tag{4.26}$$

where $E\left(Z^k\right)$ is the kth moment of SN distribution given in (4.18).
The mean and variance of ST distribution are, respectively, given by

$$\begin{aligned} E(X) &= \frac{\left(\frac{\upsilon}{\pi}\right)^{\frac{1}{2}}\Gamma\left(\frac{(\upsilon-1)}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)} \frac{\lambda}{\sqrt{1+\lambda^2}}, \\ Var(X) &= \left(\frac{\upsilon}{\upsilon-2} - \mu^2\right). \end{aligned} \tag{4.27}$$

Let $\varepsilon = (X-\mu)/\sigma$, the random variable X can be expressed as $X = \varepsilon\sigma + \mu$. Thus, the pdf of standardized ST distribution is given by

$$f(\varepsilon;\lambda,\upsilon) = 2\sigma t\left((\varepsilon\sigma+\mu);\upsilon\right) T\left(\sqrt{\frac{1+\upsilon}{(\varepsilon\sigma+\mu)^2+\upsilon}}\lambda(\varepsilon\sigma+\mu);\upsilon+1\right), \upsilon > 2 \tag{4.28}$$

where μ and σ are given in (4.27), respectively. The ll function of GARCH model with the ST innovation distribution is given by

$$\begin{aligned} \ell(\boldsymbol{\psi}) &= T\ln(2) + T\ln(\sigma) + \sum_{t=1}^{T}\ln\left[t\left((\varepsilon_t\sigma+\mu);\upsilon\right)\right] \\ &+ \sum_{t=1}^{T}\ln\left[T\left(\sqrt{\frac{1+\upsilon}{(\varepsilon_t\sigma+\mu)^2+\upsilon}}\lambda(\varepsilon_t\sigma+\mu);\upsilon+1\right)\right] - \frac{1}{2}\sum_{t=1}^{T}\ln(h_t^2). \end{aligned} \tag{4.29}$$

where $\boldsymbol{\psi} = (m,\omega,\gamma_1,\gamma_2,\lambda,\upsilon)$ is the parameter vector. The one-day-ahead VaR forecast of GARCH-ST model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_p^{-1}(\varepsilon_t;\lambda,\upsilon)\hat{h}_{t+1}. \tag{4.30}$$

where $F_p^{-1}(\varepsilon_t;\lambda,\upsilon)$ is the qf of ST distribution at the p level. The **qst** function of R is used to obtain quantile estimation of standard SN distribution.

4.1.6. Asymmetric Student-t Distribution

Zhu and Galbraith [43] proposed the asymmetric Student-t (AST) distribution. The pdf of AST distribution is given by

$$f(x;\alpha,\upsilon_1,\upsilon_2)=\begin{cases}\frac{\alpha}{\alpha^*}K(\upsilon_1)\left[1+\frac{1}{\upsilon_1}\left(\frac{x}{2\alpha^*}\right)^2\right]^{-\frac{\upsilon_1+1}{2}}, x\leq 0\\ \frac{1-\alpha}{1-\alpha^*}K(\upsilon_2)\left[1+\frac{1}{\upsilon_2}\left(\frac{x}{2(1-\alpha^*)}\right)^2\right]^{-\frac{\upsilon_2+1}{2}}, x>0\end{cases} \tag{4.31}$$

where the parameter $\alpha\in(0,1)$ controls the skewness of distribution, $\upsilon_1>0$ and $\upsilon_2>0$ are the tail parameters for left and right tails, respectively, and $K(\upsilon)\equiv\Gamma((\upsilon+1)/2)/[\sqrt{\pi\upsilon}\Gamma(\upsilon/2)]$, $\alpha^*=\alpha K(\upsilon_1)/[\alpha K(\upsilon_1)+(1-\alpha)K(\upsilon_2)]$. The AST distribution is an alternative generalization of Student-t distribution with two tail parameters and one skewness parameter. The tail parameters, υ_1 and υ_2, provide more flexibility in tails of the distribution.

Let the random variable X follows AST distribution, the cdf and qf of AST distribution are, respectively, given by,

$$F(x)=2\alpha F_t\left(\frac{x\wedge 0}{2\alpha^*};\upsilon_1\right)+2(1-\alpha)\left[F_t\left(\frac{x\vee 0}{2(1-\alpha^*)};\upsilon_2\right)-\frac{1}{2}\right]. \tag{4.32}$$

$$F^{-1}(p)=2\alpha^*F_t^{-1}\left(\frac{p\wedge\alpha}{2\alpha};\upsilon_1\right)+2(1-\alpha)F_t^{-1}\left(\frac{p\vee\alpha+1-2\alpha}{2(1-\alpha)};\upsilon_2\right). \tag{4.33}$$

where $p\wedge\alpha=\min(p,a)$, $p\vee\alpha=\max(p,a)$ and $F_t(\cdot;\upsilon)$, $F_t^{-1}(\cdot;\upsilon)$ are cdf and inverse cdf functions of Student-t distribution with parameter υ.

The mean and variance of AST distribution are, respectively, given by,

$$E(X)=4\left[-\alpha\alpha^*\frac{\upsilon_1K(\upsilon_1)}{\upsilon_1-1}+(1-\alpha)(1-\alpha^*)\frac{\upsilon_2K(\upsilon_2)}{\upsilon_2-1}\right]. \tag{4.34}$$

$$
\begin{aligned}
&Var\left(X\right) = 4\left[\alpha(\alpha^*)^2\tfrac{\upsilon_1}{\upsilon_1-2} + \left(1-\alpha\right)\left(1-\alpha^*\right)^2\tfrac{\upsilon_2}{\upsilon_2-2}\right] \\
&-16B^2\left[-(\alpha^*)^2\tfrac{\upsilon_1}{\upsilon_1-1} + \left(1-\alpha^*\right)^2\tfrac{\upsilon_2}{\upsilon_2-1}\right]^2.
\end{aligned} \tag{4.35}
$$

where $B = \alpha K\left(\upsilon_1\right) + \left(1-\alpha\right)K\left(\upsilon_2\right)$ and $K\left(\cdot\right)$ is defined in (4.31).

Using the standardized AST distribution, the ll function of GARCH model with the AST innovation distribution is given by

$$
\begin{aligned}
&\ell\left(\boldsymbol{\psi}\right) = \sum_{t=1}^{T}\ln\left(\tfrac{\sigma\alpha}{\alpha^*}\right) - \tfrac{\upsilon_1+1}{2}\sum_{t=1}^{T}\ln\left\{K\left(\upsilon_1\right)\left[1+\tfrac{1}{\upsilon_1}\left(\tfrac{\varepsilon_t\sigma+\mu}{2\alpha^*}\right)^2\right]\right\}_{y\leq\mu/\sigma} \\
&\sum_{t=1}^{T}\ln\left(\tfrac{\sigma(1-\alpha)}{1-\alpha^*}\right) - \tfrac{\upsilon_2+1}{2}\sum_{t=1}^{T}\ln\left\{K\left(\upsilon_2\right)\left[1+\tfrac{1}{\upsilon_2}\left(\tfrac{\varepsilon_t\sigma+\mu}{2(1-\alpha^*)}\right)^2\right]\right\}_{y>\mu/\sigma} - \tfrac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right).
\end{aligned} \tag{4.36}
$$

where μ and σ are given in (4.34) and (4.35), respectively. Here, $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \alpha, \upsilon_1, \upsilon_2)$ is the parameter vector.

The one-day-ahead VaR forecast of GARCH-AST model is given by

$$
VaR_{t+1} = \hat{m}_{t+1} + {F_p}^{-1}(\varepsilon_t, \alpha, \upsilon_1, \upsilon_2)\hat{h}_{t+1}. \tag{4.37}
$$

where ${F_p}^{-1}(\varepsilon_t, \alpha, \upsilon_1, \upsilon_2)$ is the qf of AST distribution at p level. The quantile estimation of AST distribution is obtained by using Equation (4.33).

4.1.7. Generalized Hyperbolic Skewed Student-t Distribution

The generalized hyperbolic skewed Student-t (GHST) is a limiting distribution of generalized hyberbolic (GH) distribution. The comprehensive properties of GHST distribution with financial applications were given by Aas and Haff [44]. Firstly, we define the pdf of GH distribution as

$$
f\left(x;\alpha,\beta,\lambda,\mu,\delta\right) = \frac{\left(\alpha^2-\beta^2\right)K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+\left(x-\mu\right)^2}\right)\exp\left(\beta\left(x-\mu\right)\right)}{\sqrt{2\pi}\alpha^{\lambda-1/2}\delta^{\lambda}K_{\lambda}\left(\delta\sqrt{\alpha^2-\beta^2}\right)\left(\sqrt{\delta^2}+\left(x-\mu\right)^2\right)^{1/2-\lambda}}, \tag{4.38}
$$

where $x \in \Re$ and K_j is the modified bessel function of the third kind of order j (see Abramowitz and Stegun [45] for details). The parameters of the GH distribution hold the following conditions:

$$\begin{aligned} &\delta \geq 0, |\beta| < \alpha \quad \text{if} \lambda > 0 \\ &\delta > 0, |\beta| < \alpha \quad \text{if} \lambda = 0 \\ &\delta > 0, |\beta| \leq \alpha \quad \text{if} \lambda < 0 \end{aligned} \tag{4.39}$$

Let $\lambda = -\upsilon/2$ and $\alpha \to |\beta|$ in (4.38), the pdf of GHST distribution is given by

$$f\left(x;\beta,\upsilon,\mu,\delta\right) = \begin{cases} \frac{2^{\frac{1-\upsilon}{2}}\delta^{\upsilon}|\beta|^{\frac{\upsilon+1}{2}}K_{(\upsilon+1)/2}\left(\sqrt{\beta^2\left(\delta^2+(x-\mu)^2\right)}\right)\exp(\beta(x-\mu))}{\Gamma\left(\frac{\upsilon}{2}\right)\sqrt{\pi}\left(\sqrt{\delta^2+(x-\mu)^2}\right)^{\frac{\upsilon+1}{2}}}, \beta \neq 0 \\ \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{\pi}\delta\Gamma\left(\frac{\upsilon}{2}\right)}\left[1+\frac{(x-\mu)^2}{\delta^2}\right]^{-\frac{(\upsilon+1)}{2}}, \beta = 0 \end{cases} \tag{4.40}$$

where $\mu \in \Re$ and $\delta > 0$ are location and scale parameters, respecetively; $\beta \in \Re$ and $\upsilon > 0$ are shape parameters. Note that when $\beta = 0$, GHST distribution is equal to non-central Student-t distribution with mean μ and variance $\delta^2/(\upsilon-2)$. The mean and variance of GHST distribution are given by

$$\begin{aligned} &E\left(X\right) = \mu + \frac{\beta\delta^2}{\upsilon-1} \\ &Var(X) = \frac{2\beta^2\delta^4}{(\upsilon-2)(\upsilon-4)} + \frac{\delta^2}{\upsilon-2}. \end{aligned} \tag{4.41}$$

As seen in (4.41), the variance of GHST distribution is valid for $\upsilon > 4$. Note that skewness and kurtosis of GHST distribution are only valid for $\upsilon > 6$ and $\upsilon > 8$, respectively.

The standardized GHST distribution can be obtained by following re-parametrization,

$$\begin{aligned} &\mu = -\frac{\delta\beta}{\upsilon-2} \\ &\delta^2 = \frac{(\upsilon-2)(\upsilon-4)}{4\beta^2}\left(-1+\sqrt{1+\frac{8\beta^2}{\upsilon-4}}\right) \end{aligned} \tag{4.42}$$

Using the standardized GHST distribution, the ll function of GARCH-GHST model is given by (for $\beta \neq 0$),

$$
\begin{aligned}
\ell(\boldsymbol{\psi}) = {} & T\ln\left(2^{\frac{1-\upsilon}{2}}\left(\frac{(\upsilon-2)(\upsilon-4)}{4\beta^2}\left(-1+\sqrt{1+\frac{8\beta^2}{\upsilon-4}}\right)\right)^{\upsilon/2}|\beta|^{\frac{\upsilon+1}{2}}\frac{1}{\Gamma(\upsilon/2)\pi}\right) \\
& + \sum_{t=1}^{T}\ln\left(K_{(\upsilon+1)/2}\left(\sqrt{\beta^2\left(\begin{array}{c}\left(\frac{(\upsilon-2)(\upsilon-4)}{4\beta^2}\left(-1+\sqrt{1+\frac{8\beta^2}{\upsilon-4}}\right)\right) \\ +\left(\varepsilon_t+\frac{\sqrt{(\upsilon-4)}}{\sqrt{4(\upsilon-2)}}\sqrt{-1+\sqrt{1+\frac{8\beta^2}{\upsilon-4}}}\right)^2\end{array}\right)}\right)\right) \\
& + \sum_{t=1}^{T}\beta\left(\varepsilon_t+\frac{\sqrt{(\upsilon-4)}}{\sqrt{4(\upsilon-2)}}\sqrt{-1+\sqrt{1+\frac{8\beta^2}{\upsilon-4}}}\right)-\frac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right) \\
& -\left(\frac{\upsilon+1}{2}\right)\sum_{t=1}^{T}\ln\left(\sqrt{\left(\frac{(\upsilon-2)(\upsilon-4)}{4\beta^2}\left(-1+\sqrt{1+\frac{8\beta^2}{\upsilon-4}}\right)\right)+\left(\varepsilon_t+\frac{\sqrt{(\upsilon-4)}}{\sqrt{4(\upsilon-2)}}\sqrt{-1+\sqrt{1+\frac{8\beta^2}{\upsilon-4}}}\right)^2}\right).
\end{aligned}
\tag{4.43}
$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \beta, \upsilon)$ is the parameter vector. The one-day-ahead VaR forecast of GARCH-GHST model is given by

$$
VaR_{t+1} = \hat{m}_{t+1} + F_p^{-1}(\varepsilon_t; \beta, \upsilon)\hat{h}_{t+1}. \tag{4.44}
$$

where $F_p{}^{-1}(\varepsilon_t; \beta, \upsilon)$ is the qf of GHST distribution at the p level. The **qghyp** function of R is used to obtain quantile estimation of standard GHST distribution.

4.1.8. Normal Inverse Gaussian Distribution

When the parameter $\lambda = -1/2$, GH distribution reduces to Normal Inverse Gaussian (NIG) distribution. NIG distribution is widely used in financial applications because of its ability to model excess kurtosis and skewness. Forsberg and Bollerslev [46] introduced the GARCH model under NIG innovation distribution to model the EUR/USD exchange rates. More recently, Chen and Lu [47] have demonstrated that forecasting VaR with NIG distribution yields more robust and accurate forecasts for daily time horizon based on the empirical results obtained from real data and simulated data sets. The pdf of NIG distribution is given by

$$f\left(x;\alpha,\beta,\mu,\delta\right)=\frac{\delta\alpha\exp\left(\delta\sqrt{\alpha^2-\beta^2}\right)\exp\left(\beta\left(x-\mu\right)\right)K_1\left(\alpha\sqrt{\delta^2+\left(x-\mu\right)^2}\right)}{\pi\sqrt{\delta^2+\left(x-\mu\right)^2}},\tag{4.45}$$

where $x\in\Re$, $0\leq|\beta|\leq\alpha$, $\delta>0$, $\mu\in\Re$. The parameters μ and δ are the location and scale parameters, respectively; α and β parameters control the shape of NIG distribution. When $\beta=0$, the distribution is symmetric. The mean and variance of NIG distribution are given by

$$\begin{aligned}E\left(X\right)&=\mu+\delta\tfrac{\beta}{\gamma}\\ Var\left(X\right)&=\delta\tfrac{\alpha^2}{\gamma^3}.\end{aligned}\tag{4.46}$$

where $\gamma=\sqrt{\alpha^2-\beta^2}$. The standardized NIG distribution can be obtained by replacing $\mu=-\delta\frac{\beta}{\gamma}$ and $\delta=\gamma^3/\alpha^2$. Then, the pdf of standardized NIG distribution with zero mean and unit variance is given by

$$f\left(\varepsilon;\alpha,\beta\right)=\frac{\frac{\gamma^3}{\alpha}\exp\left(\frac{\gamma^3}{\alpha^2}\sqrt{\alpha^2-\beta^2}\right)\exp\left(\beta\left(\varepsilon+\frac{\gamma^2\beta}{\alpha^2}\right)\right)K_1\left(\alpha\sqrt{\left(\frac{\gamma^3}{\alpha^2}\right)^2+\left(\varepsilon+\frac{\gamma^2\beta}{\alpha^2}\right)^2}\right)}{\pi\sqrt{\left(\frac{\gamma^3}{\alpha^2}\right)^2+\left(\varepsilon+\frac{\gamma^2\beta}{\alpha^2}\right)^2}},\tag{4.47}$$

Using the standardized NIG distribution, the ll function of GARCH-NIG model is given by

$$\begin{aligned}\ell\left(\boldsymbol{\psi}\right)&=T\ln\left(\tfrac{\gamma^3}{\alpha\pi}\right)+T\left(\tfrac{\gamma^3}{\alpha^2}\sqrt{\alpha^2-\beta^2}\right)+\beta\sum_{t=1}^{T}\left(\varepsilon_t+\tfrac{\gamma^2\beta}{\alpha^2}\right)-\tfrac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right)\\&+\sum_{t=1}^{T}\ln\left(K_1\left(\alpha\sqrt{\left(\tfrac{\gamma^3}{\alpha^2}\right)^2+\left(\varepsilon_t+\tfrac{\gamma^2\beta}{\alpha^2}\right)^2}\right)\right)-\sum_{t=1}^{T}\ln\left(\sqrt{\left(\tfrac{\gamma^3}{\alpha^2}\right)^2+\left(\varepsilon_t+\tfrac{\gamma^2\beta}{\alpha^2}\right)^2}\right).\end{aligned}\tag{4.48}$$

where $\boldsymbol{\psi}=(m,\omega,\gamma_1,\gamma_2,\alpha,\beta)$ is the parameter vector. The one-day-ahead VaR forecast of GARCH-NIG model is given by

$$VaR_{t+1}=\hat{m}_{t+1}+F_p^{-1}(\varepsilon_t;\alpha,\beta)\hat{h}_{t+1}.\tag{4.49}$$

where ${F_p}^{-1}\left(\varepsilon_t;\alpha,\beta\right)$ is the qf of NIG distribution at the p level. The **qnig** function of R is used to obtain quantile estimation of standard NIG distribution.

4.1.9. Generalized Error Distribution

Nelson [36] proposed the GED instead of assuming ε_t is normally distributed. Under this specification, the ll function of GARCH-GED model is given by

$$\ell(\boldsymbol{\psi}) = \sum_{t=1}^{T}\left[\ln\left(\frac{\kappa}{2}\right) - \frac{1}{2}\left|\frac{\varepsilon_t}{\delta}\right|^{\kappa} - (1+\kappa^{-1})\ln(2) - \ln\Gamma\left(\frac{1}{2}\right) - \frac{1}{2}\ln\left(h_t^2\right)\right], \tag{4.50}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \kappa)$ denotes the parameter vector, κ is tail-thickness parameter and,

$$\delta = \left(\frac{\Gamma\left(\frac{1}{\kappa}\right)}{2^{\frac{2}{\kappa}}\Gamma\left(\frac{3}{\kappa}\right)}\right)^{\frac{1}{2}}, \tag{4.51}$$

When the parameter $\kappa = 2$, the GED reduces to standard normal distribution. When $\kappa < 2$, GED has heavier tails than Gaussian distribution. The one-day-ahead VaR forecast of GARCH-GED model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + {F_p}^{-1}(\varepsilon_t, \kappa)\hat{h}_{t+1}. \tag{4.52}$$

where ${F_p}^{-1}(\varepsilon_t, \kappa)$ is the qf of GED at p level. The **qged** function of R is used to obtain quantile estimation of standardized GED distribution.

4.1.10. Skewed Generalized Error Distribution

The SGED provides an opportunity to model the skewness and kurtosis simultaneously. Lee et al. [5] introduced the GARCH-SGED model and concluded that GARCH model with SGED innovation process outperformed the GARCH-N model for all confidence levels. The pdf of standardized SGED is given by

$$f(\varepsilon;\kappa,\lambda) = C\exp\left(-\frac{|\varepsilon+\delta|^{\kappa}}{[1+sign(\varepsilon+\delta)\lambda]^{\kappa}\theta^{\kappa}}\right), \quad \varepsilon \in \Re \tag{4.53}$$

where

$$
\begin{aligned}
&C=\tfrac{\kappa}{2\theta}\Gamma\left(\tfrac{1}{\kappa}\right)^{-1},\theta=\Gamma\left(\tfrac{1}{\kappa}\right)^{0.5}\Gamma\left(\tfrac{3}{\kappa}\right)^{0.5}S(\lambda)^{-1}\\
&S(\lambda)=\sqrt{1+3\lambda^2-4A^2\lambda^2},\delta=\tfrac{2\lambda A}{S(\lambda)}\\
&A=\Gamma\left(\tfrac{2}{\kappa}\right)\Gamma\left(\tfrac{1}{\kappa}\right)^{-0.5}\Gamma\left(\tfrac{3}{\kappa}\right)^{-0.5},
\end{aligned}
\tag{4.54}
$$

where $\kappa > 0$ is the shape parameter, $-1 < \lambda < 1$ is skewness parameter. When the parameters $\kappa = 2$ and $\lambda = 0$, SGED reduces to standard normal distribution. Note that when the parameter $\lambda = 0$, SGED reduces to GED distribution. The ll function of GARCH-SGED model is given by

$$
\ell(\boldsymbol{\psi}) = +T\ln(C) - \sum_{t=1}^{T}\left(\frac{|\varepsilon_t+\delta|^{\kappa}}{[1+sign\,(\varepsilon_t+\delta)\,\lambda]^{\kappa}\theta^{\kappa}}\right) - \frac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right). \tag{4.55}
$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \lambda, \kappa)$ is the parameter vector. The one-day-ahead VaR forecast of GARCH-SGED model is given by

$$
VaR_{t+1} = \hat{m}_{t+1} + F_p^{-1}(\varepsilon_t, \kappa, \lambda)\hat{h}_{t+1}. \tag{4.56}
$$

where $F_p^{-1}(\varepsilon_t, \kappa, \lambda)$ is the qf of SGED at p level. The **qsged** function of R is used to obtain quantile estimation of standardized SGED distribution.

4.1.11. Generalized-T Distribution

The GT distribution was introduced by McDonald and Newey [48]. The GT distribution was employed to GARCH models by Bollerslev et al. [49]. The additional flexibility of the GT distribution enables it to capture a variety of shapes at the peak of the distribution as well as in the tails. The pdf of standardized GT distribution is given by

$$
f(\varepsilon;\kappa,\upsilon) = \frac{\gamma\kappa}{2\upsilon^{1/\kappa}B(1/\kappa,\upsilon)}\left(1+\frac{|\varepsilon\gamma|^{\kappa}}{\upsilon}\right)^{-(\upsilon+1/\kappa)}, \quad \varepsilon \in \Re \tag{4.57}
$$

where

$$
\gamma = \left[\frac{\upsilon^{2/\kappa}\Gamma\left(\frac{3}{\kappa}\right)\Gamma\left(\upsilon-\frac{2}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)\Gamma(\upsilon)}\right]^{1/2}. \tag{4.58}
$$

GT distribution contains Student-t and GED distributions as its submodels. When $\kappa = 2$, GT distribution reduces to Student-t distribution. When $\upsilon \to \infty$, GT distribution reduces to GED distribution. The ll function of GARCH-GT model is given by

$$\ell(\boldsymbol{\psi}) = T\ln(\gamma\kappa) - T\ln\left(2\upsilon^{1/\kappa}B(1/\kappa,\upsilon)\right) - \frac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right) - (\upsilon+1/\kappa)\sum_{t=1}^{T}\ln\left(1+\frac{|\varepsilon_t\gamma|^\kappa}{\upsilon}\right). \tag{4.59}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \upsilon, \kappa)$ denotes the parameter vector, $B[\cdot]$ is the beta function. The one-day-ahead VaR forecast of GARCH-GT model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_p{}^{-1}(\varepsilon_t, \kappa, \upsilon)\hat{h}_{t+1}. \tag{4.60}$$

where $F_p{}^{-1}(\varepsilon_t, \kappa, \upsilon)$ is the qf of GT at p level. The **qGT** function of R is used to obtain quantile estimation of standardized GT distribution.

4.1.12. Skewed Generalized-T Distribution

The SGT distribution was introduced by Theodossiou [50]. For the first time, SGT distribution was applied to GARCH models by Bali and Theodossiou [6]. The advantage of SGT distribution over the GT distribution is to provide an opportunity to model the skewness. The standardized SGT distribution is given by

$$f(\varepsilon;\eta,\kappa,\upsilon) = C\left(1+\frac{|\varepsilon+m|^\kappa}{((\upsilon+1)/\kappa)\left(1+sign(\varepsilon+m)\eta\right)^\kappa\tau^\kappa}\right)^{-\frac{\upsilon+1}{\kappa}}, \quad \varepsilon \in \Re \tag{4.61}$$

where

$$\begin{aligned}
&C = \tfrac{\kappa}{2}\left(\tfrac{\upsilon+1}{\kappa}\right)^{-\frac{1}{\kappa}}B\left(\tfrac{\upsilon}{\kappa},\tfrac{1}{\kappa}\right)^{-1}\tau^{-1}\\
&\tau = 1\Big/\sqrt{g-\rho^2}\\
&m = \rho\tau\\
&\rho = 2\eta B\left(\tfrac{\upsilon}{\kappa},\tfrac{1}{\kappa}\right)^{-1}\left(\tfrac{\upsilon+1}{\kappa}\right)^{\frac{1}{\kappa}}B\left(\tfrac{\upsilon-1}{\kappa},\tfrac{2}{\kappa}\right)\\
&g = \left(1+3\eta^2\right)B\left(\tfrac{\upsilon}{\kappa},\tfrac{1}{\kappa}\right)^{-1}\left(\tfrac{\upsilon+1}{\kappa}\right)^{\frac{2}{\kappa}}B\left(\tfrac{\upsilon-2}{\kappa},\tfrac{3}{\kappa}\right)
\end{aligned} \tag{4.62}$$

Here, the parameter $-1 < \eta < 1$ controls the skewness and the parameters $\kappa > 0$ and $\upsilon > 0$ control the kurtosis of SGT distribution. The ll function of GARCH model with SGT innovation distribution is given by

$$\begin{aligned}\ell\left(\boldsymbol{\psi}\right) = T\ln\left(C\right) - \left(\tfrac{\upsilon+1}{\kappa}\right)\sum_{t=1}^{T}\ln\left(1+\tfrac{|\varepsilon_t+m|^{\kappa}}{((\upsilon+1)/\kappa)(1+sign(\varepsilon_t+m)\eta)^{\kappa}\tau^{\kappa}}\right)\\ -\tfrac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right)\end{aligned}. \tag{4.63}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \eta, \kappa, \upsilon)$ denotes the parameter vector. The one-day-ahead VaR forecast of GARCH-SGT model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + {F_p}^{-1}(\varepsilon_t, \eta, \kappa, \upsilon)\hat{h}_{t+1}. \tag{4.64}$$

where ${F_p}^{-1}(\varepsilon_t, \eta, \kappa, \upsilon)$ is the left quantile of SGT at p level. The **qsgt** function of R is used to obtain quantile estimation of standardized SGT distribution. When $\eta = 0$, GARCH-SGT model reduces to GARCH-GT model.

4.1.13. Alpha Skew Generalized-T Distribution

It is commonly known that the normality assumption is unrealistic for many cases in real life problems. Therefore, researchers are interested to construct more flexible distributions as alternative to normal distribution to model both skewness and kurtosis. Acitas et al. [51] proposed the *alpha skew generalized-T* (ASGT) distribution which is the new skew extension of GT distribution. ASGT distribution provides new opportunities to model both skewness and fat-tailed structure of the data sets in many fields. Since the most of the financial time series possess skewness and leptokurtotic properties, ASGT distribution can be good candidate to remove lack of modeling ability of many distributions in terms of VaR forecast. The main idea behind to ASGT distribution is based on the alpha-skew normal (ASN) distribution. (see Elal-Olivero [52] for details).

The pdf of ASGT distribution is given by

$$f_{ASGT}(z;\alpha,\kappa,\upsilon)=\frac{(1-\alpha z)^2+1}{2+\alpha^2 c(\kappa,\upsilon)}\underbrace{\left[\frac{\kappa}{2\upsilon^{1/\kappa}B(1/\kappa,\upsilon)}\left(1+\frac{|z|^\kappa}{\upsilon}\right)^{-(\upsilon+1/\kappa)}\right]}_{f_{GT}(z;\kappa,\upsilon)},\quad z\in\Re \tag{4.65}$$

where $\kappa\upsilon>2$, $-\infty<\alpha<\infty$ is skewness and uni-bimodality parameter and $c(\kappa,\upsilon)$ is given by

$$c(\kappa,\upsilon)=\frac{\upsilon^{2/\kappa}\Gamma\left(\frac{3}{\kappa}\right)\Gamma\left(\upsilon-\frac{2}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)\Gamma(\upsilon)} \tag{4.66}$$

When $\alpha=0$, it is clear that ASGT distribution reduces to GT distribution with parameters κ and υ. ASGT and GT are nested distributions. Therefore, some mathematical properties of ASGT distribution can be obtained through GT distribution. The raw moments of ASGT distribution are given by

$$\mu_{2k}=E\left(Z^{2k}\right)=\frac{1}{2+\alpha^2c(\kappa,\upsilon)}\left[\frac{2\upsilon^{2k/\kappa}\Gamma\left(\frac{2k+1}{\kappa}\right)\Gamma\left(\upsilon-\frac{2k}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)\Gamma(\upsilon)}+\frac{\alpha^2\upsilon^{(2k+2)/\kappa}\Gamma\left(\frac{2k+3}{\kappa}\right)\Gamma\left(\upsilon-\frac{2k+2}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)\Gamma(\upsilon)}\right],\kappa\upsilon>2k+2$$

$$\mu_{2k-1}=E\left(Z^{2k-1}\right)=\frac{1}{2+\alpha^2c(\kappa,\upsilon)}\left[\frac{-2\alpha\upsilon^{2k/\kappa}\Gamma\left(\frac{2k+1}{\kappa}\right)\Gamma\left(\upsilon-\frac{2k}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)\Gamma(\upsilon)}\right],\kappa\upsilon>2k \tag{4.67}$$

where $k\in\mathbb{Z}^+$.

Figure 2 displays the location, spread and shape measures of ASGT distribution. Note that equations of these measures can be found in Acitas et al. [51]. Based on these plots, the following results are concluded for fixed parameter $\upsilon=4$: the parameter α has more significant effect on mean, skewness and kurtosis measures than parameter κ; when the parameter α increases, variance increases; when the parameter κ increases, variance decreases.

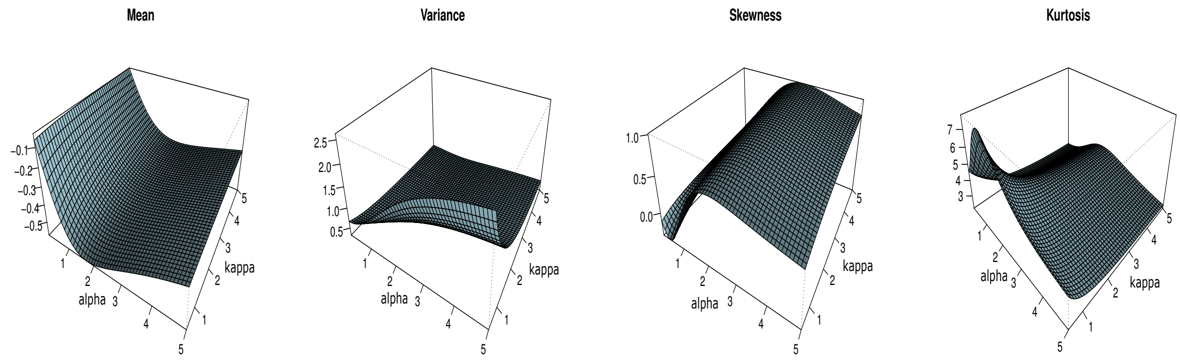


Figure 2: The plots for location, spread and shape measures of the ASGT distribution for $\upsilon = 4$

Here, GARCH-ASGT model is introduced by means of standardized ASGT distribution. Let $\varepsilon = (Z - \mu)/\sigma$, the random variable Z can be expressed as $Z = \varepsilon\sigma + \mu$. Thus, the pdf of standardized ASGT distribution is given by

$$f(\varepsilon;\alpha,\kappa,\upsilon) = \frac{\sigma\left[(1-\alpha(\varepsilon\sigma+\mu))^2+1\right]}{2+\alpha^2 c(\kappa,\upsilon)}\left[\frac{\kappa}{2\upsilon^{1/\kappa}B(1/\kappa,\upsilon)}\left(1+\frac{|\varepsilon\sigma+\mu|^\kappa}{\upsilon}\right)^{-(\upsilon+1/\kappa)}\right], \tag{4.68}$$

where μ and σ are mean and standard deviation of ASGT distribution, respectively. The mean and variance of ASGT distribution can be easily obtained in closed form using Equation (4.67). The ll function of GARCH model with ASGT innovation distribution is given by

$$\begin{aligned}\ell(\boldsymbol{\psi}) = &\sum_{t=1}^{T}\ln\left[\sigma(1-\alpha(\varepsilon_t\sigma+\mu))^2+1\right] - T\left\{2+\alpha^2 c(\kappa,\upsilon)\right\} - \tfrac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right)\\ &+\sum_{t=1}^{T}\ln\left[\tfrac{\kappa}{2\upsilon^{1/\kappa}B(1/\kappa,\upsilon)}\left(1+\tfrac{|\varepsilon_t\sigma+\mu|^\kappa}{\upsilon}\right)^{-(\upsilon+1/\kappa)}\right].\end{aligned} \tag{4.69}$$

where $\boldsymbol{\psi} = (m,\omega,\gamma_1,\gamma_2,\alpha,\kappa,\upsilon)$ denotes the parameter vector. The one-day-ahead VaR forecast of GARCH-ASGT model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + {F_p}^{-1}(\varepsilon_t, \alpha, \kappa, \upsilon)\hat{h}_{t+1}. \tag{4.70}$$

where ${F_p}^{-1}(\varepsilon_t, \alpha, \kappa, \upsilon)$ is the qf of ASGT at p level. The quantile estimation of ASGT distribution can be obtained by means of numerical integration methods. The *integrate* function of stats package of R software can be used. This function uses the globally adaptive interval subdivision method in connection with extrapolation by Wynn's Epsilon algorithm (see for details: Piessens et al. [53]). GARCH-ASGT model contains the following models as its submodels:

- ✓ If $\alpha = 0$, GARCH-ASGT reduces to GARCH-GT
- ✓ If $\alpha = 0$ and $\upsilon \to \infty$, GARCH-ASGT reduces to GARCH-GED
- ✓ If $\alpha = 0$ and $\kappa = 2$, GARCH-ASGT reduces to GARCH-Student-t
- ✓ If $\alpha = 0$, $\upsilon \to \infty$ and $\kappa = 2$, GARCH-ASGT reduces to GARCH-N

4.1.14. A New Skew Extension of Generalized Normal Distribution

The researchers show great interest to generalize the well-known distribution to increase the flexibility of the distribution. On this basis, SN distribution was studied and generalized extensively by many researchers [54, 55, 56, 57, 58]. The alternative normal distribution, enables to model skewness, called *alpha-skew-normal* (ASN), was introduced by Elal-Olivero [52]. The pdf of ASN distribution is given by

$$f(x;\alpha) = \frac{(1-\alpha x)^2 + 1}{2+\alpha^2}\phi(x), x \in \Re, \alpha \in \Re, \tag{4.71}$$

where α is an additional parameter controls the both skewness and uni-bimodal shapes.

Now, a new skew extension of the generalized-Normal (GN) distribution, introduced by Nadarajah [59], is proposed by means of Elal-Olivero [52] approach. Note that GN distribution is also known as GED. The pdf of the normal distribution is given by

$$f(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \tag{4.72}$$

where $x\in\Re$, $\mu\in\Re$ and $\sigma>0$. A generalization of (4.72), called as the GN distribution, was introduced by Nadarajah [59] with replacing the power 2 with $\kappa>0$. The pdf of GN distribution is given by

$$f(x;\mu,\sigma,\kappa)=K\exp\left\{-\left|\frac{x-\mu}{\sigma}\right|^{\kappa}\right\}, \tag{4.73}$$

where $K=\frac{\kappa}{2\sigma\Gamma(1/\kappa)}$. The GN distribution reduces to normal distribution for $\kappa=2$. Using the standardized random variable $Z=(X-\mu)/\sigma$, the pdf of standardized GN (SGN) distribution is given by

$$f_{SGN}(z;\kappa)=\frac{\kappa\exp\{-|z|^{\kappa}\}}{2\Gamma(1/\kappa)}. \tag{4.74}$$

The *nth* moment of standardized GN distribution is given by

$$E\left(Z^k\right)=\frac{1+(-1)^k}{2\Gamma(1/\kappa)}\Gamma\left(\frac{1+k}{\kappa}\right). \tag{4.75}$$

Definition 1. *Using the approach of Elal-Olivero [52], a new skew extension of GN distribution, called Alpha Skew Generalized Normal (ASGN) is proposed. The pdf of ASGN distribution is obtained by*

$$f(x;\alpha,\kappa)=\frac{(1-\alpha x)^2+1}{2+\alpha^2\frac{\Gamma(3/\kappa)}{\Gamma(1/\kappa)}}\frac{\kappa\exp\{-|x|^{\kappa}\}}{2\Gamma(1/\kappa)},\quad x\in\Re, \tag{4.76}$$

where $\alpha\in\Re$ controls the skewness and bi-modality and $\kappa>0$ controls the kurtosis of the ASGN distribution.

The ASGN distribution contains following distributions as its sub-models:

✓ If $\alpha=0$, the ASGN reduces to the GN.

✓ If $\kappa=2$, the ASGN reduces to the ASN.

✓ If $\alpha = 0$ and $\kappa = 2$, the ASGN reduces to normal.

✓ If $\kappa = 1$, the ASGN reduces to Alpha-Skew Laplace.

✓ If $\alpha = 0$ and $\kappa = 1$, the ASGN reduces to Laplace.

✓ If $\alpha = 0$ and $\kappa = 2$, the ASGN reduces to normal.

Figure 3 displays the plots of density functions of the ASGN distribution. As seen in Figure 3, the ASGN distribution provides new opportunities to model skewness, bimodality and fat-tailed structure of the data sets in many fields. Since the most of the financial time series possess skewness and leptokurtotic properties, the ASGN distribution could be a good alternative to remove lack of modeling ability of many distributions in terms of VaR forecast.

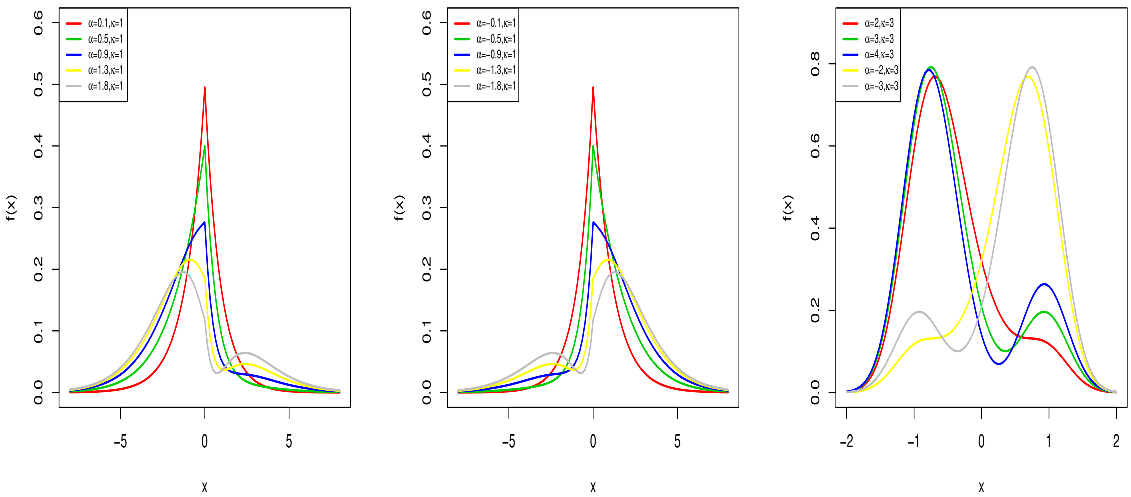


Figure 3: Plots of density functions for the ASGN distribution for several parameter values.

4.1.14.1. Moments

Proposition 1. *Let $X \sim ASGN(\alpha,\kappa)$, then kth moment of X is given by*

$$\begin{array}{l} E\left(X^k\right)_{ASGN} = \int\limits_{-\infty}^{\infty} \frac{x^k\left(2-2\alpha x+\alpha^2x^2\right)}{c(\alpha,\kappa)} f_{SGN}(x)\,dx \\ = \frac{2}{c(\alpha,\kappa)}\left[E\left(Z^k\right)_{SGN} - \alpha E\left(Z^{k+1}\right)_{SGN} + \frac{\alpha^2}{2}E\left(Z^{k+2}\right)_{SGN}\right], \end{array} \tag{4.77}$$

where $c(\alpha,\kappa) = 2+\alpha^2\left[\Gamma(3/\kappa)/\Gamma(1/\kappa)\right]$ and $E\left(Z^k\right)_{SGN}$ is the kth moment of standartized $GN(\kappa)$, defined in (4.75).

Proof.

$$\begin{array}{l} E\left(X^k\right) = \frac{1}{c(\alpha,\kappa)} \int\limits_{-\infty}^{\infty} x^k\left(2-2\alpha x+\alpha^2x^2\right) f_{SGN}(x)\,dx \\ = \frac{1}{c(\alpha,\kappa)}\left\{\int\limits_{-\infty}^{\infty} 2x^k f_{SGN}(x)\,dx - 2\alpha\int\limits_{-\infty}^{\infty} x^{k+1} f_{SGN}(x)\,dx + \alpha^2\int\limits_{-\infty}^{\infty} x^{k+2} f_{SGN}(x)\,dx\right\} \\ = \frac{2}{c(\alpha,\kappa)}\left[E\left(Z^k\right) - \alpha E\left(Z^{k+1}\right) + \frac{\alpha^2}{2}E\left(Z^{k+2}\right)\right]. \end{array} \tag{4.78}$$

□

The first four moments of the ASGN distribution are obtained using (4.77) as follows:

$$\begin{array}{l} E(X) = \frac{2}{c(\alpha,\kappa)}\left[-\alpha\frac{\Gamma(3/\kappa)}{\Gamma(1/\kappa)}\right], \\ E\left(X^2\right) = \frac{2}{c(\alpha,\kappa)}\left[\frac{\Gamma(3/\kappa)}{\Gamma(1/\kappa)} + \frac{\alpha^2}{2}\frac{\Gamma(5/\kappa)}{\Gamma(1/\kappa)}\right], \\ E\left(X^3\right) = \frac{2}{c(\alpha,\kappa)}\left[-\alpha\frac{\Gamma(5/\kappa)}{\Gamma(1/\kappa)}\right], \\ E\left(X^4\right) = \frac{2}{c(\alpha,\kappa)}\left[\frac{\Gamma(5/\kappa)}{\Gamma(1/\kappa)} + \frac{\alpha^2}{2}\frac{\Gamma(7/\kappa)}{\Gamma(1/\kappa)}\right]. \end{array} \tag{4.79}$$

Using the first four moments of the ASGN distribution, skewness and kurtosis can be obtained by

$$\begin{array}{l} \gamma_1 = \frac{E(X-\mu)^3}{\sigma^3}, \\ \gamma_2 = \frac{E(X-\mu)^4}{\sigma^4} - 3. \end{array} \tag{4.80}$$

where $\mu = E(X)$ and $\sigma = \sqrt{E\left(X^2\right)-\mu^2}$.

Figure 4 displays the location, spread and shape measures of ASGN distribution. It is clear from Figure 4 that the parameter α has more effect on mean, variance, skewness, and kurtosis than parameter λ.

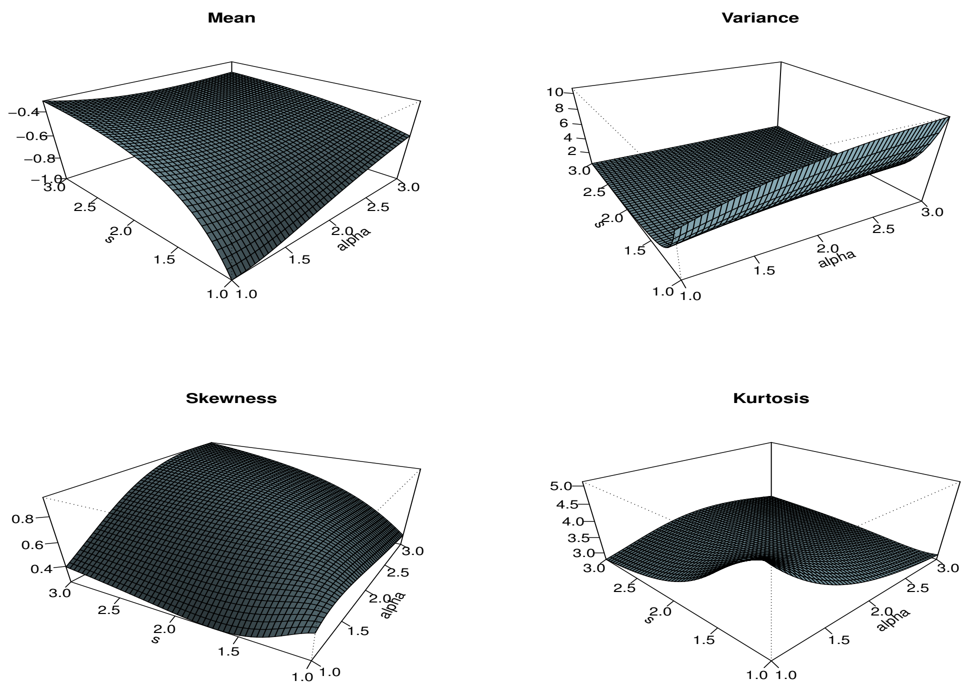


Figure 4: The plots for location, spread and shape measures of the ASGN distribution.

4.1.14.2. Distribution Function

Proposition 2. *Let $X \sim ASGN\left(\alpha, \kappa\right)$, then the cdf of X is given by*

$$F\left(x\right) = \frac{1}{c\left(\alpha, \kappa\right)}\left[2F_{SGN}\left(x\right) - 2\alpha F_1\left(x\right) + \alpha^2 F_2\left(x\right)\right], \tag{4.81}$$

where $\Gamma\left(a, b\right)$ denotes the incomplete gamma function and

$$F_{SGN}(x) = \begin{cases} \frac{\Gamma(1/\kappa,(-x)^{\kappa})}{2\Gamma(1/\kappa)}, x \leq 0 \\ 1 - \frac{\Gamma(1/\kappa,(-x)^{\kappa})}{2\Gamma(1/\kappa)}, x > 0 \end{cases}$$

$$F_1(x) = \begin{cases} \frac{\Gamma(2/\kappa,(-x)^{\kappa})}{2\Gamma(1/\kappa)}, x \leq 0 \\ 1 - \frac{\Gamma(2/\kappa,(-x)^{\kappa})}{2\Gamma(1/\kappa)}, x > 0 \end{cases}$$

$$F_2(x) = \begin{cases} \frac{\Gamma(3/\kappa,(-x)^{\kappa})}{2\Gamma(1/\kappa)}, x \leq 0 \\ 1 - \frac{\Gamma(3/\kappa,(-x)^{\kappa})}{2\Gamma(1/\kappa)}, x > 0 \end{cases}$$

Proof.

$$\begin{aligned} F(x) &= \tfrac{1}{c(\alpha,\kappa)} \int_{-\infty}^{x} (2 - 2\alpha t + \alpha^2 t^2) f_{SGN}(t)\, dt \\ &= \tfrac{1}{c(\alpha,\kappa)} \left\{ \int_{-\infty}^{x} 2 f_{SGN}(t)\, dt - 2\alpha \int_{-\infty}^{x} t f_{SGN}(t)\, dt + \alpha^2 \int_{-\infty}^{x} t^2 f_{SGN}(t)\, dt \right\} \\ &= \tfrac{2}{c(\alpha,\kappa)} \left[2F_{SGN}(x) - 2\alpha F_1(x) + \alpha^2 F_2(x) \right]. \end{aligned} \tag{4.82}$$

Here, $F_1(x)$, $F_2(x)$ and $F_{SGN}(x)$ can be obtained easily by following the results given in Nadarajah [59]. The main idea is based on the $z = (-x)^{\kappa}$ transformation. □

4.1.14.3. Stochastic Representation

Definition 2. *Let the random variable X has the following pdf*

$$f_{SC-ASGN}(x;\alpha,\kappa) = \frac{(2 + \alpha^2 x^2)}{2 + \alpha^2 \frac{\Gamma(3/\kappa)}{\Gamma(1/\kappa)}} f_{GN}(x), \tag{4.83}$$

where $-\infty < \alpha < \infty$. This density function is called as symmetric-component random variable of $ASGN(\alpha,\kappa)$. The pdf in (4.83) is denoted as $S \sim SCASGN(\alpha,\kappa)$. Note that the SCASGN reduces to $GN(\kappa)$ for $\alpha = 0$.

Here, two algorithms are given for generating random observations from the ASGN distribution.

Proposition 3. *(Acceptance-Rejection Algorithm)*

Let $f(x)$ denotes the pdf of ASGN and $f_1(x)$ denotes the pdf of SCASGN, then the supremum of these two functions is given by

$$M = \sup_x \frac{f(x)}{f_1(x)} = \frac{\sqrt{2}+1}{\sqrt{2}}. \tag{4.84}$$

The below algorithm is useful to generate the random observations from $X \sim ASGN(\alpha, \kappa)$.

1. *Generate $U \sim uniform(0,1)$ and $S \sim SCASGN(\alpha, \kappa)$,*

2. *If $U < \frac{1}{M}\frac{f(x)}{f_1(x)}$, then $X = S$, otherwise, return to Step 1.*

Proposition 4. *(Inverse Transform Algorithm)*

Let $F(x)$ denotes the cdf of ASGN distribution. The below algorithm is useful to generate the random observations from $X \sim ASGN(\alpha, \kappa)$.

1. *Generate $U \sim uniform(0,1)$,*

2. *Solve non-linear equation $F(x) - u = 0$, then $X = x$.*

*The **uniroot** function of R software can be used to solve given non-linear equation.*

4.1.14.4. Standardized ASGN Distribution

Let $\varepsilon = (X - \mu)/\sigma$, the random variable X can be expressed as $X = \varepsilon\sigma + \mu$. Thus, the pdf of standardized ASGN distribution is given by

$$f(\varepsilon; \alpha, \kappa) = \frac{\sigma\left\{(1 - \alpha[\varepsilon\sigma + \mu])^2 + 1\right\}}{2 - \alpha^2\frac{\Gamma(3/\kappa)}{\Gamma(1/\kappa)}} \frac{\kappa \exp\{-|\varepsilon\sigma + \mu|^\kappa\}}{2\Gamma(1/\kappa)}. \tag{4.85}$$

where μ and σ are given in (4.79), respectively.

4.1.14.5. Estimation and Simulation Study

Estimation

Let $x_1, x_2, \ldots, x_n$ be a random sample from $ASGN\left(\alpha,\kappa\right)$ distribution. Using Equation (4.76), the ll function of ASGN distribution is given by

$$\ell\left(\mathbf{\Theta}\right)=\sum_{i=1}^{n}\ln\left[\frac{\left(1-\alpha x_i\right)^2+1}{c\left(\alpha,\kappa\right)}\right]+\sum_{i=1}^{n}\ln\left[\frac{\kappa\exp\left\{-|x_i|^{\kappa}\right\}}{2\Gamma\left(1/\kappa\right)}\right]. \tag{4.86}$$

where $\mathbf{\Theta}=\left(\alpha,\kappa\right)$ denotes the parameter vector. Taking partial derivatives of (4.86) with respect to parameters, the following normal equations are obtained as

$$\frac{\partial\ell}{\partial\alpha}=\sum_{i=1}^{n}\frac{4x_i\left(\alpha x_i-1\right)\Gamma\left(1/\kappa\right)+2\alpha\left(\alpha x_i-2\right)\Gamma\left(3/\kappa\right)}{\left(\alpha^2x_i^2-2\alpha x_i+2\right)\left(\alpha^2\Gamma\left(3/\kappa\right)+2\Gamma\left(1/\kappa\right)\right)},$$

$$\begin{aligned}\frac{\partial\ell}{\partial\kappa}=\sum_{i=1}^{n}&\left(\begin{array}{l}2\Gamma\left(1/\kappa\right)\frac{\exp\left(-|x_i|^{\kappa}\right)}{2k\Gamma\left(1/\kappa\right)}\\ \times\left(\exp\left(-|x_i|^{\kappa}\right)-\kappa\exp\left(-|x_i|^{\kappa}\right)|x_i|^{\kappa}\log\left(|x_i|\right)+\exp\left(-|x_i|^{\kappa}\right)\psi^{(0)}\left(1/\kappa\right)\right)\end{array}\right)\\ -n&\left(\frac{\left(\kappa^2\Gamma(1/\kappa)\right)^{-1}\alpha^2\Gamma(3/\kappa)\left(\psi^{(0)}(1/\kappa)-3\psi^{(0)}(3/\kappa)\right)}{\frac{\alpha^2\Gamma(3/\kappa)}{\Gamma(1/\kappa)}+2}\right).\end{aligned}$$

where $\psi^{n}\left(x\right)$ denotes the n_{th} derivative of digamma function. The simultaneous solutions of the $\frac{\partial\ell}{\partial\alpha}=0$, and $\frac{\partial\ell}{\partial\kappa}=0$ equations give the maximum likelihood estimates (MLEs) of $\left(\alpha,\kappa\right)$, say, $\left(\hat{\alpha},\hat{\kappa}\right)$. Since the likelihood equations contain non-linear functions, it is not possible to obtain explicit forms of the MLEs. Therefore, they have to be solved by using numerical methods. Note that S-Plus, R or MATLAB can be used for obtaining the MLEs of the parameters.

For interval estimation of the model parameters, the observed information matrix $J\left(\Theta\right)$ is required. Under standard regularity conditions, when $n\to\infty$, the distribution of $\hat{\Theta}$ can be approximated by multivariate normal distribution, $N_p\left(0,J\left(\hat{\Theta}\right)^{-1}\right)$ where $J\left(\hat{\Theta}\right)$ is the total observed information matrix evaluated at $\hat{\Theta}$. The elements of $J\left(\hat{\Theta}\right)$ are given in Appendix.

Simulation Study

Now, the simulation study is conducted to investigate the performance of the MLEs of ASGN distribution. For this goal, $10,000$ samples of sizes $n = 50, 150$ and 500 are generated from ASGN distribution by means of inverse transform method. The following measures are used to evaluate the precision of the MLEs: averages of the estimates (AEs), biases and mean square errors (MSEs).

The results of simulations study is given in Table 1. Table 1 shows that the AEs are closed to nominal values. It means that estimates are quite stable. Moreover, biases and MSEs approach to zero when n increases. It means that consistency property of the MLEs holds. Figure 5 displays the true densities and the density functions evaluated at average values of the MLEs given in Table 1 for $n = 500$. Figure 6 displays the Q-Q plots of MLEs of ASGN distributions for $n = 500$. Figure 6 verifies the asymptotic normality property of MLE. It is clear that the MLEs of ASGN distribution are near to normal distribution.

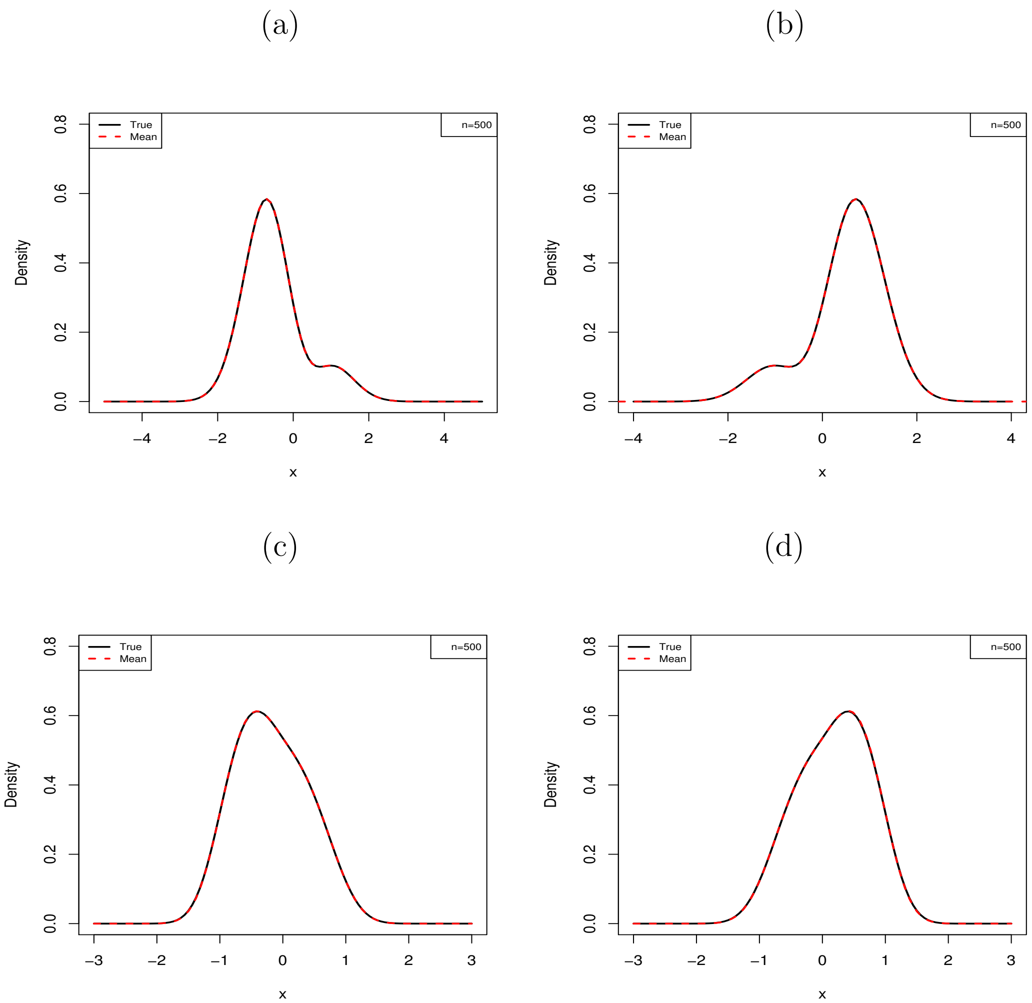


Figure 5: The pdf plots of ASGN distribution evaluated at the true parameter values and the AEs for $n = 500$, (a) $\alpha = 2, \kappa = 2$ (b) $\alpha = -2, \kappa = 2$ (c) $\alpha = 0.5, \kappa = 3$ and (d) $\alpha = -0.5, \kappa = 3$

Table 1: The simulation results of the ASGN distribution for $n = 50, 150$ and 500.

n	Parameters	$\alpha = 2$	$\kappa = 2$		n	Parameters	$\alpha = -2$	$\kappa = 2$	
		AE	Bias	MSE			AE	Bias	MSE
50	α	2.0718	0.0718	0.3096	50	α	-2.1417	-0.1417	0.3558
	κ	2.0374	0.0374	0.0526		κ	2.0628	0.0628	0.0470
150	α	2.0276	0.0276	0.0897	150	α	-2.0139	-0.0139	0.0876
	κ	2.0205	0.0205	0.0130		κ	2.0069	0.0069	0.0130
500	α	2.0050	0.0050	0.0272	500	α	-2.0028	-0.0028	0.0248
	κ	2.0005	0.0005	0.0039		κ	2.0002	0.0002	0.0034
n	Parameters	$\alpha = 0.5$	$\kappa = 3$		n	Parameters	$\alpha = -0.5$	$\kappa = 3$	
		AE	Bias	MSE			AE	Bias	MSE
50	α	0.5341	0.0341	0.0904	50	α	-0.5468	-0.0468	0.1033
	κ	3.4125	0.4125	0.4580		κ	3.5423	0.5423	0.3969
150	α	0.5013	0.0013	0.0246	150	α	-0.5049	-0.0049	0.0261
	κ	3.1527	0.1527	0.2360		κ	3.0685	0.0685	0.1492
500	α	0.4992	-0.0008	0.0071	500	α	-0.5032	-0.0032	0.0074
	κ	3.0236	0.0236	0.0414		κ	3.0353	0.0353	0.0424

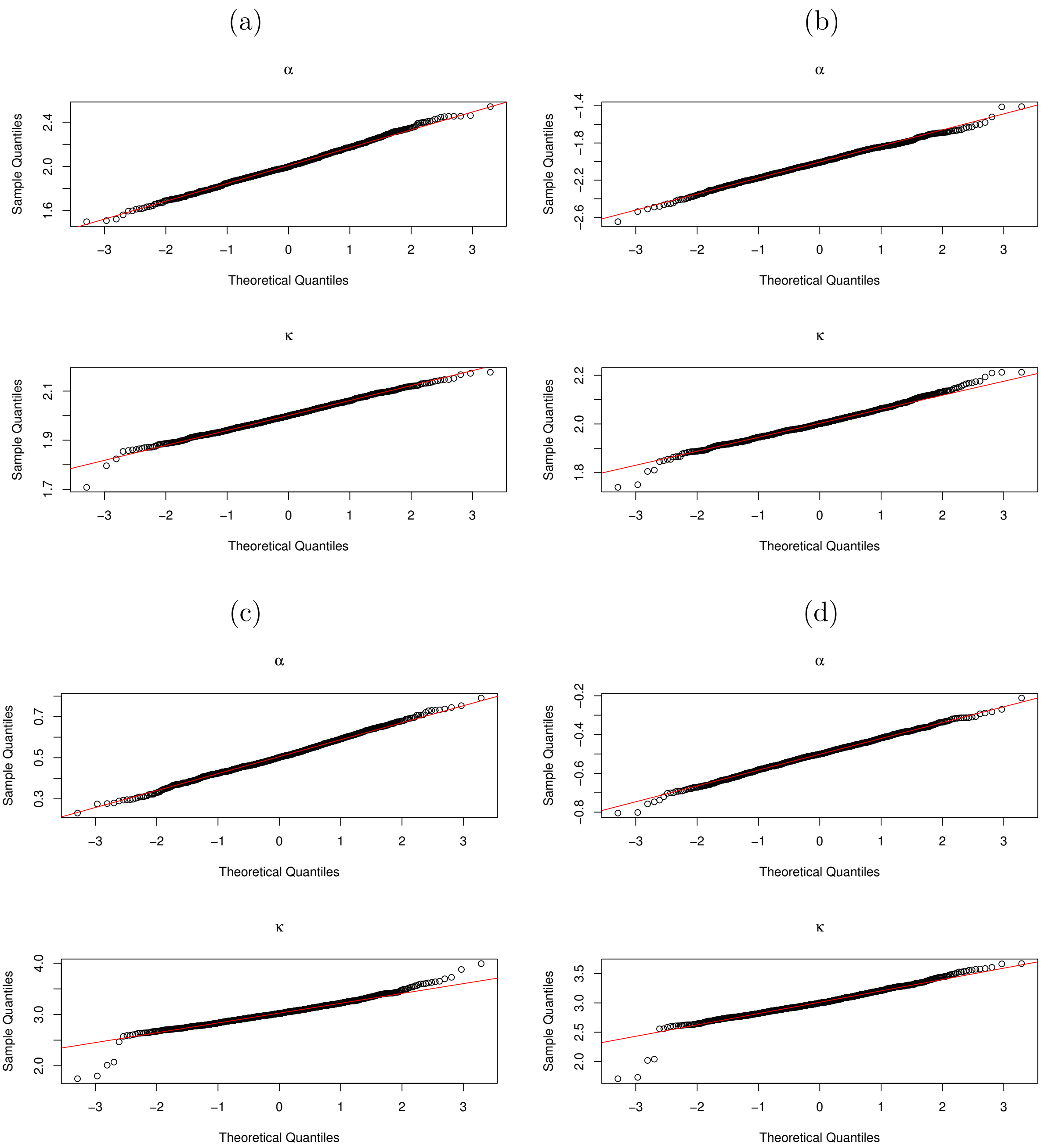


Figure 6: QQ normality plots of MLEs of ASGN for $n = 500$, (a) $\alpha = 2, \kappa = 2$ (b) $\alpha = -2, \kappa = 2$ (c) $\alpha = 0.5, \kappa = 3$ and (d) $\alpha = -0.5, \kappa = 3$

Here, GARCH-ASGN model is introduced by means of standardized ASGN distribution given in (4.85). The ll function of GARCH model with the ASGN innovation distribution is given by

$$\ell(\boldsymbol{\psi}) = \sum_{t=1}^{T} \ln\left(\sigma\left\{(1-\alpha\left[\varepsilon_t\sigma+\mu\right])^2+1\right\}\right) - T\ln\left(2-\alpha^2\frac{\Gamma(3/\kappa)}{\Gamma(1/\kappa)}\right) - T\ln\left(2\Gamma\left(1/\kappa\right)\right)$$
$$+\sum_{t=1}^{T}\ln\left(\kappa\exp\left\{-|\varepsilon_t\sigma+\mu|^{\kappa}\right\}\right) - \frac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right) \tag{4.87}$$

where $\boldsymbol{\psi} = (m, \omega, \gamma_1, \gamma_2, \alpha, \kappa)$. The one-day-ahead VaR forecast of GARCH-ASGN model is given by

$$VaR_{t+1} = \hat{m}_{t+1} + F_p^{-1}(\varepsilon_t, \alpha, \kappa)\hat{h}_{t+1}, \tag{4.88}$$

where $F_p^{-1}(\varepsilon_t, \alpha, \kappa)$ is the qf of ASGN distribution at p level. The quantile estimations of ASGN distribution is obtained by means of numerical integration methods. Note that when the parameter $\alpha = 0$, GARCH-ASGN model reduces to GARCH-GN model.

4.1.15. A New Generalized Skew-T Distribution

In this section, a new generalization of skew-T distribution is introduced by means of combining the approaches of Azzalini [40] and Elal-Olivero [52]. Azzalini and Capitanio [42] proposed the ST distribution which is the skew generalization of Student-t distribution. ST distribution has some advantages over the normal and Student-t distributions in many application fields, such as financial risk modeling and management. ST distribution enables to model the skewness and kurtosis, simultaneously. This property of ST distribution gives an opportunity to increase its popularity in forecasting VaR.

Motivated by the approaches of Azzalini and Capitanio [42] and Elal-Olivero [52], a new generalization of Skew-T, called Generalized Alpha Skew-T (GAST), is proposed.

Definition 3. *The pdf of GAST distribution is obtained as*

$$f(x;\alpha,\lambda,\upsilon) = \frac{(1-\alpha x)^2+1}{c(\alpha,\lambda,\upsilon)} t(x;\upsilon)\, T\left(\sqrt{\frac{1+\upsilon}{x^2+\upsilon}}\lambda x; \upsilon+1\right), \quad \upsilon > 2, x \in \Re, \tag{4.89}$$

where

$$c\,(\alpha,\lambda,\upsilon) = 1-\alpha\underbrace{\left[\delta(\upsilon/\pi)^{1/2}\Gamma\left((\upsilon-1)/2\right)\Big/\Gamma\left(\upsilon/2\right)\right]}_{E[Y]}+\frac{\alpha^2}{2}\underbrace{\left[\upsilon/(\upsilon-2)\right]}_{E[Y^2]}. \tag{4.90}$$

Here, $E\left[Y\right]$ *and* $E\left[Y^2\right]$ *are the moments of ST distribution.*

GAST distribution contains some important distributions as follows:

- ✓ If $\alpha = 0$, GAST reduces to ST.
- ✓ If $\lambda = 0$, GAST reduces to Alpha-skew-T.
- ✓ If $\nu \to \infty$, GAST reduces to ASN.
- ✓ If $\alpha = 0$ and $\nu \to \infty$, GAST reduces to SN.
- ✓ If $\alpha = 0$ and $\lambda = 0$, GAST reduces to Student-t.
- ✓ If $\alpha = 0$, $\nu \to \infty$ and $\lambda = 0$, GAST reduces to normal.

Note that alpha-skew-T (AST) distribution is the sub model of Alpha-Skew Generalized-T distribution introduced by Acitas et al. [51].

Figure 7 displays the plots for the pdf of GAST distribution. As seen in Figure 7, GAST distribution provides new opportunities to model skewness, bimodality and fat-tailed structure of the data sets in many fields. Since the most of the financial time series possess skewness and leptokurtotic properties, GAST distribution could be a good candidate to remove lack of modeling ability of many distributions in terms of VaR forecast.

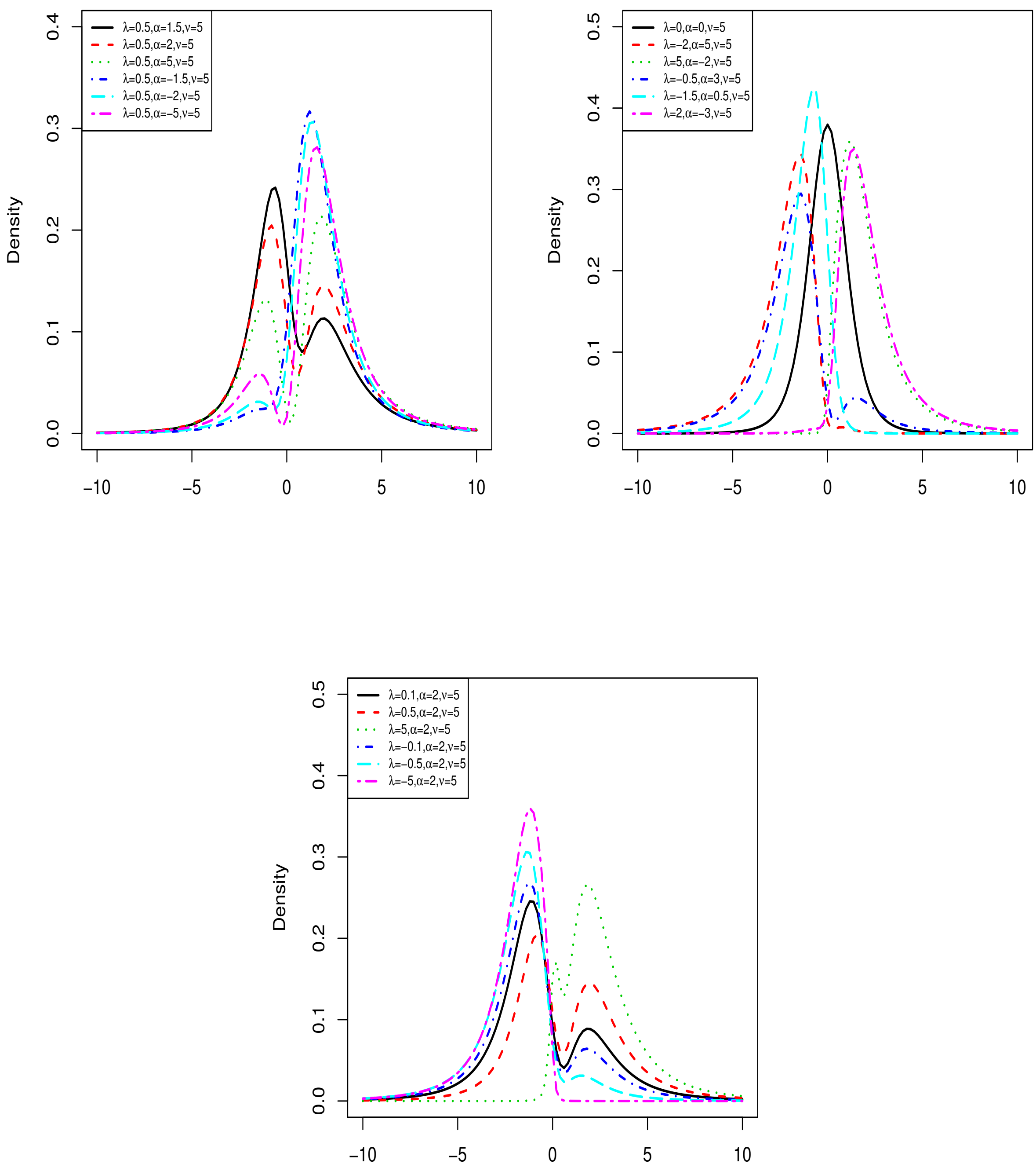


Figure 7: Probability density plots of GAST distribution for several parameter values

4.1.15.1. Moments

Proposition 5. *The moments of GAST distribution can be obtained from the moments of ST distribution. Let $X \sim GAST(\alpha, \lambda, \upsilon)$, then kth moment of X is given by*

$$E\left[X^k\right] = \frac{\left\{E\left[Y^k\right] - \alpha E\left[Y^{k+1}\right] + \frac{\alpha^2}{2} E\left[Y^{k+2}\right]\right\}}{c(\alpha, \lambda, \upsilon)}, \tag{4.91}$$

where $E\left[Y^k\right]$ is the kth moment of $ST(\lambda, \nu)$, defined in Equation (4.26) *and* (4.18).

The first four moments of GAST distribution are obtained using Equation (4.77) as follows:

$$E[X] = \frac{\left\{\delta(\upsilon/\pi)^{1/2}\Gamma((\upsilon-1)/2)/\Gamma(\upsilon/2) - \alpha\upsilon/(\upsilon-2) + \frac{\alpha^2}{2}\left(\frac{\upsilon}{2}\right)^{3/2}\left(\Gamma\left(\frac{\upsilon-3}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right)\sqrt{\frac{2}{\pi}}3\left(\delta-\delta^3\right)\right\}}{c(\alpha,\lambda,\upsilon)},$$

$$E[X^2] = \frac{\left\{\upsilon/(\upsilon-2) - \alpha\left(\frac{\upsilon}{2}\right)^{3/2}\left(\Gamma\left(\frac{\upsilon-3}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right)\sqrt{\frac{2}{\pi}}3\left(\delta-\delta^3\right) + \frac{3\alpha^2}{2}\left(\frac{\upsilon}{2}\right)^{4/2}\left(\Gamma\left(\frac{\upsilon-4}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right)\right\}}{c(\alpha,\lambda,\upsilon)},$$

$$E[X^3] = \frac{\left\{\begin{array}{l} \left(\frac{\upsilon}{2}\right)^{3/2}\left(\Gamma\left(\frac{\upsilon-3}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right)\sqrt{\frac{2}{\pi}}3\left(\delta-\delta^3\right) - 3\alpha\left(\frac{\upsilon}{2}\right)^{4/2}\left(\Gamma\left(\frac{\upsilon-4}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right) \\ +\frac{\alpha^2}{2}\left(\frac{\upsilon}{2}\right)^{5/2}\left(\Gamma\left(\frac{\upsilon-5}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right)\sqrt{\frac{2}{\pi}}\left(15\delta-10\delta^3+3\delta^5\right) \end{array}\right\}}{c(\alpha,\lambda,\upsilon)},$$

$$E[X^4] = \frac{\left\{\begin{array}{l} 3\left(\frac{\upsilon}{2}\right)^{4/2}\left(\Gamma\left(\frac{\upsilon-4}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right) - \alpha\left(\frac{\upsilon}{2}\right)^{5/2}\left(\Gamma\left(\frac{\upsilon-5}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right)\sqrt{\frac{2}{\pi}}\left(15\delta-10\delta^3+3\delta^5\right) \\ +\frac{15\alpha^2}{2}\left(\frac{\upsilon}{2}\right)^{6/2}\left(\Gamma\left(\frac{\upsilon-6}{2}\right)/\Gamma\left(\frac{\upsilon}{2}\right)\right) \end{array}\right\}}{c(\alpha,\lambda,\upsilon)}.$$

The skewness and kurtosis measures of GAST distribution can be obtained by means of first four raw moments given above. Figure 8 displays location, spread and shape measures of GAST distribution. Figure 8 provides graphical interpretation about how the additional shape parameter effects the location and dispersion measures. It is clear in Figure 8 that the parameter α has more effect on mean, variance, skewness, and kurtosis than parameter λ.

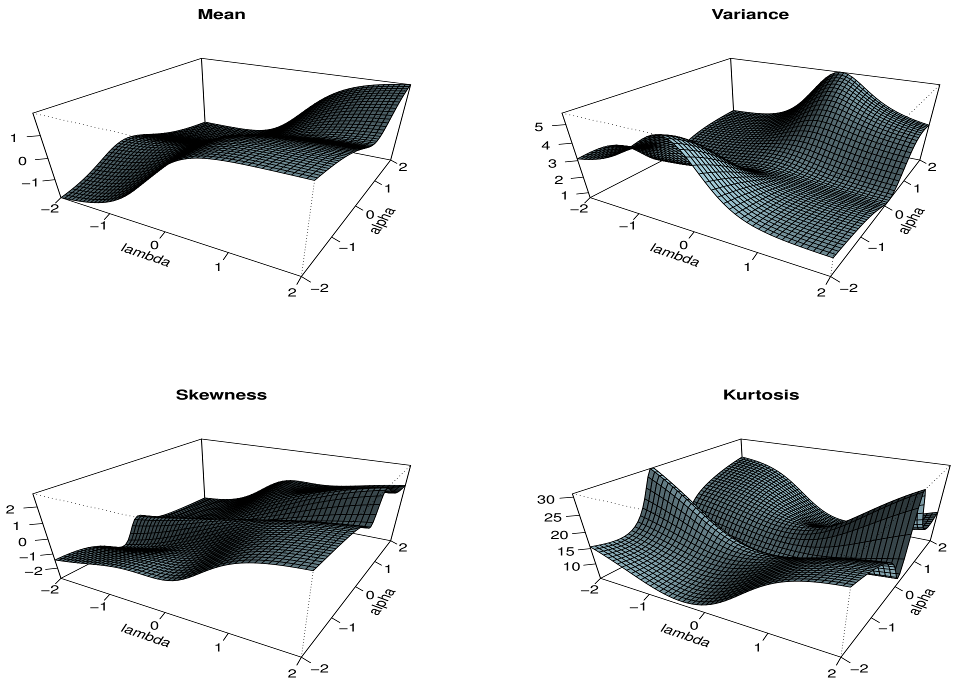


Figure 8: The plots for location, spread and shape measures of GAST distribution for $\nu = 7$

4.1.15.2. Stochastic Representation

Proposition 6. *If the random variables* $W \sim AST\left(\alpha, \upsilon\right)$ *and* $Z \sim t\left(\upsilon+1\right)$ *are independent, then we have*

$$W| \left\{ \sqrt{\frac{1+\upsilon}{W^2+\upsilon}} \lambda W > Z \right\} \sim GAST\left(\lambda, \alpha, \upsilon\right). \tag{4.92}$$

Proof. Let $X = W| \left\{ \sqrt{\frac{1+\upsilon}{W^2+\upsilon}} \lambda W > Z \right\}$, we obtain

$P\left(X \leq x\right) = P\left(W \leq x | \sqrt{\frac{1+\upsilon}{W^2+\upsilon}} \lambda W > Z\right) = \frac{P\left(W \leq x, \sqrt{\frac{1+\upsilon}{W^2+\upsilon}} \lambda W > Z\right)}{P\left(\sqrt{\frac{1+\upsilon}{W^2+\upsilon}} \lambda W > Z\right)}$

Then, we get

$$P\left(W \leq x, \sqrt{\frac{1+\upsilon}{W^2+\upsilon}}\lambda W > Z\right) = \int\limits_{-\infty}^{x} \frac{(1-\alpha u)^2+1}{2+\alpha^2\frac{\upsilon}{\upsilon-2}} t\left(u,\upsilon\right) T\left(\sqrt{\frac{1+\upsilon}{u^2+\upsilon}}\lambda u, \upsilon+1\right) du$$

$$P\left(\sqrt{\frac{1+\upsilon}{W^2+\upsilon}}\lambda W > Z\right) = \int\limits_{-\infty}^{\infty} \frac{(1-\alpha u)^2+1}{2+\alpha^2\frac{\upsilon}{\upsilon-2}} t\left(u,\upsilon\right) T\left(\sqrt{\frac{1+\upsilon}{u^2+\upsilon}}\lambda u, \upsilon+1\right) du = \frac{1}{2+\alpha^2\frac{\upsilon}{\upsilon-2}} c\left(\alpha,\lambda,\upsilon\right)$$

Thus, we obtain

$$P\left(X \leq x\right) = \frac{\int\limits_{-\infty}^{x} \frac{(1-\alpha u)^2+1}{2+\alpha^2\frac{\upsilon}{\upsilon-2}} t(u,\upsilon) T\left(\sqrt{\frac{1+\upsilon}{u^2+\upsilon}}\lambda u,\upsilon+1\right) du}{\frac{1}{2+\alpha^2\frac{\upsilon}{\upsilon-2}} c(\alpha,\lambda,\upsilon)} = \int\limits_{-\infty}^{x} \frac{(1-\alpha u)^2+1}{c(\alpha,\lambda,\upsilon)} t\left(u,\upsilon\right) T\left(\sqrt{\frac{1+\upsilon}{u^2+\upsilon}}\lambda u, \upsilon+1\right) du . \tag{4.93}$$

Then, the pdf corresponding to Equation (4.93) is given by

$$f\left(x;\alpha,\lambda,\upsilon\right) = \frac{(1-\alpha x)^2+1}{c\left(\alpha,\lambda,\upsilon\right)} t\left(x;\upsilon\right) T\left(\sqrt{\frac{1+\upsilon}{x^2+\upsilon}}\lambda t; \upsilon+1\right). \tag{4.94}$$

It is clear that Equation (4.94) is the pdf of GAST distribution. □

Proposition 7. *From Equation (4.94), the algorithm to generate data from GAST distribution can be given as follows:*

1. *Generate* $W \sim AST\left(\alpha,\upsilon\right)$ *and* $Z \sim t\left(\upsilon+1\right)$
2. *If* $\sqrt{\frac{1+\upsilon}{W^2+\upsilon}}\lambda W > Z$, $X = W$, *otherwise go to Step 1.*

4.1.15.3. Estimation and Simulation Study

Estimation

Let $x_1, x_2, \ldots, x_n$ be a random sample from $GAST(\alpha, \lambda, \upsilon)$ distribution. Using Equation (4.89), the ll function of GAST distribution is given by

$$\ell(\Theta) = \sum_{i=1}^{n} \log\left[\frac{(1-\alpha x_i)^2 + 1}{c(\alpha, \lambda, \upsilon)}\right] + \sum_{i=1}^{n} \log\left[t(x_i; \upsilon)\right] + \sum_{i=1}^{n} \log\left[T\left(\sqrt{\frac{1+\upsilon}{x_i^2+\upsilon}}\lambda x_i; \upsilon+1\right)\right], \tag{4.95}$$

where $\mathbf{\Theta} = (\alpha, \lambda, \upsilon)$ is the parameter vector and $c(\alpha, \lambda, \upsilon)$ is defined in Equation (4.90). Taking partial derivatives from Equation (4.95) with respect to parameters, the following normal equations are obtained as

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} \frac{(2-\alpha x_i^2)\, c_\alpha(\alpha, \lambda, \upsilon) + x_i^2 c(\alpha, \lambda, \upsilon)}{(\alpha x_i^2 - 2)\, c(\alpha, \lambda, \upsilon)},$$

$$\frac{\partial \ell}{\partial \lambda} = -\frac{n c_\lambda(\alpha, \lambda, \upsilon)}{c(\alpha, \lambda, \upsilon)} + \sum_{i=1}^{n} x_i \sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\omega^*,$$

$$\frac{\partial \ell}{\partial \upsilon} = -\frac{n c_\upsilon(\alpha, \lambda, \upsilon)}{c(\alpha, \lambda, \upsilon)} + \sum_{i=1}^{n} \tau_i^* + \sum_{i=1}^{n}\left[\frac{\lambda x_i (x_i^2 - 1)}{2\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}(x_i^2+\upsilon)^2}\right]\omega_i^*,$$

where

$$\omega_i^* = \frac{t\left(\sqrt{\frac{1+\upsilon}{x_i^2+\upsilon}}\lambda x_i; \upsilon+1\right)}{T\left(\sqrt{\frac{1+\upsilon}{x_i^2+\upsilon}}\lambda x_i; \upsilon+1\right)},$$

$$\tau_i^* = \frac{t_\upsilon(x_i; \upsilon)}{t(x_i; \upsilon)}.$$

Here, $c_\upsilon(\alpha, \lambda, \upsilon)$, $c_\lambda(\alpha, \lambda, \upsilon)$, $c_\alpha(\alpha, \lambda, \upsilon)$ and $t_\upsilon(x_i; \upsilon)$ are the partial derivatives of $c(\alpha, \lambda, \upsilon)$ and $t(x_i; \upsilon)$ with respect to υ, λ, α and υ. The maximum likelihood estimates (MLEs) of $(\alpha, \lambda, \upsilon)$, say, $\left(\hat{\alpha}, \hat{\lambda}, \hat{\upsilon}\right)$, are the simultaneous solutions of the equations: $\frac{\partial \ell}{\partial \alpha} = 0$, $\frac{\partial \ell}{\partial \lambda} = 0$ and $\frac{\partial \ell}{\partial \upsilon} = 0$. Since the likelihood equations contain non-linear functions, it is not possible to obtain explicit forms of the MLEs. Therefore, they have to be solved by using numerical methods. S-Plus, R or MATLAB can be used for obtaining the MLEs of the parameters.

For interval estimation of the model parameters, the observed information matrix $J(\Theta)$ is required. Under standard regularity conditions, when $n \to \infty$, the distribution of $\hat{\Theta}$ can be approximated by multivariate normal distribution, $N_p\left(0, J\left(\hat{\Theta}\right)^{-1}\right)$ where $J\left(\hat{\Theta}\right)$ is the total observed information matrix evaluated at $\hat{\Theta}$. The elements of $J\left(\hat{\Theta}\right)$ are given in Appendix.

Simulation Study

Here, the simulation study is conducted to investigate the performance of the MLEs of GAST distribution. For this goal, $10,000$ samples of sizes $n = 50, 150$ and 500 are generated from GAST distribution by means of inverse transform method. The following measures are used to evaluate the precision of the MLEs: AEs, biases, mean relative errors (MREs) and MSEs. The results of simulations study are given in Table 2. Table 2 shows that the AEs are closed to nominal values. It means that estimates are quite stable. Moreover, biases and MSEs approach to zero when n increases. It means that consistency property of the MLEs holds. Figure 9 displays the true densities and the density functions evaluated at average values of the MLEs given in Table 2 for $n = 500$. Figure 10 displays the Q-Q plots of MLEs of ASGN distributions for $n = 500$. Figure 10 verifies the asymptotic normality property of MLE. It is clear that the MLEs of ASGN distribution are near to normal distribution.

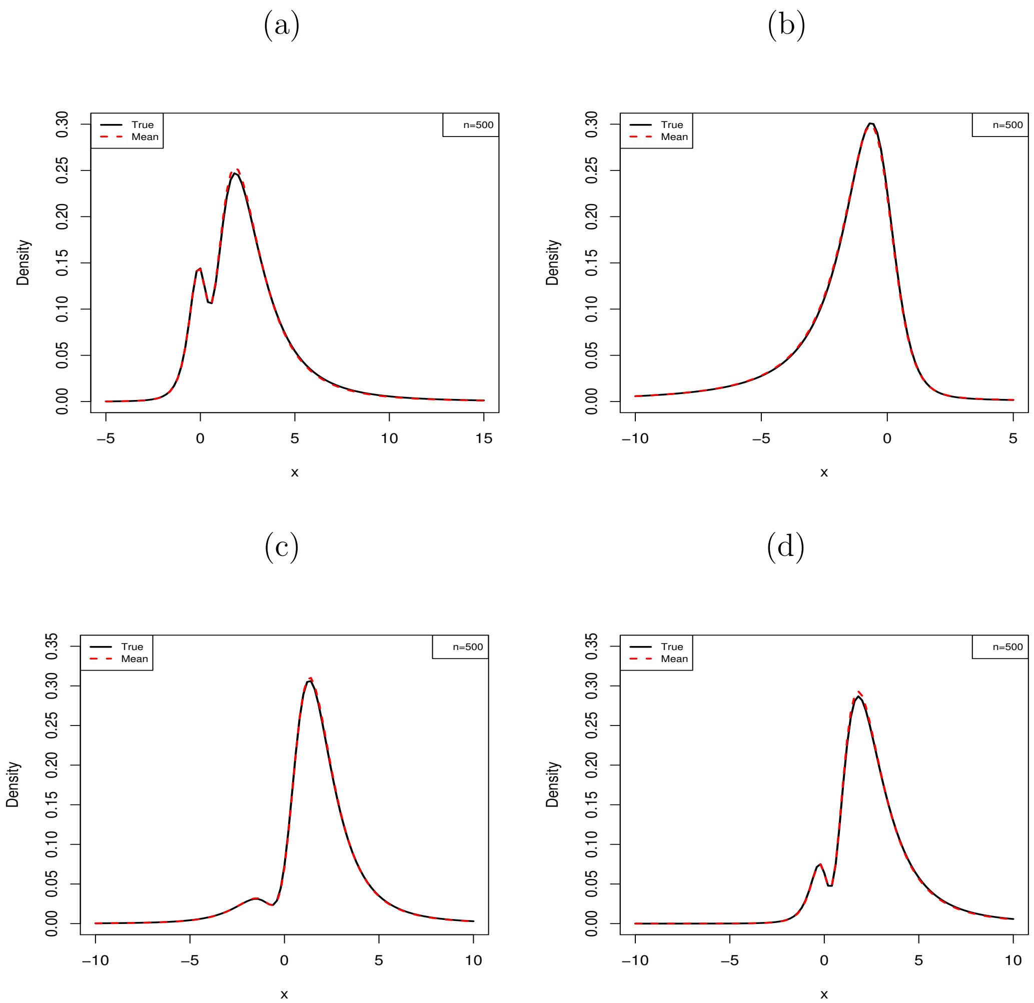


Figure 9: Plots of the density functions for the GAST distribution at the true parameter values and at the AEs for $n = 500$, (a) $\lambda = 2, \alpha = 2, \upsilon = 5$ (b) $\lambda = -0.5, \alpha = 0.5, \upsilon = 3$ (c) $\lambda = 0.5, \alpha = -2, \upsilon = 5$ and (d) $\lambda = 2, \alpha = 3, \upsilon = 5$

Table 2: The simulation results of the GAST distribution for $n = 50, 150$ and 500.

n	Parameters	$\lambda = 2$	$\alpha = 2$		$\upsilon = 5$	n	Parameters	$\lambda = -0.5$	$\alpha = 0.5$		$\upsilon = 3$
		AE	Bias	MRE	MSE			AE	Bias	MRE	MSE
50	λ	2.215	0.215	1.108	0.559	50	λ	-0.554	-0.054	1.108	0.146
	α	2.116	0.116	1.058	0.231		α	0.634	0.134	1.268	0.059
	υ	5.234	0.234	1.047	0.471		υ	3.497	0.497	1.166	0.421
150	λ	2.039	0.038	1.020	0.095	150	λ	-0.453	0.046	0.906	0.029
	α	2.038	0.038	1.019	0.064		α	0.641	0.141	1.282	0.032
	υ	5.136	0.137	1.027	0.121		υ	3.495	0.495	1.165	0.292
500	λ	2.005	0.005	1.003	0.021	500	λ	-0.481	0.078	0.962	0.012
	α	2.026	0.026	1.013	0.015		α	0.531	0.131	1.062	0.020
	υ	5.130	0.130	1.026	0.048		υ	3.060	0.060	1.020	0.002
n	Parameters	$\lambda = 0.5$	$\alpha = -2$		$\upsilon = 5$	n	Parameters	$\lambda = 2$	$\alpha = 3$		$\upsilon = 5$
		AE	Bias	MRE	MSE			AE	Bias	MRE	MSE
50	λ	0.558	0.058	1.116	0.098	50	λ	2.174	0.174	1.087	0.628
	α	-2.322	-0.322	1.161	1.241		α	3.337	0.337	1.112	1.511
	υ	5.291	0.291	1.058	0.972		υ	5.198	0.198	1.040	0.286
150	λ	0.514	0.014	1.028	0.014	150	λ	2.060	0.060	1.030	0.137
	α	-2.152	-0.152	1.076	0.359		α	3.187	0.187	1.062	0.562
	υ	5.182	0.182	1.036	0.248		υ	5.138	0.138	1.028	0.125
500	λ	0.509	0.009	1.018	0.004	500	λ	2.022	0.022	1.011	0.026
	α	-2.068	-0.068	1.034	0.057		α	3.037	0.037	1.012	0.103
	υ	5.018	0.018	1.004	0.012		υ	5.121	0.121	1.024	0.042

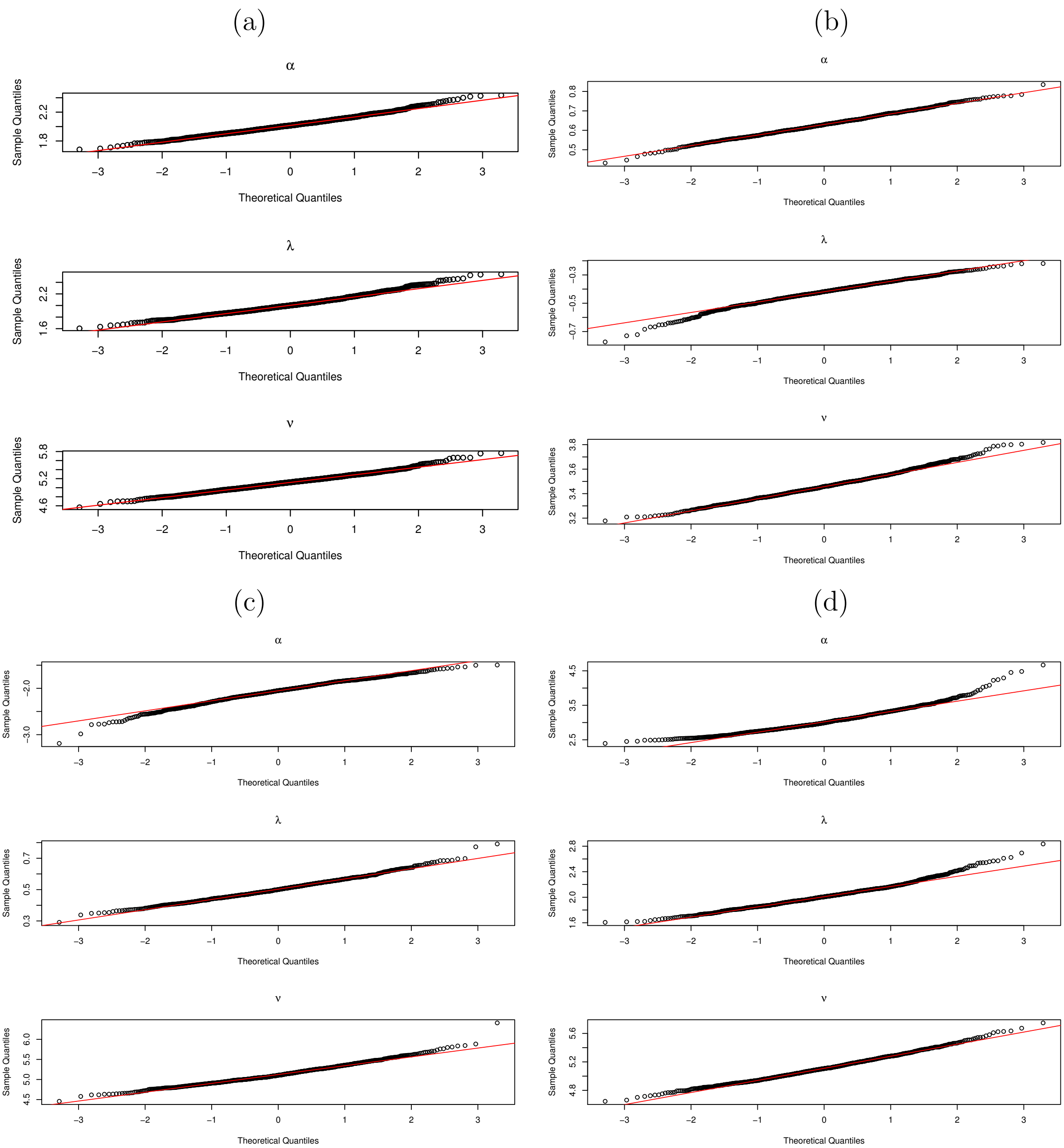


Figure 10: QQ normality plots of MLEs of GAST for $n = 500$, (a) $\lambda = 2, \alpha = 2, \upsilon = 5$ (b) $\lambda = -0.5, \alpha = 0.5, \upsilon = 3$ (c) $\lambda = 0.5, \alpha = -2, \upsilon = 5$ and (d) $\lambda = 2, \alpha = 3, \upsilon = 5$

Some of mathematical properties of the GAST distribution are presented above. Here, GARCH-GAST model is introduced by means of standardized GAST distribution. The standardized GAST distribution is given by

$$f\left(\varepsilon;\alpha,\lambda,\upsilon\right)=\frac{\sigma(1-\alpha\left(\varepsilon\sigma+\mu\right))^{2}+1}{c\left(\alpha,\lambda,\upsilon\right)}t\left(\left(\varepsilon\sigma+\mu\right);\upsilon\right)T\left(\sqrt{\frac{1+\upsilon}{\left(\varepsilon\sigma+\mu\right)^{2}+\upsilon}}\lambda\left(\varepsilon\sigma+\mu\right);\upsilon+1\right), \tag{4.96}$$

where $\upsilon > 4$, μ and σ can be obtained by (4.91). Mean and variance of GAST distribution can be obtained by Equation (4.91). The ll function of GARCH model with GAST innovation distribution is given by

$$\begin{aligned}\ell\left(\psi\right)=&\sum_{t=1}^{T}\ln\left[\frac{\sigma(1-\alpha(\varepsilon_t\sigma+\mu))^2+1}{c(\alpha,\lambda,\upsilon)}\right]+\sum_{t=1}^{T}\ln\left[t\left(\left(\varepsilon_t\sigma+\mu\right);\upsilon\right)\right]\\&+\sum_{t=1}^{T}\ln\left(T\left(\sqrt{\frac{1+\upsilon}{(\varepsilon_t\sigma+\mu)^2+\upsilon}}\lambda\left(\varepsilon_t\sigma+\mu\right);\upsilon+1\right)\right)-\frac{1}{2}\sum_{t=1}^{T}\ln\left(h_t^2\right).\end{aligned} \tag{4.97}$$

where $\psi=(m,\omega,\gamma_1,\gamma_2,\alpha,\lambda,\upsilon)$. The one-day-ahead VaR forecast of GARCH-GAST model is given by

$$VaR_{t+1}=\hat{m}_{t+1}+F_p^{-1}(\varepsilon_t,\alpha,\lambda,\nu)\hat{h}_{t+1}. \tag{4.98}$$

where $F_p^{-1}(\varepsilon_t,\alpha,\lambda,\nu)$ is the qf of the GAST at p level. The quantile estimation of GAST distribution is obtained by means of numerical integration method.

4.2. Semiparameteric Value-at-Risk Models

In this section, the most popular semi-parametric VaR model, based on the EVT, is given comprehensively.

4.2.1. Extreme Value Theory

EVT has numerous applications in actuarial sciences, engineering, finance and environmental fields. Two methods are widely used to apply EVT. The first method is Block-Maxima (BM). BM is a traditional method and used to model the observed maximum (minimum) values in blocks. The lengths of blocks can be determined as monthly or yearly. Generalized Extreme Value (GEV) distribution is used to model maximum (minimum) values observed in the blocks. The other method is Peaks-over-Threshold (POT). POT method is widely used in finance and econometric modeling and deals with the conditional excess distribution over a given threshold value. The POT method is based on the modeling of tail distribution with generalized Pareto distribution (GPD). Let $F_u(y)$ denotes the conditional excess distribution. $F_u(y)$ is defined as follows

$$F_u(y) = P(X - u \leq y / X > u), \tag{4.99}$$

where $y = X - u$ denotes the excess values over a given threshold value. $F_u(y)$ can be re-defined as follows

$$F_u(y) = \frac{\Pr\{X - u \leq y, X > u\}}{\Pr(X > u)} = \frac{F(x) - F(u)}{1 - F(u)}. \tag{4.100}$$

A theorem given by [60] and [61] shows that the excess distribution function, $F_u(y)$, can be approximated by GPD for a sufficiently high threshold u. The pdf of GPD is given by

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - (1 + \xi \frac{y}{\sigma})^{-1/\xi}, & \xi \neq 0 \\ 1 - e^{-y/\sigma}, & \xi = 0 \end{cases} \tag{4.101}$$

where $y \geq 0$ for $\xi \geq 0$ and $0 \leq y \leq \frac{\sigma}{\xi}$ for $\xi < 0$ and ξ and σ are shape and scale parameters of GP distribution, respectively. Isolating $F(x)$ from (4.100), we get

$$F(x) = (1 - F(u))F_u(y) + F(u), \tag{4.102}$$

where $F_u(y)$ is the cdf of GPD and $F(u) = (n - N_u)/n$. Then, substituting (4.101) in (4.102), the following estimate for $F(x)$ is obtained

$$\hat{F}(x) = 1 - \frac{N_u}{n}(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x - \hat{u}))^{-1/\hat{\xi}}, \tag{4.103}$$

where $\hat{\xi}$ and $\hat{\sigma}$ are MLEs of ξ and σ, respectively. Note that N_u is the number of observations over the defined threshold value u. Inverting (4.103) for a given probability p, VaR_p can be obtained as

$$VaR_p = \hat{u} + \frac{\hat{\sigma}}{\hat{\xi}}\left[\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\xi}} - 1\right]. \qquad (4.104)$$

In VaR estimation using POT method is applied to raw return data assuming the distribution to be stationary or unconditional without considering the time-varying volatility. POT method can also be considered as a dynamic framework, where the conditional distribution of F is taken into account and the volatility of returns is captured. The dynamic POT method is given comprehensively at below.

4.2.1.1. Dynamic POT-GPD Approach

McNeil and Frey [17] proposed the dynamic POT-GPD model combining GARCH model with POT-GPD method. This model is known as GARCH-EVT model and consists of two-step estimation procedure, can be given as follows:

1. The benchmark model, GARCH (1,1) is fitted to raw returns by the quasi maximum likelihood estimation (QMLE) to maximize the ll function and obtain the one day ahead forecasts of m_{t+1} and h_{t+1} from the fitted model and extract the standardized residuals for the next step. QML estimation of ψ is defined as:

$$\hat{\psi}_T = \arg\max\ L_T(\psi) = argmin\ \hat{\text{I}}_T(\psi),$$
$$\hat{I}_T(\psi) = T^{-1}\sum_{t=1}^{T} l_t \qquad and \qquad l_t = \frac{\varepsilon_t^2}{h_t^2} + \log h_t^2.$$

2. Standardized residual obtained from the GARCH(1,1) is modeled with POT method to estimate $VaR_p(\varepsilon_t)$. Using the one-day-ahead forecasts of m_{t+1} and m_{t+1} , VaR_{t+1} forecast can be obtained by

$$VaR_{t+1} = \hat{m}_{t+1} + VaR_p(\varepsilon_t)\,\hat{h}_{t+1}.$$

where $VaR_p(\varepsilon_t)$ can be obtained from (4.104).

Threshold selection is difficult task and essential part for tail modeling with GPD. The below sub-section is devoted to detail information about threshold selection tools in POT-GPD method. Scarrott and MacDonald [62] provided comprehensive survey on threshold estimation methods.

Here, the well-known threshold determination tools, Mean Excess (ME), Hill and Threshold Stability plots, are examined.

✓ The first tool is ME plot for determination of threshold. ME function is given by

$$e_n(u) = \frac{\sum_{i=1}^{n} (X_i - u)}{\sum_{i=1}^{n} I_{\{X_i > u\}}}. \tag{4.105}$$

Figure 11 displays the ME plot of ISE-100 index. As seen from Figure 11, 0.02 can be chosen as optimal threshold value for this data set.

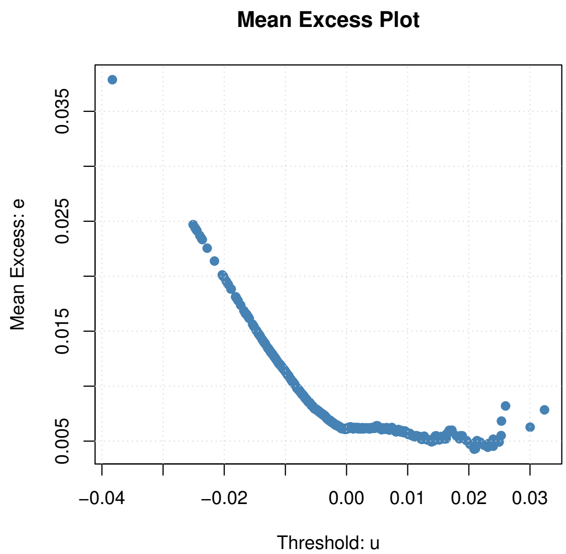


Figure 11: The mean excess plot

Note that the upward linear trend above a given threshold value indicates that the financial returns follows GPD with positive shape parameter, ξ. The opposite case indicates that he financial returns follows GPD with negative shape parameter, ξ.

✓ The second tool is Hill plot, introduced by Hill [63], for determination of threshold. Let $X_{(1)}, X_{(2)}, ..., X_{(n)}$ ascending ordered sequence, the Hill estimator is defined by

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^{k} \log\left(\frac{X_{(i)}}{X_{(k+1)}}\right). \tag{4.106}$$

where n is the sample size, k is the exceedances, $\alpha = \frac{1}{\xi}$ is the tail index. The Hill plot shows the estimated $\hat{\xi}$ against to threshold value or upper order statistics. The optimal threshold value can be determined where the parameter $\hat{\xi}$ is nearly stable.

✓ The third tool is threshold stability plot, described in Coles [64]. This plot examines the sensitivity of maximum likelihood estimates of GPD parameters to selected threshold value. The optimal threshold can be determined using the similar approach in Hill plot.

These tools only provide a graphical information about the optimal threshold value. It is not possible to determine the optimal threshold value for rolling window estimation method or any continuous process with these type tools. Many traditional threshold selection methods were proposed in literature, for instance, upper 10% rule by DuMouchel [65], $k = \sqrt{n}$ rule by Ferreira et al. [66] and $k = n^{2/3}/\log(\log(n))$ by Loretan and Philips [67].
Here, we use the method proposed by Reiss and Thomas [68]. The method is based on the minimization of below function:

$$\frac{1}{k} \sum_{i=1}^{k} i^{\beta} \left|\hat{\delta}_i - median\left(\hat{\delta}_1, \hat{\delta}_2, \hat{\delta}_3, \ldots, \hat{\delta}_k\right)\right|. \tag{4.107}$$

where $\hat{\delta}$ is the estimated shape parameter of GPD and $0 \leq \beta \leq 0.5$ is the tuning parameter. To apply the threshold determination tools and obtain GPD parameter estimation, R packages can be used. The most used R packages for EVT can be given as follows: **fextRemes**, **evir**, **evmix** and **tea**. The **tea** package contains almost all threshold estimation methods in literature. In this study, the **tea** is used to determine optimal threshold value by methods of Reiss and Thomas [68]. The parameter estimation of GPD is obtained by minimizing the negative ll function of GPD model by **optim** function of R software.

4.3. Nonparametric Value-at-Risk Models

4.3.1. Historical Simulation

In this section, the non-parametric VaR models are investigated. The most simple VaR model is HS model. The VaR based on the HS model is given by

$$VaR_{t+1} = Quantile\left\{\{X_t\}_{t=1}^{n}, p\right\}. \tag{4.108}$$

where p is the quantile at which VaR is calculated.

4.3.2. Filtered Historical Simulation

In this section, FHS model is defined. The steps of FHS model are given as follows:

✓ Let r_t denote the daily log-returns. The benchmark GARCH(1,1) model, introduced by Bollerslev [35], is defined as follows:

$$\begin{aligned} r_t &= m_t + e_t, \\ e_t &= \varepsilon_t\, h_t,\ \varepsilon_t \sim i.i.d. \\ h_t^2 &= \omega + \gamma_1\, e_{t-1}^2 + \gamma_2 h_{t-1}^2, \end{aligned} \tag{4.109}$$

The ll function of r_t with $f(\varepsilon_t;\tau)$ density function is given by

$$\ell(\psi) = \sum_{t=1}^{T}\left[\ln\left(f(\varepsilon_t;\tau)\right) - \frac{1}{2}\ln\left(h_t^2\right)\right], \tag{4.110}$$

where $\psi = (m, \omega, \gamma_1, \gamma_2, \tau)$ is the parameter vector, τ is the shape parameter(s) of $f(\varepsilon_t;\tau)$ and $\varepsilon_t = \frac{e_t}{h_t}$.

- ✓ The standardized residuals of estimated GARCH(1,1) model are extracted as follows:
$$\varepsilon_t = \frac{\hat{e}_t}{\hat{h}_t}, \tag{4.111}$$
where $\hat{e}_t$ is the estimated residual and $\hat{h}_t$ is the corresponding daily estimated volatility.

- ✓ Now, we can generate the first simulated residual by randomly (with replacement) draw standardized residuals from the dataset with multiplying the one-day ahead volatility forecast:

$$e^*_{t+1} = \varepsilon^*_1 \hat{h}_{t+1}, \tag{4.112}$$

- ✓ The first simulated return for period $t + 1$ can be obtained as follows:

$$R^*_{t+1} = \hat{m}_{t+1} + e^*_{t+1}, \tag{4.113}$$

where e^*_{t+1} is the first simulated residual for period $t + 1$.

This procedure is repeated B times of length T. Here, B represents the number of bootstrapped samples and T represents the each of bootstrapped sample size. Then, VaR for period $t + 1$ can be forecasted as follows:

$$VaR_{t+1} = \frac{\sum_{b=1}^{B} Quantile\left\{\{R^*_t\}_{t=1}^{T}, 100p\right\}}{B}. \tag{4.114}$$

5. EVALUATION OF VALUE-AT-RISK MODELS

Backtesting methodology is used to decide the best performed VaR model based on the accuracy of VaR forecast. The two well-known and widely used backtests are Unconditional Likelihood Ratio test, introduced by [69] and Conditional Likelihood Ratio test, introduced by [70]. Here, these two backtest methods are given comprehensively. Moreover, the loss functions have also critical importance for model evaluation. Even if all models achieve to pass two backtest methods, we should compare their forecasting errors. The following procedure is used to decide best model in terms of VaR accuracy.

1. The VaR forecasts of all candidate models are obtained.
2. Unconditional and conditional Likelihood Ratio tests are applied to decide which model produce consistent VaR forecasts.
3. The forecasting errors of VaR models achieved to pass stage 2 is compared by means of loss functions.
4. The lowest values of the loss functions represents the best model.

This section is devoted to model comparison tests and metrics of VaR models.

5.1. Unconditional Coverage

The first backtesting method is Likelihood Ratio test of unconditional coverage (LR_{uc}) proposed by Kupiec [69]. The aim of this method is to test the equality of failure rate and its expected value. The LR_{uc} test statistic is given by

$$LR_{uc} = -2\ln\left[\frac{p^{n_1}(1-p)^{n_0}}{\hat{\pi}^{n_1}(1-\pi)^{n_0}}\right] \sim \chi^2_1. \tag{5.1}$$

where $\hat{\pi} = n_1/(n_0+n_1)$ is the MLE of p, n_1 represents the total violation and n_0 represents the total non-violations forecasts. Note that if $VaR_t > r_t$ violation occurs, opposite case indicates the non-violation. Under the null hypothesis ($H_0 : p = \hat{\pi}$), the LR statistic is asymptotically distributed as χ^2 distribution with one degree of freedom.

5.2. Conditional Coverage

Christoffersen [70] proposed a likelihood ratio test of conditional coverage LR_{cc} to remove the lack of Kupiec's [69] test. The LR_{cc} test investigates both the equality of failure rate and expected one and also independently distributed violations. The LR_{cc} test statistic is given by

$$LR_{cc} = -2\ln\left[\frac{(1-p)^{n_0}p^{n_1}}{(1-\pi_{01})^{n_{00}}\pi_{01}{}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}{}^{n_{11}}}\right] \sim \chi_2^2. \tag{5.2}$$

where n_{ij} is the number of observations with value i followed by j for $i, j = 0, 1$ and $\pi_{ij} = n_{ij}/\sum_j n_{ij}$ are the probabilities, $i, j = 1$ denotes that the violation has been occurred, otherwise indicates the opposite case. The LR_{cc} statistic is asymptotically distributed as χ^2 distribution with two degree of freedom.

5.3. Loss Functions

The LR_{uc} annd LR_{cc} backtesting methods provide limited information to compare the VaR models. To decide the best VaR model, the magnitude of violation should be investigated whether the VaR model produce accurate forecasts. For this goal, the loss functions are very useful in decision of the best model. In this section, the most used loss functions are given.

5.3.1. Quadratic Loss Function

The performance evaluation of VaR models by means of LR_{uc} and LR_{cc} test can not provide the sufficient evidence to decide which model produce the most accurate VaR forecasts among others. For instance, some of the models may have the same violation number with different forecast errors.

Lopez [71] proposed a test based on the quadratic loss function (QLF) to take into account differences between realized returns and VaR forecasts. The QLF defined by

Lopez [71] is given by

$$QLF_{t+1} = \begin{cases} 1 + (r_{t+1} - VaR_{t+1})^2, \text{if} & r_{t+1} < VaR_{t+1} \\ 0, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases} \tag{5.3}$$

where r_{t+1} and VaR_{t+1} are the realized return and VaR forecast at time $t+1$, respectively.

5.3.2. Regulator's Loss Function

The another loss function, defined by Sarma et al. [72], is regulator's loss function (RLF). The RLF is given by

$$RLF_{t+1} = \begin{cases} (r_{t+1} - VaR_{t+1})^2, \text{if} & r_{t+1} < VaR_{t+1} \\ 0, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases} \tag{5.4}$$

where r_{t+1} and VaR_{t+1} are the realized return and VaR forecast at time $t+1$, respectively.

5.3.3. Unexpected Loss

The unexpected loss (UL) is equal to average value of differences between realized return and VaR forecasts. The magnitude of the violation is given by

$$\text{UL}_{t+1} = \begin{cases} (r_{t+1} - VaR_{t+1}), \text{if} & r_{t+1} < VaR_{t+1} \\ 0, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases} \tag{5.5}$$

where UL_{t+1} is the one-day-ahead magnitude of the violation at time $t+1$.

5.3.4. Firm's Loss Functions

The QLF and UL loss functions do not consider the case in which the realized returns exceed the VaR forecast. The appropriate loss function should take into consideration the cost of excess capital. Because, overestimated VaR forecasts yield firms to hold much more capital value than required one. The main objective of any firm is to maximize the their profits. For this reason, Sarma et al. [72] proposed the new loss function, called Firm's Loss Function (FLF). The FLF is given by

$$FLF_{t+1} = \begin{cases} (r_{t+1} - VaR_{t+1})^2, & \text{if } \; r_{t+1} < VaR_{t+1} \\ -\beta VaR_{t+1}, & \text{if } \; r_{t+1} \geq VaR_{t+1} \end{cases} \tag{5.6}$$

where β is the cost of excess capital. The another loss function was proposed by Abad et al. [73] in the same way with Sarma et al. [72] The loss function, proposed by Abad et al. [73] is given by

$$FABL_{t+1} = \begin{cases} (r_{t+1} - VaR_{t+1})^2, & \text{if } \; r_{t+1} < VaR_{t+1} \\ -\beta \left(r_{t+1} - VaR_{t+1}\right), & \text{if } \; r_{t+1} \geq VaR_{t+1} \end{cases} \tag{5.7}$$

where β is the cost of excess capital.

6. EMPIRICAL RESULTS

In this section, the usefulness of proposed VaR models are demonstrated by means of five applications to real data sets. The first four application is related to published and/or under review papers given in Section 1. The comprehensive empirical study is given in fifth application contains all the models introduced in Section 4. The results of backtesting and loss functions are used to decide the best performed model.

6.1. Application of ASGT Model

In this section, empirical results of ASGT study are presented.

6.1.1. Data Description of ASGT Model

The major index of Turkey, ISE-100 index, is used to demonstrate the forecasting performance of GARCH-ASGT model in terms of accuracy of VaR. The used time series data contains 1278 daily observations from 09.02.2012 to 07.03.2017. The descriptive statistics of the log-returns of ISE-100 index are given in Table 3. Figure 12 displays the daily log-returns of ISE-100 index and histogram of daily log-returns. The p-values of skewness and kurtosis measures are obtained by the hypothesis tests of D'Agostino [74] and Anscombe and Glynn [75].

Table 3: Summary statistics of ISE-100 index (09.02.2012-07.03.2017)

ISE-100					
Number of observations	Minimum	Maximum	Mean	Median	Std. Deviation
1278	-0.1106	0.0623	0.0003	0.0007	0.0140
Skewness	Kurtosis	Jarque-Bera			
-0.5930 (p=<0.0001)	5.0700 (p=<0.0001)	1450.7610 (p=<0.0001)			

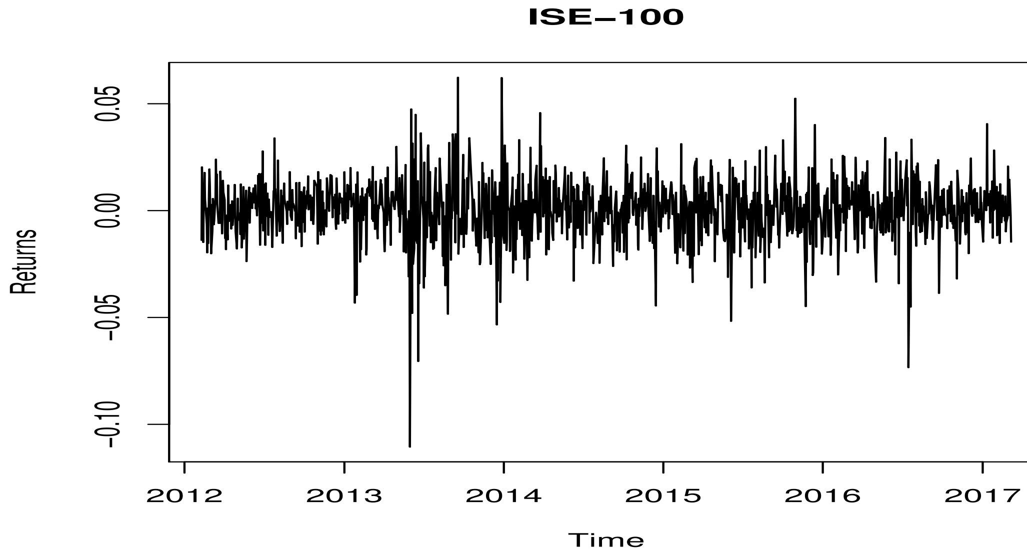


Figure 12: Daily log-returns of ISE-100 index

Table 3 shows that the mean return is close to 0. The Jarque-Bera (JB) test shows that the ISE-100 index is not normally distributed. It can be viewed as an evidence for significant skewness and excess kurtosis. Thus, it is clear that log return of ISE-100 index has non-normal characteristics, excess kurtosis, and fat tails.

The benchmark model, GARCH(1,1), defined in Equation (4.1), is estimated with six different innovation distributions: normal, Student-t, GED, GT, SGT and ASGT distributions. Table 4 shows the estimated parameters of GARCH models. The **rugarch** package in R is used to obtain parameter estimation of GARCH-N, GARCH-T and GARCH-GED models and **constrOptim** function in R is used to minimize negative ll function of GARCH-GT, GARCH-SGT and GARCH-ASGT models.

Table 4: Estimated parameters of GARCH(1,1) models for ISE-100 index, the corresponding SEs are in parenthesis and p-values in $[\cdot]$.

Parameters	ISE-100					
	Normal	Student-t	GED	GT	SGT	ASGT
m	7.29×10^{-4}	8.70×10^{-4}	8.53×10^{-4}	9×10^{-4}	9.58×10^{-4}	6.19×10^{-4}
	(3.68×10^{-4})	(3.34×10^{-4})	(3.29×10^{-4})	(3.32×10^{-4})	(3.70×10^{-4})	(3.60×10^{-4})
	[0.048]	[0.009]	[0.007]	[0.007]	[0.010]	[0.031]
ω	1.06×10^{-5}	2.99×10^{-6}	5.88×10^{-6}	1×10^{-5}	9.92×10^{-6}	3.03×10^{-6}
	(4.12×10^{-6})	(1.91×10^{-6})	(4.59×10^{-6})	(8.48×10^{-6})	(1.35×10^{-6})	(2.19×10^{-6})
	[0.012]	[0.021]	[0.341]	[0.238]	[<0.001]	[0.166]
γ_1	0.064	0.032	0.045	0.123	0.123	0.033
	(0.017)	(0.019)	(0.02)	(0.084)	(0.052)	(0.012)
	[<0.001]	[<0.001]	[0.088]	[0.143]	[0.018]	[0.006]
γ_2	0.881	0.951	0.923	0.856	0.8562731	0.951
	(0.033)	(0.018)	(0.042)	(0.083)	(0.031)	(0.021)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
υ	-	5.569	-	1.642	1.706	2.448
		(0.803)		(0.476)	(0.510)	(0.859)
		[<0.001]		[<0.001]	[<0.001]	[0.004]
κ	-	-	1.279	2.303	2.271974	2.155
			(0.063)	(0.379)	(0.370)	(0.289)
			[<0.001]	[<0.001]	[<0.001]	[<0.001]
α	-	-	-	-	-	0.157
						(0.011)
						[<0.001]
η	-	-	-	-	-0.045	-
					(0.021)	
					[0.031]	
$-\ell$	-3683.52	-3740.26	-3730.03	-3732.78	-3733.498	-3742.12

Based on the results given in Table 4, GARCH-ASGT model has the lowest ll value among others and exhibits superior fit to standardized residuals. The γ_1 and γ_2 parameters are statistically significant. Moreover, skewness parameters of SGT and ASGT distributions are found statistically significant. It is an evidence that the distribution of ε_t exhibits skewness and kurtosis. Therefore, it is easy to conclude that the normal distribution is unsuitable for innovation process of GARCH models. Figure 13 displays the estimated densities of innovation distributions for standardized residuals.

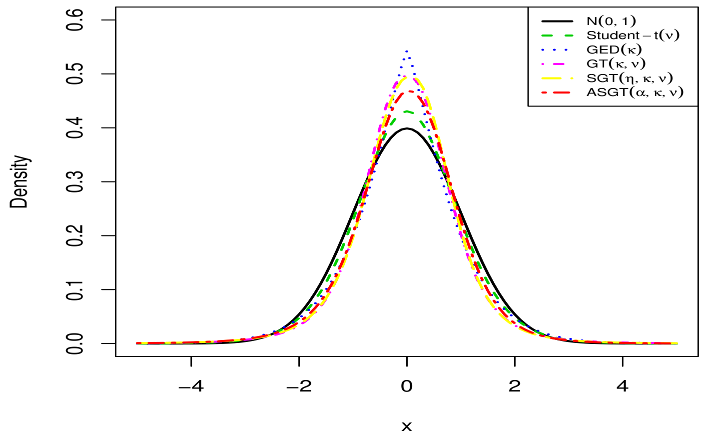


Figure 13: Estimated densities for standardized residuals

6.1.2. Backtesting Results of ASGT Model

In this subsection, rolling window estimation procedure is used to compare the out-of-sample performance of GARCH models in terms of VaR accuracy. Rolling window estimation produce allows us to capture time-varying characteristics of the time series in different time periods. Window length is determined as 678 and next 600 daily returns are used to evaluate the out-of-sample performance of VaR models.

Table 5 and 6 show the bactesting results of VaR models. All used VaR models are evaluated by comparing the failure rates and using four backtesting results. According to failure rates, GARCH models specified under the normal, Student-t and GED innovation distributions perform poorly for both confidence levels, whereas GARCH-GT, GARCH-SGT and GARCH-ASGT outperform to other models and shows great consistency for both confidence levels. Since normal, Student-t and GED distributions are symmetric about the zero, these distributions fail to capture skewness in financial returns.

Table 5: Backtesting results of VaR models for 97.5% confidence level ($p = 0.025$)

Models	Mean VaR (%)	Failure rate	LR_{uc}	LR_{cc}	AQLF	UL
GARCH-N	-2.696	0.032	1.010 (0.314)	2.255 (0.323)	0.10865	-0.0353
GARCH-Student-t	-2.701	0.032	1.010 (0.314)	2.255 (0.323)	0.10501	-0.0337
GARCH-GED	-2.793	0.028	0.262 (0.608)	1.255 (0.533)	0.09592	-0.0312
GARCH-GT	-2.930	0.025	0 (1)	0.771(0.680)	0.08290	-0.0274
GARCH-SGT	-2.933	0.025	0 (1)	0.771(0.680)	0.08301	-0.0273
GARCH-ASGT	-2.966	0.023	0.069 (0.791)	0.740 (0.691)	0.08247	-0.0272

p-values of LR_{uc} and LR_{cc} tests are presented in parentheses.

GARCH-GT, GARCH-SGT and GARCH-ASGT models produce similar results in the basis of failure rates. To understand the differences between these three models deeply, LR_{uc}, LR_{cc}, average QLF ($AQLF$) and UL are used. Based on the LR_{cc}, $AQLF$ and UL criteria, GARCH-ASGT model outperforms to GARCH-GT and GARCH-SGT models for both confidence levels. GARCH-GT and GARCH-SGT models outperform to GARCH-ASGT model only in terms of LR_{uc} results for 97.5% confidence level. Consequently, GARCH-ASGT model generates the most reliable VaR forecasts among others for both confidence levels on the basis of backtesting results. Moreover, it is clear that the VaR forecasts of GARCH-ASGT model has the lowest forecast errors among others. This fact reveals that the ASGT innovation distribution provides the superior fits to conditional return series and enables to model both skewness and kurtosis in financial return series.

Table 6: Backtesting results of VaR models for 99% confidence level ($p = 0.01$)

Models	Mean VaR (%)	Failure rate	LR_{uc}	LR_{cc}	AQLF	UL
GARCH-N	-3.212	0.022	6.185 (0.012)	6.762 (0.034)	0.07042	-0.02256
GARCH-Student-t	-3.496	0.017	2.243 (0.134)	2.583(0.274)	0.05376	-0.01625
GARCH-GED	-3.517	0.017	2.243 (0.134)	2.583(0.274)	0.05284	-0.01585
GARCH-GT	-3.798	0.012	0.159 (0.689)	0.325 (0.849)	0.03943	-0.01280
GARCH-SGT	-3.804	0.012	0.159 (0.689)	0.325 (0.849)	0.03870	-0.01266
GARCH-ASGT	-3.846	0.010	0 (1)	0.121 (0.941)	0.03850	-0.01237

p values of LR_{uc} and LR_{cc}tests are presented in parentheses.

Figures 14 and 15 display the VaR forecasts of models for both confidence levels. Based on these plots and backtesting results, it is clear that GARCH-ASGT model outperforms among others especially for high quantiles.

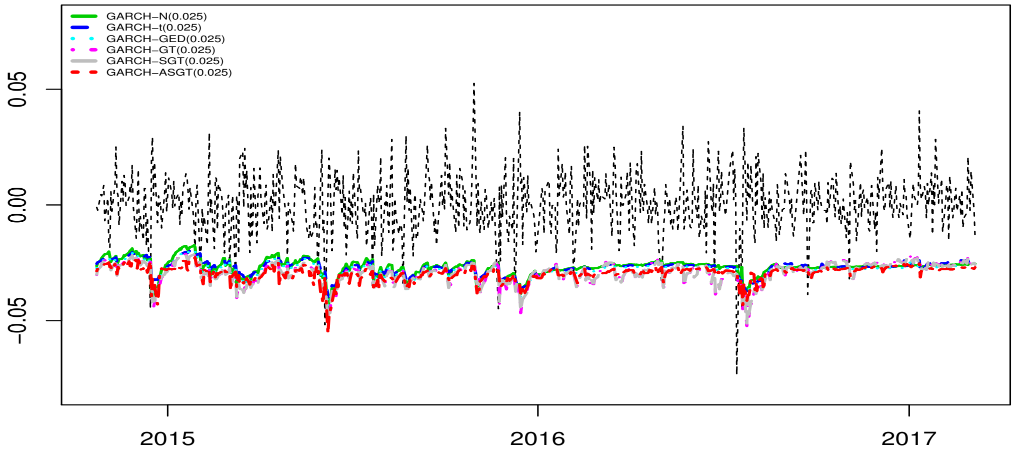


Figure 14: Daily VaR forecast of GARCH models with different innovation distributions for 97.5% confidence level

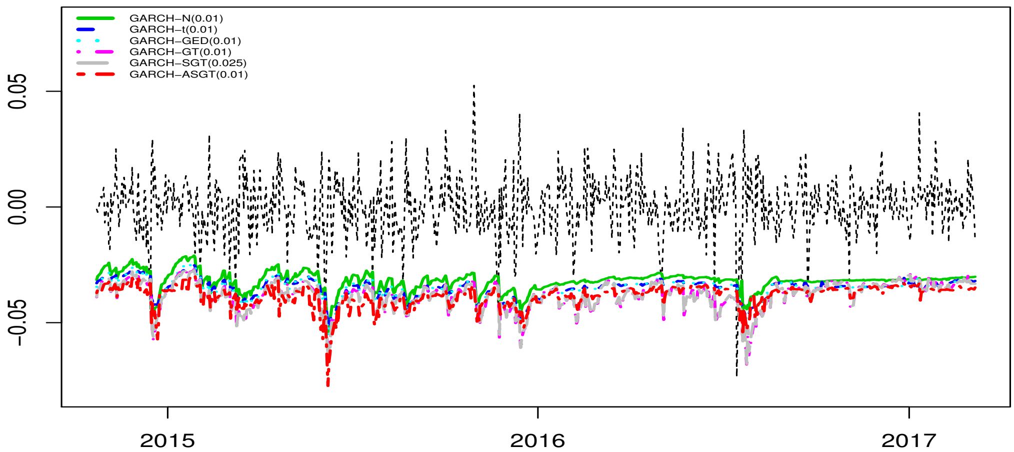


Figure 15: Daily VaR forecast of GARCH models with different innovation distributions for 99% confidence level

6.2. Application of GAST model

In this section, empirical results of GAST study are presented.

6.2.1. Data Description of GAST Model

In this section, S&P-500 index is used. The used time series data contains 792 daily log-returns from 03.01.2014 to 24.02.2017. The descriptive statistics of the log-returns of S&P-500 index are given in Table 7. Figure 16 displays the daily log-returns of S&P-500 index.

Table 7: Summary statistics of S&P-500 index

SP-500	
Number of observations	792
Minimum	-0.0402
Maximum	0.0382
Mean	0.0003
Median	0.0003
Std. Deviation	0.0082
Skewness	-0.3736
Kurtosis	5.4270
Jarque-Bera	212.9143 ($p<0.001$)

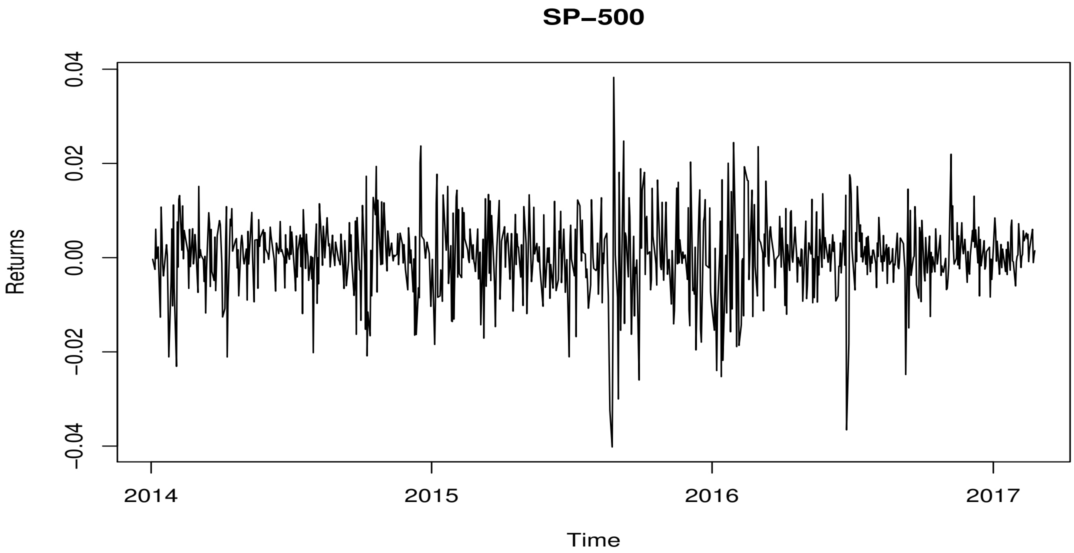


Figure 16: Daily log-returns of S&P-500 index

Table 7 shows that the mean return is closed to 0. The JB test shows that the ISE-100 index is not normally distributed. It can be viewed as an evidence for significant skewness and excess kurtosis. Thus, it is clear that log return of S&P-500 index has non-normal characteristics, excess kurtosis, and fat tails.

The benchmark models, GARCH(1,1) and GJR-GARCH(1,1), are estimated with four different innovation distributions: Normal, Student-t, ST, and GAST. Table 8 shows the estimated parameters of models specified under four different innovation distributions. The **rugarch** package in R software is used to obtain parameter estimation of normal and Student-t models. The **constrOptim** function in R software is used to minimize negative ll function of GARCH-ST, GJR-GARCH-ST, GARCH-GAST and GJR-GARCH-GAST models.

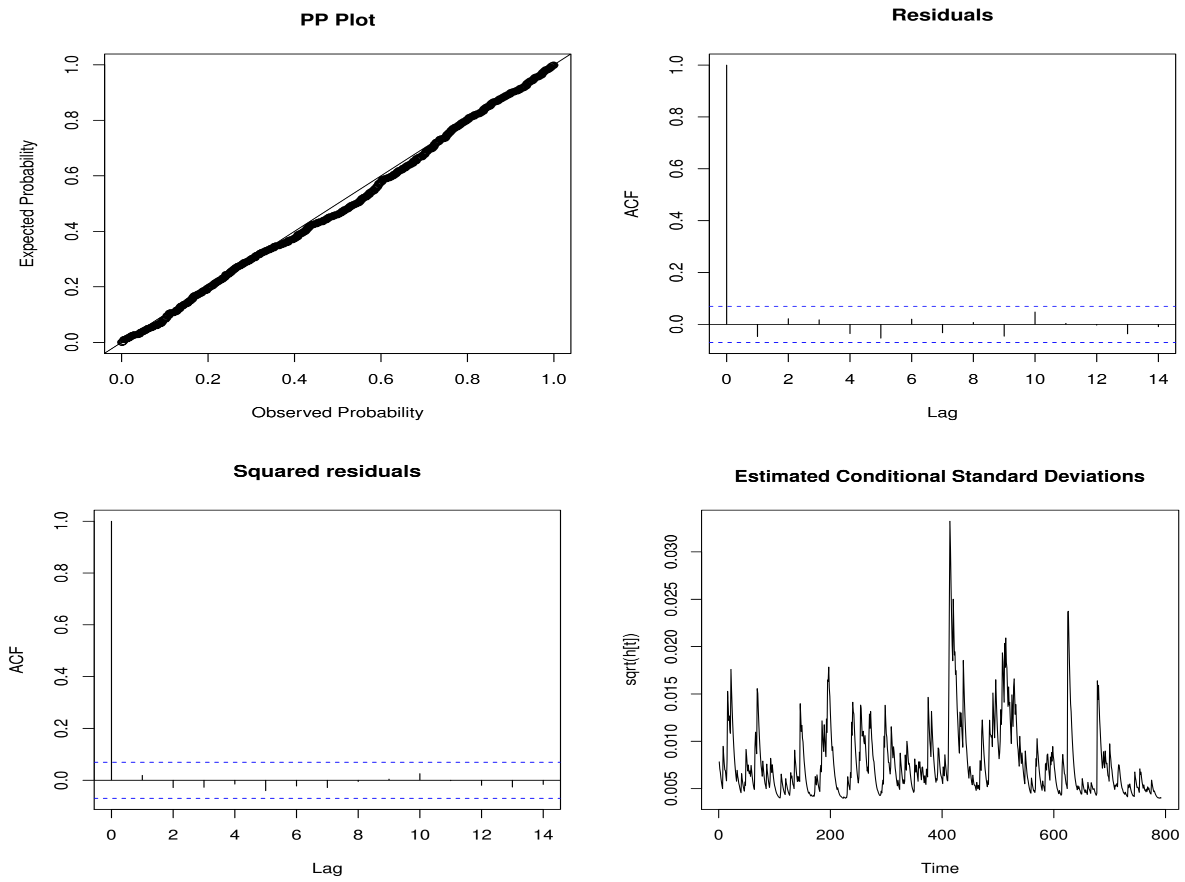


Figure 17: Diagnostic plots of GJR-GARCH-GAST model

Table 8 shows that GJR-GARCH-GAST model has the lowest ll value among others and therefore, GJR-GARCH-GAST model could be chosen as the best model for the dataset. Based on the estimated parameters, the γ_2 and γ_3 are statistically significant. Therefore, it is clear that bad news have more impact on volatility than good news. Figure 17 displays the diagnostic plots of GJR-GARCH-GAST model. Figure 17 proves that GJR-GARCH-GAST model is statistically valid and exhibits superior fits to innovations of GJR-GARCH(1,1) models.

Table 8: Estimated parameters of GARCH(1,1) and GJR-GARCH(1,1) models for S&P-500 index, the corresponding standard errors are given in second line.

Parameters	GARCH				GJR-GARCH			
	Normal	Student-t	ST	GAST	Normal	Student-t	ST	GAST
m	5.88×10^{-4}	7.15×10^{-4}	7.07×10^{-4}	3.98×10^{-4}	2.22×10^{-4}	4.20×10^{-4}	4.21×10^{-4}	4.20×10^{-4}
	2.36×10^{-4}	2.14×10^{-4}	2.18×10^{-4}	2.23×10^{-4}	2.29×10^{-4}	2.02×10^{-4}	2.13×10^{-4}	2.13×10^{-4}
ω	6.52×10^{-6}	4.29×10^{-6}	4.90×10^{-6}	4.15×10^{-6}	5.00×10^{-6}	4.00×10^{-6}	3.51×10^{-6}	3.51×10^{-6}
	1.65×10^{-6}	1.46×10^{-6}	1.61×10^{-6}	1.35×10^{-6}	2.00×10^{-6}	1.00×10^{-6}	9.13×10^{-7}	9.35×10^{-7}
γ_1	0.204	0.218	0.240	0.233	1×10^{-7}	1×10^{-7}	1.70×10^{-7}	6.11×10^{-8}
	0.038	0.046	0.047	0.046	2.68×10^{-4}	1.1×10^{-5}	1.05×10^{-5}	2.01×10^{-5}
γ_2	0.704	0.739	0.697	0.725	0.759	0.759	0.769	0.768
	0.048	0.047	0.053	0.047	0.041	0.035	0.032	0.032
γ_3	-	-	-	-	0.330	0.402	0.364	0.372
	-	-	-	-	0.084	0.077	0.064	0.070
ν	-	5.754	7.261	7.407	-	6.455	8.031	8.446
	-	1.193	1.792	1.929	-	1.687	0.332	2.143
λ	-	-	-0.820	-0.468	-	-	-1.022	-0.673
	-	-	0.290	0.053	-	-	0.242	0.049
α	-	-	-	0.211	-	-	-	0.222
	-	-	-	0.029	-	-	-	0.029
$-\ell$	-2751.69	-2773.020	-2776.720	**-2779.34**	-2770.02	-2794.36	-2797.03	**-2801.04**

6.2.2. Backtesting Results of GAST Model

Rolling window estimation procedure is used to estimate parameters of GARCH models. Window length is determined as 292 and next 500 daily returns are used to compare the performance of VaR models.

Table 9 shows the backtesting results for GARCH(1,1)-N, GJR-GARCH(1,1)-N, GARCH(1,1)-T,GJR-GARCH(1,1)-T, GARCH(1,1)-ST, GJR-GARCH(1,1)-ST, GARCH(1,1)-GAST and GJR-GARCH(1,1)-GAST models for long position. Table 10 shows the backtesting results of VaR models for the short position. All used VaR models are evaluated by comparing the failure rates and using LR_{uc}, LR_{cc}, average RLF (ARLF) and UL backtesting results.

When considered the out-of-sample performance of VaR models for long position, GARCH models specified under normal, Student-t and ST distributions perform poorly at 0.025% level and produce underestimated VaR forecasts on the basis of failure rates and GJR-GARCH-GAST model produces the overestimated VaR forecast at 0.025% level. GARCH-GAST model produces more accurate VaR forecast and has the minimum $ARLF$ and UL values among others. Consequently, GARCH-GAST model could be chosen the best model for 0.025% level.

GJR-GARCH-ST, GJR-GARCH-GAST and GARCH-GAST models produce the similar results for 0.01% level on the basis of failure rates. When taking into consideration the backtesting results of these models, it is clear that GARCH-GAST and GJR-GARCH-GAST models, respectively, have the minimum $ARLF$ and UL values. Therefore, GJR-GARCH-GAST model could be chosen as the best model for 0.01% level.

Table 9: Backtesting results of VaR models for long position ($p = 0.025$ and $p = 0.01$)

p=0.025						
Models	Mean VaR	Failure rate	LR_{uc}	LR_{uc}	ARLF	UL
GARCH-N	-1.577	0.042	4.938 (0.026)	6.102 (0.047)	0.0267	-0.0233
GJR-GARCH-N	-1.662	0.032	0.925 (0.336)	1.317 (0.518)	0.0224	-0.0162
GARCH-T	-1.617	0.038	2.998 (0.083)	4.698 (0.095)	0.0247	-0.0212
GJR-GARCH-T	-1.582	0.040	3.916 (0.047)	3.965 (0.137)	0.0283	-0.0201
GARCH-ST	-1.725	0.032	0.924 (0.336)	1.317 (0.517)	0.0220	-0.0182
GJR-GARCH-ST	-1.722	0.034	1.496 (0.221)	1.768 (0.412)	0.0251	-0.0160
GARCH-GAST	-1.854	**0.022**	0.192 (0.661)	1.621 (0.444)	**0.0175**	-0.0139
GJR-GARCH-GAST	-1.811	0.016	1.901(0.167)	2.166(0.339)	0.0189	**-0.0121**
p=0.01						
Models	Mean VaR	Failure rate	LR_{uc}	LR_{uc}	ARLF	UL
GARCH-N	-1.881	0.024	7.111 (0.008)	8.263 (0.016)	0.0169	-0.0141
GJR-GARCH-N	-1.974	0.016	1.538 (0.214)	1.798 (0.406)	0.0166	-0.0114
GARCH-T	-2.045	0.014	0.718 (0.396)	3.804 (0.149)	0.0140	-0.0115
GJR-GARCH-T	-2.000	0.020	3.913 (0.048)	4.322 (0.115)	0.0206	-0.0116
GARCH-ST	-2.171	0.014	0.718 (0.396)	3.804 (0.149)	0.0128	-0.0104
GJR-GARCH-ST	-2.168	**0.010**	0(1)	0.101 (0.951)	0.0194	-0.0100
GARCH-GAST	-2.306	0.012	0.189 (0.663)	3.901 (0.142)	**0.0103**	-0.0087
GJR-GARCH-GAST	-2.258	**0.010**	0(1)	0.102(0.951)	0.0114	**-0.0082**

p values of LR-uc and LR-cc tests are presented in parentheses.

When considered the out-of-sample performance of VaR models for short position, all models pass LR_{uc} and LR_{cc} tests for both confidence levels. However, GARCH-N, GJR-GARCH-N, GARCH-T, GJR-GARCH-T and GARCH-ST models produce overestimated VaR forecasts at 97.5% level on the basis of failure rates. GJR-GARCH-ST and GARCH-GAST models produce more accurate VaR forecast at 97.5% level. Since GARCH-GAST has the minimum $ARLF$ and UL values among others, GARCH-GAST model could be chosen as the best model for 97.5% level.

All VaR models produce similar results at 99% level on the basis of failure rates, LR_{uc} and LR_{cc} tests results. When taking into consideration the $ARLF$ and UL results, it is clear that GARCH-GAST model has the minimum values for these statistics. Therefore, GARCH-GAST model could be chosen as the best model for 99% level.

Table 10: Backtesting results of VaR models for short position ($p = 0.975$ and $p = 0.99$)

p=0.975						
Models	Mean VaR	Failure rate	LR_{uc}	LR_{uc}	ARLF	UL
GARCH-N	1.671	0.984	1.901 (0.167)	2.161 (0.339)	0.0024	0.0047
GJR-GARCH-N	1.678	0.980	0.549 (0.458)	0.958 (0.619)	0.0031	0.0057
GARCH-T	1.729	0.988	4.278 (0.038)	4.424 (0.109)	0.0028	0.0043
GJR-GARCH-T	1.696	0.980	0.549 (0.458)	0.958 (0.619)	0.0030	0.0049
GARCH-ST	1.602	0.982	1.111 (0.291)	1.442 (0.486)	0.0031	0.0053
GJR-GARCH-ST	1.604	**0.972**	0.177 (0.673)	0.986 (0.611)	0.0053	0.0086
GARCH-GAST	1.614	**0.978**	0.192 (0.661)	0.688 (0.708)	**0.0023**	**0.0042**
GJR-GARCH-GAST	1.537	0.966	1.496 (0.221)	1.768 (0.419)	0.0037	0.0079
p=0.99						
Models	Mean VaR	Failure rate	LR_{uc}	LR_{uc}	ARLF	UL
GARCH-N	1.975	0.996	2.352 (0.125)	2.369 (0.305)	0.0010	0.0016
GJR-GARCH-N	1.989	**0.994**	0.943 (0.331)	0.979 (0.613)	0.0011	0.0023
GARCH-T	2.157	0.996	2.352 (0.125)	2.369 (0.305)	0.0010	0.0015
GJR-GARCH-T	2.113	**0.994**	0.943 (0.331)	0.979 (0.613)	0.0011	0.0018
GARCH-ST	1.950	0.996	2.352 (0.125)	2.369 (0.305)	0.0012	0.0017
GJR-GARCH-ST	1.793	0.986	0.718 (0.396)	0.917 (0.632	0.0023	0.0043
GARCH-GAST	1.912	**0.994**	0.943 (0.331)	0.979 (0.613)	**0.0007**	**0.0013**
GJR-GARCH-GAST	1.787	0.986	0.718(0.396)	0.917(0.632)	0.0009	0.0026

p values of LR-uc and LR-cc tests are presented in parentheses.

Figures 18 and 19 display the VaR forecasts of the models for long and short positions. Based on the Figures 18 and 19, it is concluded that the GARCH models with GAST innovation distribution exhibit great consistency for estimating the true quantile value of innovation distribution.

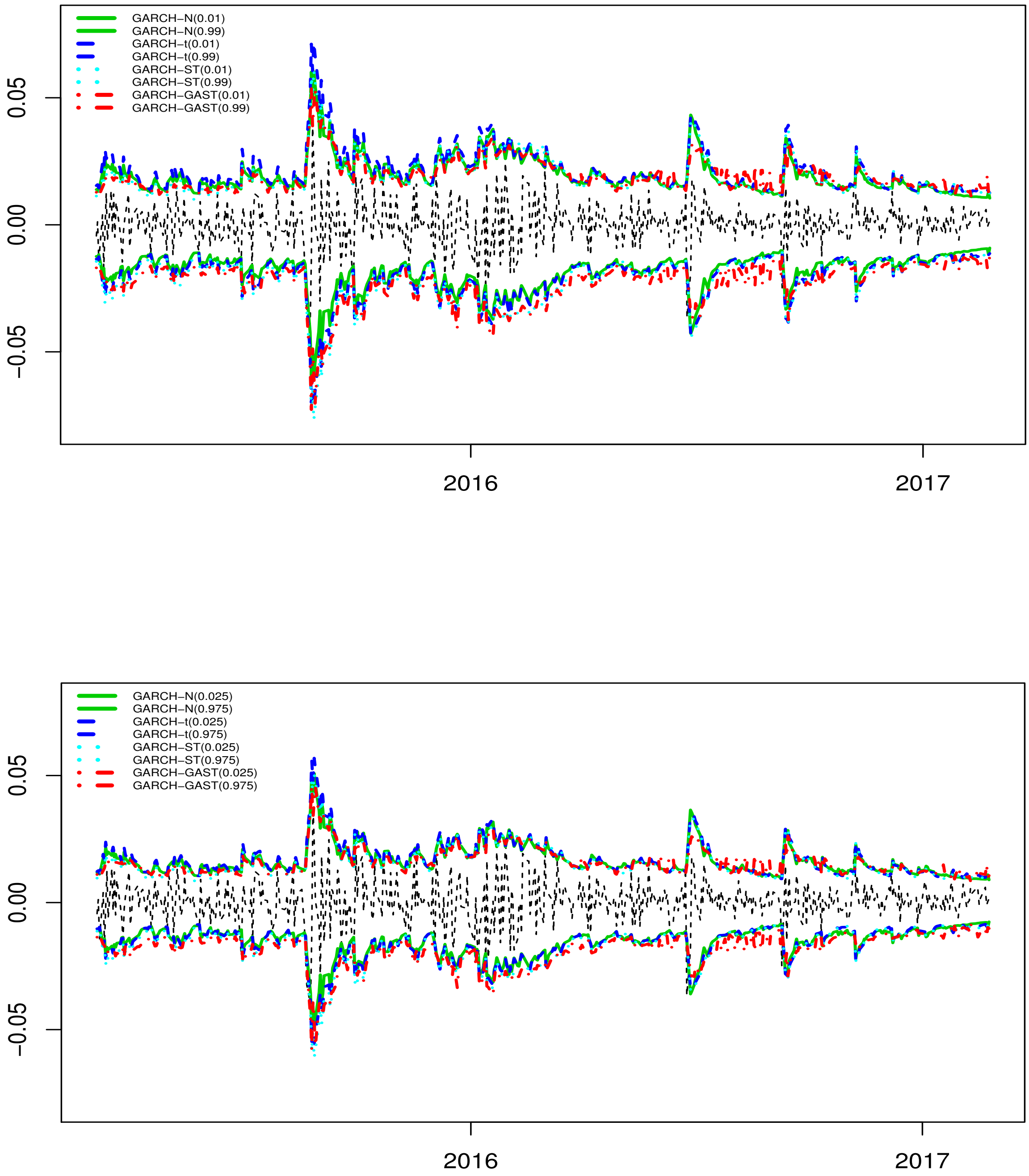


Figure 18: Daily VaR forecast of GARCH models with different innovation distributions for 97.5% and 99% confidence levels

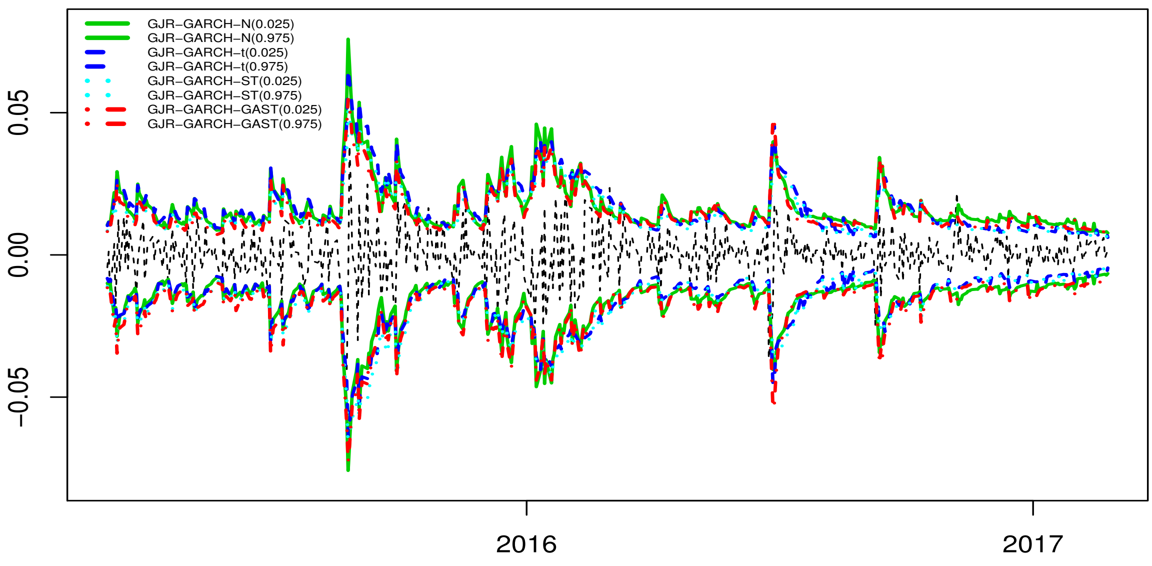


Figure 19: Daily VaR forecast of GJR-GARCH models with different innovation distributions for 97.5% and 99% confidence levels

6.3. Application of FHS Model

In this section, empirical results of FHS study are presented. The goal of this empirical study is to show whether the distributional assumption on innovation process effects the accuracy of VaR forecast for FHS model.

6.3.1. Data description of FHS Model

ISE-100 index of Turkey is used to evaluate the performance of FHS models in terms of accuracy of VaR forecasts. The time series data contains 1092 daily log-returns from 03.01.2013 to 04.05.2017. The descriptive statistics of the log-returns of ISE-100 index are given in Table 11.

Table 11: Summary statistics of ISE-100 index (03.01.2013-04.05.2017)

ISE-100	
Number of observations	1092
Minimum	-0.048
Maximum	0.027
Mean	6.6×10^{-5}
Median	2×10^{-4}
Std. Deviation	0.006
Skewness	-0.603
Kurtosis	4.957
Jarque-Bera	1190.970 (p <0.001)

Table 11 shows that the mean return is close to 0. The JB test shows that the ISE-100 index is not normally distributed. It can be viewed as an evidence for significant skewness and excess kurtosis. Thus, it is clear that log return of ISE-100 index has non-normal characteristics, excess kurtosis, and fat tails. Figure 20 displays the daily log-returns of ISE-100 index.

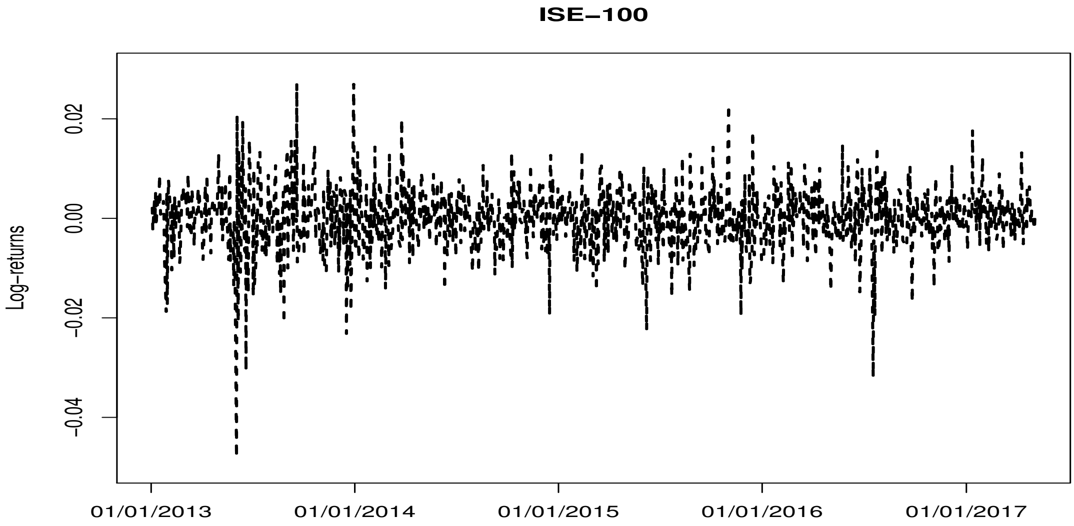


Figure 20: Daily log-returns of ISE-100 index

Figure 21 displays the time-varying skewness and kurtosis of ISE-100. For Figure 21, window length is determined as 392 and rolling window procedure is used. Based on Figure 21, it is clear that skewness and kurtosis of ISE-100 index exhibit great variability across the time.

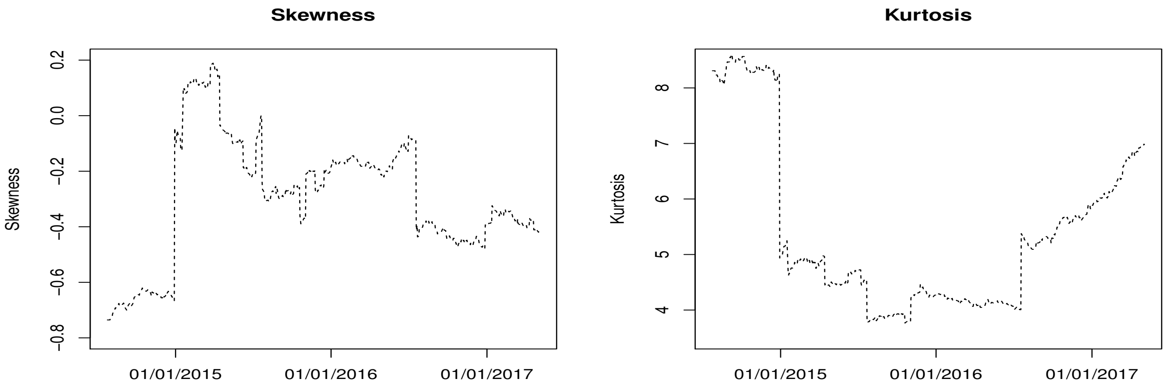


Figure 21: Time varying skewness and kurtosis plots of ISE-100 index

The benchmark model, GARCH(1,1), is estimated with six different innovation distributions: Normal, SN, Student-t, ST, GED and SGED. Table 12 lists the estimated parameters of GARCH models. The **rugarch** package in R software is used to obtain parameter estimation of normal, Student-t, GED and SGED models. The **constrOptim** function in R software is used to minimize negative ll functions of GARCH-ST and GARCH-SN models.

Table 12: In-sample performance of GARCH models under skewed and fat-tailed innovation distributions

Parameters	Normal	Student-t	ST	SN	GED	SGED
m	5.24×10^{-4}	5.67×10^{-4}	8.56×10^{-4}	4.05×10^{-4}	5.67×10^{-4}	3.51×10^{-4}
	3.41×10^{-4}	3.04×10^{-4}	3.60×10^{-4}	3.34×10^{-4}	3.77×10^{-4}	3.26×10^{-4}
ω	3.69×10^{-6}	2.01×10^{-6}	2.27×10^{-6}	3.65×10^{-6}	2.72×10^{-6}	2.63×10^{-6}
	1.50×10^{-6}	1.53×10^{-6}	6.71×10^{-6}	1.73×10^{-6}	1.66×10^{-6}	1.57×10^{-6}
γ_1	0.1194	0.0759	0.1510	0.1460	0.0930	0.0873
	0.0413	0.0336	0.2940	0.0560	0.0393	0.0353
γ_2	0.8234	0.8908	0.8240	0.7950	0.8600	0.8650
	0.0449	0.0485	0.3130	0.0662	0.0518	0.0486
ν	-	4.7490	4.8760	-	1.2020	-
	-	1.1600	0.5620	-	0.1040	-
λ	-	-	-0.2750	-1.5050	-	0.8860
	-	-	2.4260	0.2630	-	0.0440
κ	-	-	-	-	-	1.2200
	-	-	-	-	-	0.1093
$-\ell$	-1381.110	-1405.003	-1402.263	-1388.180	-1401.060	-1402.730

Based on figures in Table 12, we conclude that GARCH-T and GARCH-SGED models have the lower ll value among others. Since GARCH-T model has the lowest ll value, it could be chosen as the best model for in-sample period. Moreover, the parameters γ_1 γ_2 are highly significant for all GARCH models.

6.3.2. Backtesting Results of FHS Model

In this subsection, rolling window estimation procedure is used to estimate parameters of GARCH models. Then, VaR forecasts of FHS models are obtained by using estimated parameters of GARCH models, one-day-ahead forecasts of conditional mean and conditional variance and standardized residuals extracted from estimated GARCH models. Window length is determined as 392 and next 700 daily returns are used to compare the performance of VaR models.

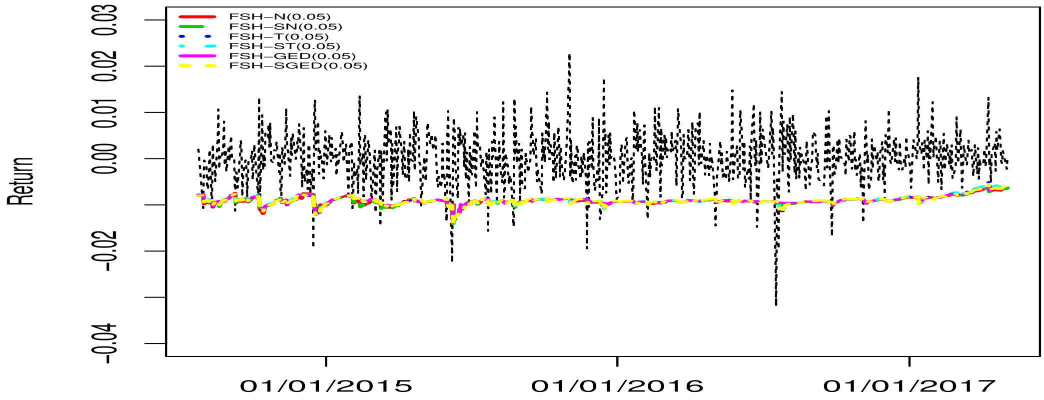
Table 13 shows the backtesting results for FHS-N, FHS-T, FHS-ST, FHS-SN, FHS-GED and FHS-SGED models. LR_{uc}, LR_{cc}, average RLF (ARLF) and UL results are used to decide the best performed FHS model on the basis of accuracy of VaR forecasts. Table 13 indicates that all FHS models perform well based on the results of LR_{uc}, LR_{cc} tests. On the other hand, even if FHS models have similar results in view of LR_{uc}, LR_{cc}, they have different failure rates and forecast errors. The failure rates of FHS-T and FHS-ST at $p = 0.05$ and FHS-GED model at $p = 0.01$ are more close to nominal values. Moreover, loss functions are useful to compare VaR models with their forecast errors. Based on the ARLF and UL results, we conclude following results: (i) FHS-SN is the best performed model at $p = 0.05$ and $p = 0.025$ levels; (ii) FHS model with GED innovation distribution provides the most accurate VaR forecasts among other at $p = 0.01$ level.

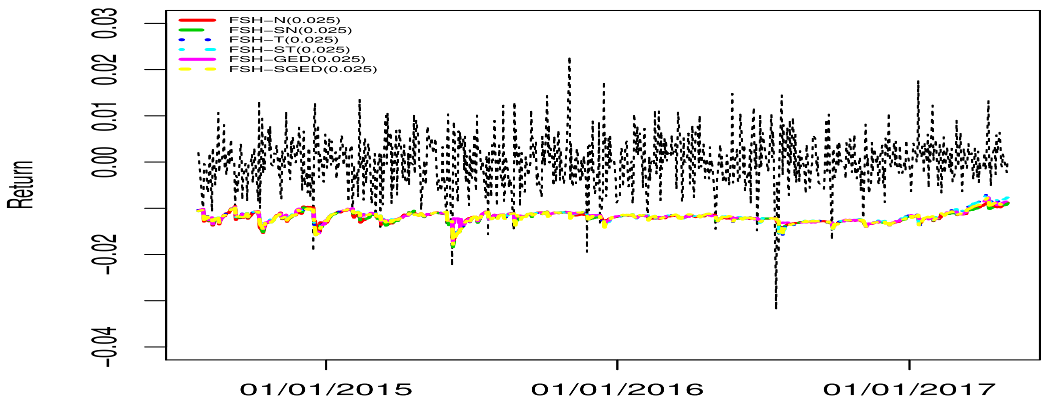
Figures 22 displays the VaR forecasts of FHS models specified under six innovation distributions. As seen in Figure 22, the assumption on innovation process does not affect the VaR forecasts of FHS model soulfully. However, the GED could be preferable to reduce the forecast error of the FHS model, especially at high quantiles.

Table 13: Backtesting results of VaR models for long position ($p = 0.05$, $p = 0.025$, and $p = 0.01$)

p=0.05							
Models	Mean VaR (%)	N. Of Vio.	Failure rate	LR-uc	LR-cc	ARLF	UL
FHS-N	-0.910	29	0.041	1.146 (0.284)	1.186 (0.552)	0.0172063	-0.0179110
FHS-SN	-0.911	29	0.041	1.146 (0.284)	1.186 (0.552)	**0.0171900**	**-0.0178643**
FHS-T	-0.897	32	**0.046**	0.278 (0.597)	0.459 (0.794)	0.0173740	-0.0181385
FHS-GED	-0.899	30	0.043	0.788 (0.374)	0.863 (0.649)	0.0172983	-0.0180858
FHS-SGED	-0.904	30	0.043	0.788 (0.374)	0.863 (0.649)	0.0173162	-0.0179867
FHS-ST	-0.897	32	**0.046**	0.278 (0.597)	0.459 (0.794)	0.0173437	-0.0181591
p=0.025							
Models	Mean VaR (%)	N. Of Vio.	Failure rate	LR-uc	LR-cc	ARLF	UL
FHS-N	-1.193	20	0.029	0.350 (0.554)	0.630 (0.729)	0.0097984	-0.0108364
FHS-SN	-1.196	20	0.029	0.350 (0.554)	0.630 (0.729)	**0.0097622**	**-0.0107498**
FHS-T	-1.177	20	0.029	0.350 (0.554)	0.630 (0.729)	0.0098350	-0.0108600
FHS-GED	-1.179	20	0.029	0.350 (0.554)	0.630 (0.729)	0.0097898	-0.0108694
FHS-SGED	-1.187	20	0.029	0.350 (0.554)	0.630 (0.729)	0.0098099	-0.0107848
FHS-ST	-1.177	20	0.029	0.350 (0.554)	0.630 (0.729)	0.0098287	-0.0108752
p=0.01							
Models	Mean VaR (%)	N. Of Vio.	Failure rate	LR-uc	LR-cc	ARLF	UL
FHS-N	-1.546	9	0.013	0.529 (0.466)	0.764 (0.682)	0.0052475	-0.0051635
FHS-SN	-1.549	9	0.013	0.529 (0.466)	0.764 (0.682)	0.0052483	-0.0051776
FHS-T	-1.526	9	0.013	0.529 (0.466)	0.764 (0.682)	0.0052399	-0.0051979
FHS-GED	-1.530	**8**	0.011	0.137 (0.710)	0.323 (0.851)	**0.0052125**	**-0.0051546**
FHS-SGED	-1.538	9	0.013	0.529 (0.466)	0.764 (0.682)	0.0052148	-0.0051849
FHS-ST	-1.526	9	0.013	0.529 (0.466)	0.764 (0.682)	0.0052135	-0.0051634

p values of LR-uc and LR-cc tests are presented in parentheses.





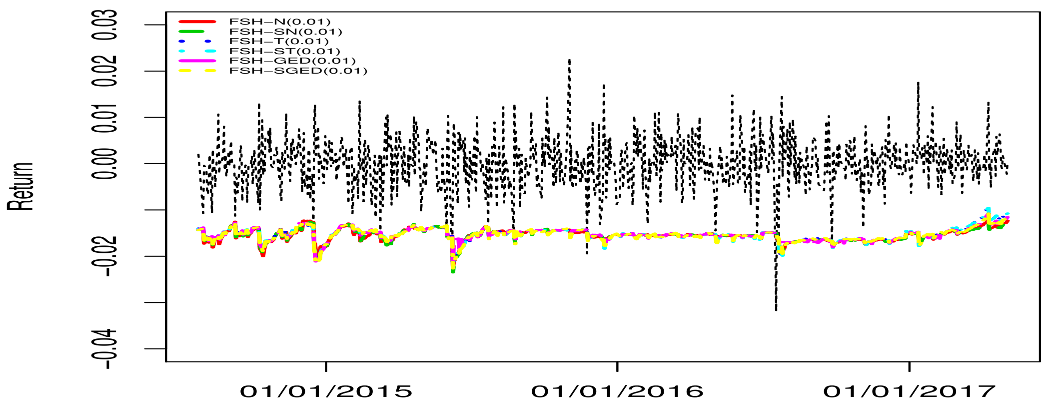


Figure 22: Daily VaR forecast of GARCH models with different innovation distributions for 97.5% and 99% confidence levels

6.4. Application of ASGN Model

In this section, empirical results of ASGN study are presented.

6.4.1. Data Description of ASGN Model

To demonstrate the forecasting performance of GARCH model with ASGN innovation distribution, ISE-100 index is used. The used time series data contains 1278 daily observations from 09.02.2012 to 07.03.2017. Here, the used time series data set is the same with ASGT model. Therefore, the descriptive statistics are omitted and corresponding statistics can be found in Section 6.1.1.

The benchmark model GARCH(1,1), defined in (4.1), is estimated with five different innovation distributions: normal, Student-t, SN, GN and ASGN. Table 14 shows the estimated parameters of GARCH models. The **rugarch** package in R is used to obtain parameter estimation of GARCH-N and GARCH-T models and **constrOptim** function in R is used to minimize negative ll function of GARCH-SN, GARCH-GN and GARCH-ASGN models.

Table 14: Estimated parameters of GARCH(1,1) model for ISE-100 index, the corresponding standard errors are given in second line.

	ISE-100				
Parameters	Normal	Student-t	GN	SN	ASGN
m	7.29×10^{-4}	8.70×10^{-4}	8.53×10^{-4}	5.36×10^{-4}	6.19×10^{-4}
	3.68×10^{-4}	3.34×10^{-4}	3.29×10^{-4}	3.72×10^{-4}	3.63×10^{-4}
	0.048	0.009	0.007	0.149	0.088
ω	1.06×10^{-5}	2.99×10^{-6}	5.88×10^{-6}	9×10^{-6}	6.52×10^{-6}
	4.12×10^{-6}	1.91×10^{-6}	4.59×10^{-6}	4.00×10^{-6}	5.28×10^{-6}
	0.012	0.021	0.341	0.031	0.217
γ_1	0.064	0.032	0.045	0.056	0.047
	0.017	0.019	0.020	0.017	0.022
	<0.001	<0.001	0.088	0.001	0.035
γ_2	0.881	0.951	0.923	0.896	0.918
	0.033	0.018	0.042	0.036	0.048
	<0.001	<0.001	<0.001	<0.001	<0.001
υ	-	5.569	-	-	
		0.803			
		<0.001			
κ	-	-	1.279	-	1.287
			0.063		0.061
			<0.001		<0.001
α	-	-	-	-	0.163
					0.076
					<0.001
λ	-	-	-	0.891	-
				0.029	
				<0.001	
$-\ell$	-3683.52	-3740.26	-3730.03	-3689.94	-3741.05

As seen in Table 14, GARCH-ASGN model has the lowest ll value among others and exhibits superior fit to standardized residuals. The parameters γ_1 and γ_2 are statistically significant. Figure 23 displays the diagnostic plots of GARCH-ASGN model. These figures reveal that standardized ASGN distribution provides the superior fits to standardized residuals of GARCH(1,1) model.

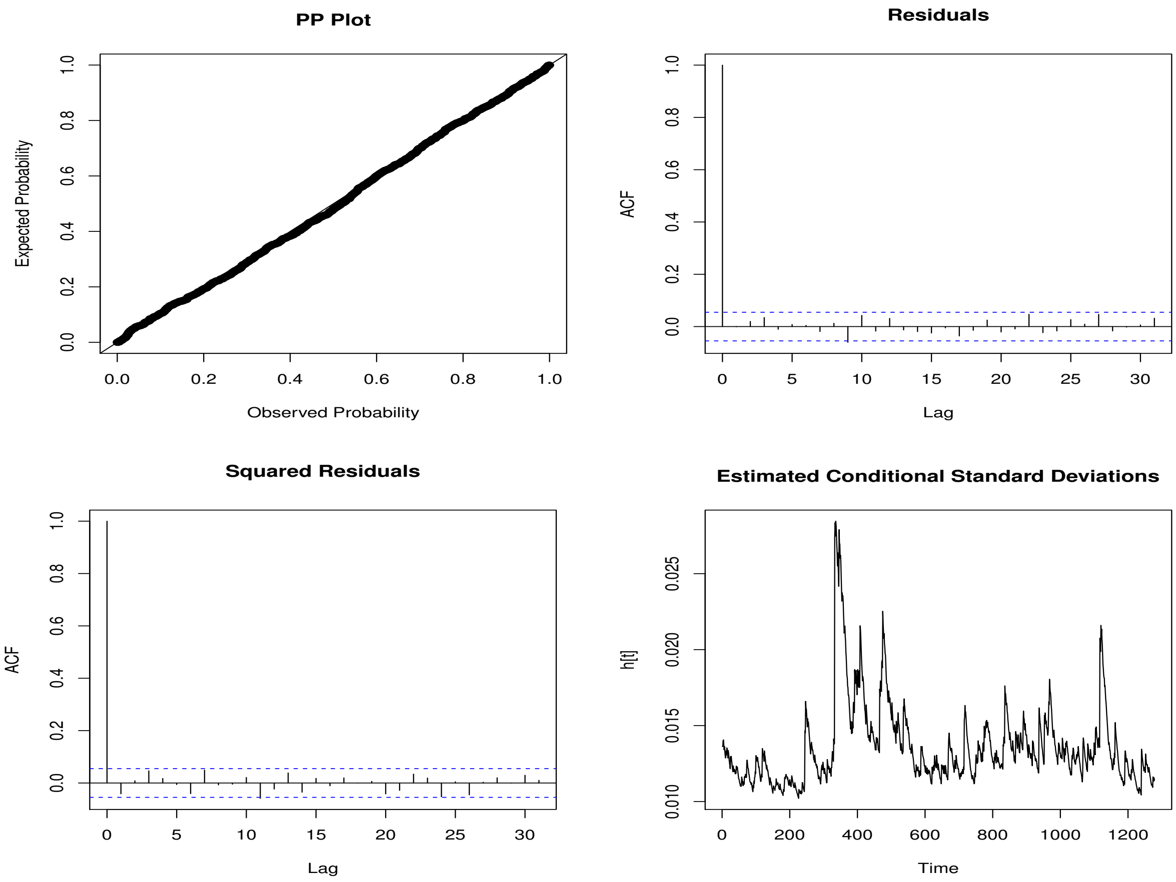


Figure 23: Diagnostic plots of GARCH-ASGN model

6.4.2. Backtesting Results of ASGN Model

Rolling window estimation procedure is used to estimate parameters of GARCH models specified under different innovation distributions. Window length is determined as 678 and next 600 daily log-returns are used to compare the performance of VaR models. The out-of-sample performance of VaR models are evaluated by comparing the failure rates, LR_{uc}, LR_{cc}, average RLF (ARLF) and UL results.

Tables 15 and 16 show the backtesting results of GARCH models for 97.5% and 99% confidence levels. As seen in Tables 15 and 16, GARCH-ASGN model yields higher VaR forecasts than other competitive models. The failure rates of GARCH-ASGN model are more closer to nominal values. According to LR_{uc} and LR_{cc} results, all models produce accurate VaR forecasts. Based on the this results, it is clear that the failure rates of the VaR models is consistent with the expected one. Hence, all models can produce accurate VaR forecasts for ISE-100 index. However, AQLF and UL results show that which model produces the more accurate VaR forecasts. The lower values of AQLF and UL, indicates the more accurate VaR forecasts. Tables 15 and 16 show that the GARCH-ASGN model has the lowest values of AQLF and UL results. Therefore, GARCH-ASGN model can be chosen as the best performed model than among others.

Table 15: Backtesting results of VaR models for ISE-100 index at $p = 0.025$

Models	Mean VaR (%)	Failure rate	LR_{uc}	LR_{cc}	AQLF	UL
GARCH-N	-2.696	0.032	1.010 (0.314)	2.255 (0.323)	0.10865	-0.0353
GARCH-t	-2.701	0.032	1.010 (0.314)	2.255 (0.323)	0.10501	-0.0337
GARCH-GN	-2.793	0.028	0.262 (0.608)	1.255 (0.533)	0.09592	-0.0312
GARCH-SN	-2.825	0.028	0.262 (0.608)	1.255 (0.533)	0.06675	-0.0308
GARCH-ASGN	-3.020	0.023	0.069 (0.791)	0.741 (0.691)	0.0575	-0.0263

p values of LR-uc and LR-cc tests are presented in parentheses.

Table 16: Backtesting results of VaR models for ISE-100 index at $p = 0.01$

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc	AQLF	UL
GARCH-N	-3.212	0.022	6.185 (0.012)	6.762 (0.034)	0.0704	-0.0226
GARCH-t	-3.496	0.017	2.243 (0.134)	2.583 (0.274)	0.0538	-0.0163
GARCH-GN	-3.517	0.017	2.243 (0.134)	2.583 (0.274)	0.0528	-0.0158
GARCH-SN	-3.387	0.018	3.377 (0.066)	3.788 (0.151)	0.0408	-0.0182
GARCH-ASGN	-3.768	0.012	0.159 (0.689)	0.325 (0.849)	0.0302	-0.0141

p values of LR-uc and LR-cc tests are presented in parentheses.

Figures 24 and 25 display the VaR forecasts of GARCH models for for 97.5% and 99% confidence levels, respectively. These figures reveal that GARCH model with ASGN innovation distribution exhibits great consistency for estimating the true quantile value of innovation distribution.

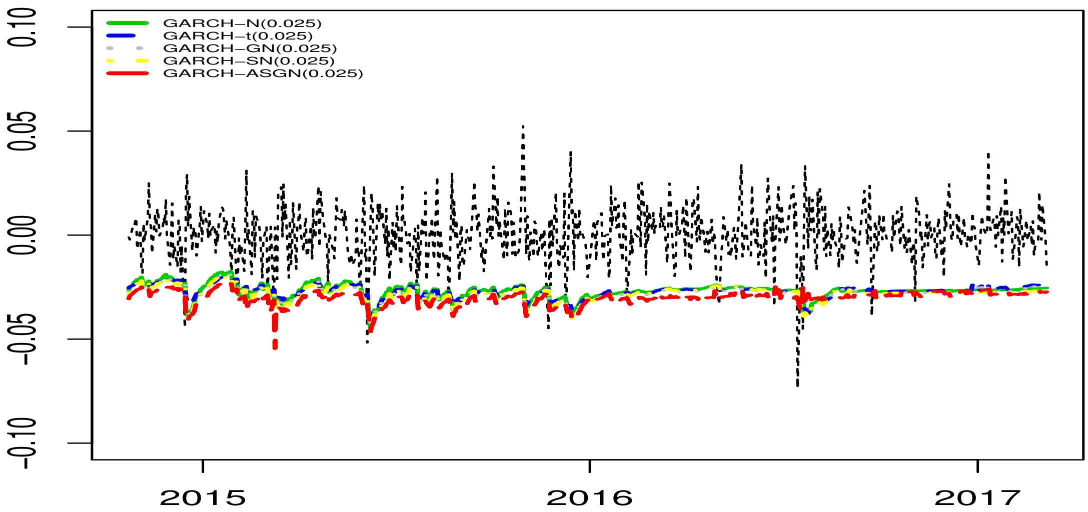


Figure 24: Daily VaR forecasts of GARCH models for 97.5% confidence level

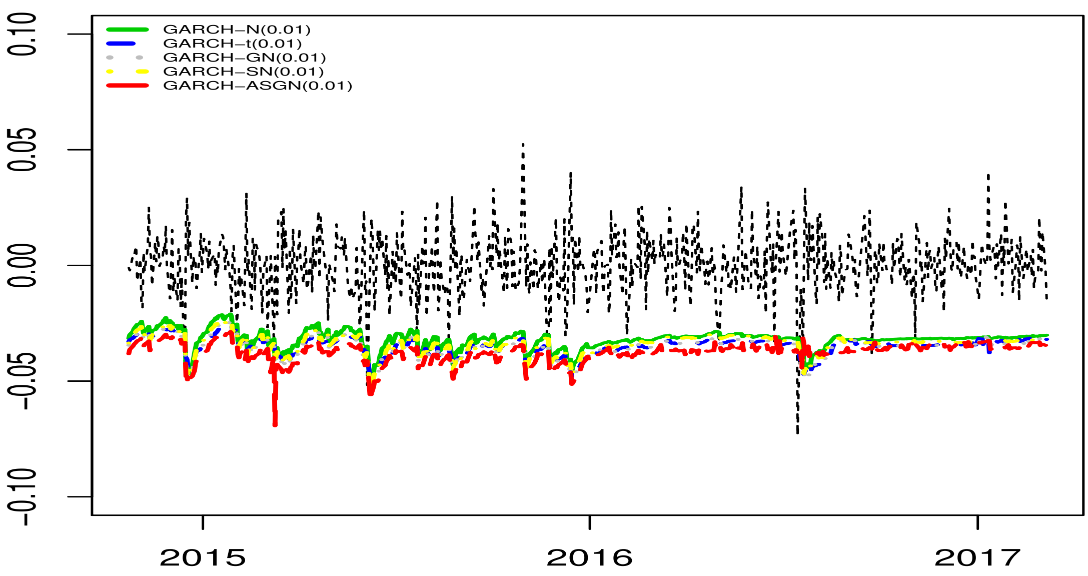


Figure 25: Daily VaR forecasts of GARCH models for 99% confidence level

6.5. Comprehensive Empirical Study

In this section, the VaR models given in Section 4 are compared by means of the results of backtesting and loss functions.

6.5.1. Data Description

Nasdaq-100 index is used to evaluate the performance of all VaR models given in Section 4 in forecasting daily VaR. The used time series data contains 1450 daily observations from 01.03.2011 to 01.12.2017. The descriptive statistics of the log-returns of Nasdaq-100 index are given in Table 17. Figure 26 displays the daily log-returns of Nasdaq-100.

Table 17: Summary statistics of Nasdaq-100 index

Nasdaq-100	
Number of Observations	1450
Minimum	-0.0437
Maximum	0.0493
Mean	0.0006
Median	0.0008
Std. Deviation	0.0092
Skewness	-0.3203
Kurtosis	2.3239
Jarque-Bera	353.2188 ($p<0.001$)
ARCH-LM test	141.3383 ($p<0.001$)
Ljung-Box	11.5365 ($p=0.3173$)

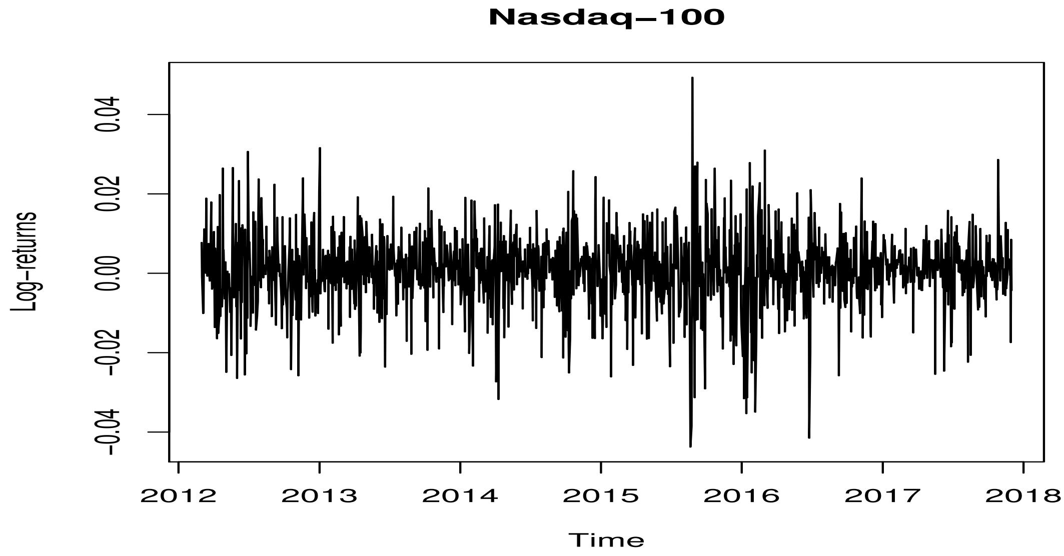


Figure 26: Daily log-returns of Nasdaq-100 index.

Based on the results given in Table 17, it is clear that the mean return is close to 0. The JB test shows that the ISE-100 index is not normally distributed. It can be viewed as an evidence for significant skewness and excess kurtosis. Thus, it is clear that log return of Nasdaq-100 index has the non-normal characteristics, excess kurtosis, and fat tails. ARCH-LM test shows that log returns have a heteroscedasticity problem and it is an evidence of significant ARCH effect. According to result of Ljung-Box test, there is no autocorrelation problem for log-returns of Nasdaq-100 index.

Figure 27 displays the time-varying skewness and kurtosis of Nasdaq-100. For Figure 27, window length is determined as 250 and rolling window procedure is used. Based on Figure 27, it is clear that skewness and kurtosis of Nasdaq-100 index exhibit great variability across the time.

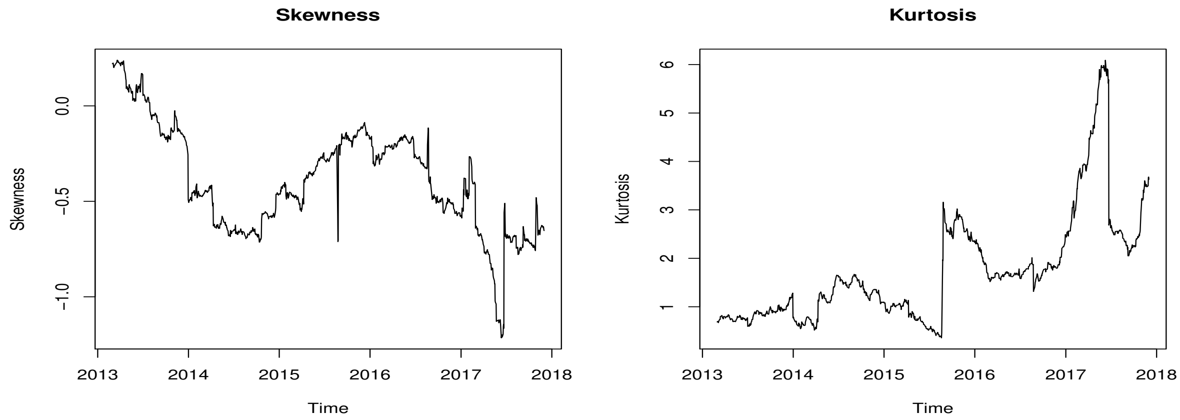


Figure 27: Time varying skewness and kurtosis plots of Nasdaq-100 index

The benchmark model GARCH(1,1), defined in (4.1), is estimated with fourteen different innovation distributions: normal, Student-t, SN, ST, AST, GHST, NIG, GED, SGED, GT, SGT, ASGT, ASGN and GAST. Table 18 shows the estimated parameters of GARCH models. The **constrOptim** function in R is used to minimize negative ll function of GARCH-SN, GARCH-ST, GARCH-AST, GARCH-GT, GARCH-SGT, GARCH-ASGT, GARCH-ASGN, and GARCH-GAST models. The parameter estimations of other models are obtained by using the **rugarch** package of R software.

Table 18: Estimated parameters of GARCH(1,1) model for Nasdaq-100 index, the corresponding standard errors and p-values are given second and third line, respectively.

	Nasdaq-100					
Parameters	Normal	SN	Student-t	ST	AST	GHST
m	8.48×10^{-4}	7.59×10^{-4}	1.09×10^{-3}	8.39×10^{-4}	7.72×10^{-4}	8.35×10^{-4}
	2.14×10^{-4}	2.11×10^{-4}	1.89×10^{-4}	2.04×10^{-4}	2.24×10^{-4}	2.05×10^{-4}
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
ω	7×10^{-6}	7×10^{-6}	5×10^{-6}	1×10^{-6}	1.37×10^{-5}	1×10^{-6}
	2×10^{-6}	2×10^{-6}	2×10^{-6}	1×10^{-6}	1.25×10^{-5}	1×10^{-6}
	<0.001	<0.001	0.007	0.152	0.273	0.049
γ_1	0.116	0.119	0.141	0.098	0.364	0.113
	0.022	0.022	0.031	0.021	0.182	0.024
	<0.001	<0.001	<0.001	<0.001	0.045	<0.001
γ_2	0.798	0.802	0.823	0.902	0.619	0.885
	0.038	0.036	0.036	0.0201	0.178	0.022
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
λ	-	0.872		0.915	-	-
		0.027		0.031		
		<0.001		<0.001		
υ	-	-	4.516	4.922	-	5.055
			0.596	0.643		0.670057
			<0.001	<0.001		<0.001
α	-	-	-	-	0.509	-
					0.041	
					<0.001	
β	-	-	-	-	-	-0.315
						0.144
						0.028
υ_1	-	-	-	-	3.148	-
					0.808	
					<0.001	
υ_2	-	-	-	-	5.875	-
					2.080	
					0.005	
κ	-	-	-	-	-	-
η	-	-	-	-	-	-
$-\ell$	-4807.875	-4817.930	-4855.610	-4858.540	-4848.120	-4856.8

Table 18: Estimated parameters of GARCH(1,1) model for Nasdaq-100 index.

	Nasdaq-100							
Parameters	NIG	GED	SGED	GT	SGT	ASGT	ASGN	GAST
m	7.94×10^{-4}	9.65×10^{-4}	7.11×10^{-4}	1.26×10^{-3}	1.26×10^{-3}	1.16×10^{-3}	7.57×10^{-4}	1.26×10^{-3}
	2.06×10^{-4}	2.26×10^{-4}	1.53×10^{-4}	1.87×10^{-4}	2.13×10^{-4}	2.06×10^{-4}	2.07×10^{-4}	2.07×10^{-4}
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
ω	1×10^{-6}	1×10^{-6}	1×10^{-6}	2.98×10^{-6}	2.83×10^{-6}	4.71×10^{-6}	5.53×10^{-6}	7.66×10^{-6}
	1×10^{-6}	1×10^{-6}	1×10^{-6}	1.08×10^{-6}	1.04×10^{-6}	1.69×10^{-6}	1.85×10^{-6}	2.85×10^{-6}
	0.054	0.048	0.516	0.006	0.006	0.005	0.003	0.007
γ_1	0.106	0.083	0.062	0.123	0.123133	0.131	0.137	0.204
	0.024	0.022	0.015	0.029	0.028	0.028	0.031	0.047
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
γ_2	0.892	0.909	0.937	0.856	0.856	0.820	0.805	0.735
	0.022	0.022	0.014	0.029	0.029	0.037	0.041	0.058
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
λ	-	-	0.924	-	-	-	-	-0.274
			0.026					0.381
			<0.001					0.471
υ	-	-	-	6.191	5.882	6.785	-	4.790
				4.565	4.071	4.864		0.745
				0.175	0.148	0.163		<0.001
α	1.018	-	-	-	-	0.105	0.109	0.118
	0.188					0.035	0.044	0.037
	<0.001					0.003	<0.001	0.001
β	-0.148	-	-	-	-	-	-	-
	0.050							
	0.003							
υ_1	-	-	-	-	-	-	-	-
υ_2	-	-	-	-	-	-	-	-
κ	-	1.180	1.199	1.461	1.503	1.496	1.212	-
		0.058	0.058	0.203	0.205	0.204	0.059	
		<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	
η	-	-	-	-	-0.045	-	-	-
					0.039			
					0.256			
$-\ell$	-4858.519	-4852.740	-4849.180	-4857.550	-4858.640	-4861.100	**-4861.400**	-4855.540

Table 18 shows that GARCH-ASGN model has the lowest ll value among others and therefore, GARCH-ASGN model could be chosen as the best model for Nasdaq-100 index. Figure 28 displays the diagnostic plots of GARCH-ASGN model. Figure 28 proves that GARCH-ASGN model is statistically valid and exhibits superior fits to innovations of GARCH model for Nasdaq-100 index.

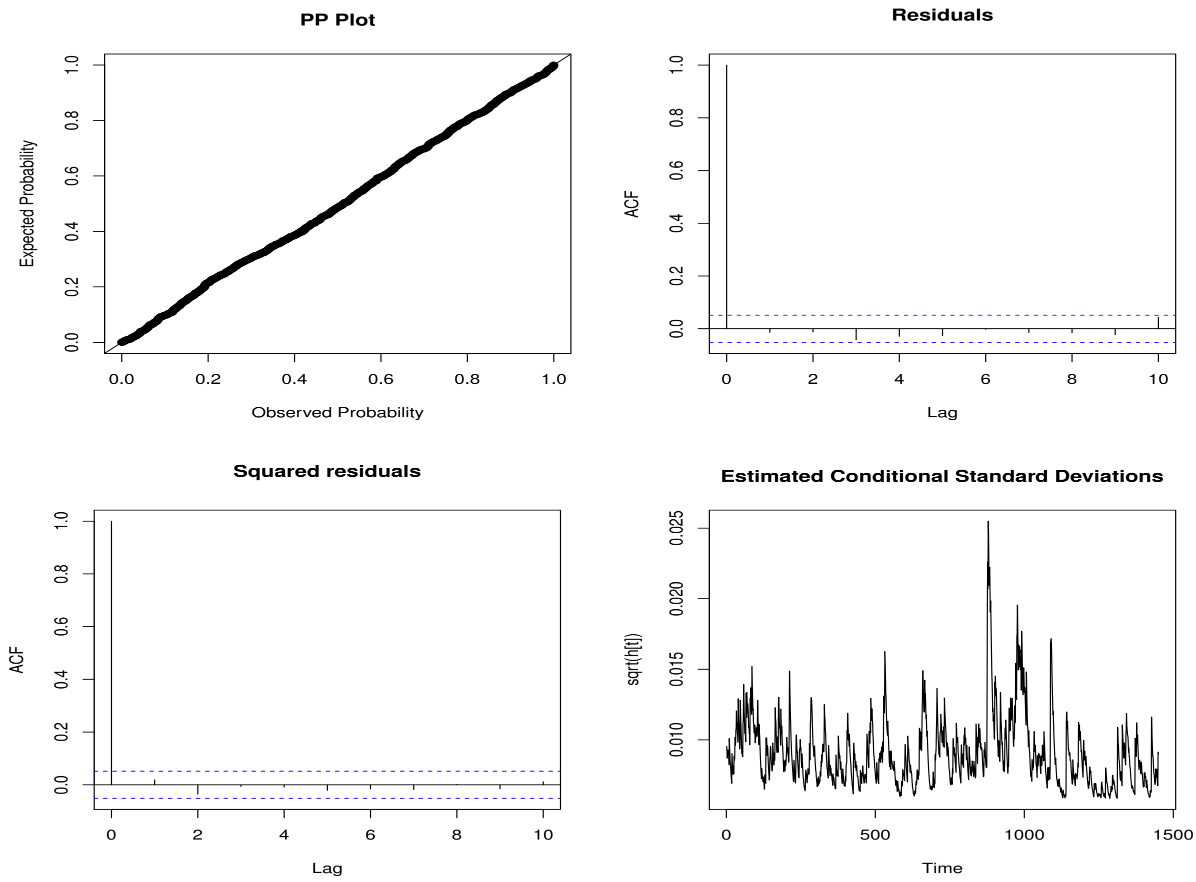


Figure 28: Diagnostic plots of GARCH-ASGN model for Nasdaq-100 index.

6.5.2. Backtesting Results

The parameters of GARCH models specified under fourteen innovation distributions such as normal, SN, Student-t, ST, AST, GHST, NIG, GED, SGED, GT, SGT, ASGT, ASGN, GAST and Riskmetrics^{TM}, GARCH-EVT and FHS models are obtained by means of rolling window estimation method. Window length is determined as 250 and next 1200 daily log-returns are used to compare the performance of VaR models. The forecasting performance of seventeen VaR models are evaluated by comparing the LR_{uc}, LR_{cc}, ARLF, UL FLF and FABL results.

Two stage evaluation process is used to decide the best model. In the first stage, LR_{uc}

and LR_{cc} backtesting results are used to determine the models which produce accurate VaR forecasts. In the second stage, the results of ARLF, UL FLF and FABL loss functions are used to determine the models which have minimum forecast errors and minimum excess capital value.

Empirical Results for Left-Tail (Long-Position) Modeling

Tables 19, 21 and 23 show the average value of VaR forecasts, failure rates, LR_{uc} and LR_{cc} results at $p = 0.01$, $p = 0.025$ and $p = 0.05$, respectively, for seventeen VaR models. Tables 20, 22 and 24 show the results of ARLF, UL FLF and FABL loss functions at $p = 0.01$, $p = 0.025$ and $p = 0.05$, respectively, for seventeen VaR models. As seen in Table 19, seven VaR models produce accurate VaR forecasts at $p = 0.01$ level on the basis of the results of LR_{uc} and LR_{cc} tests. These are GARCH-AST, GARCH-GAST, GARCH-EVT, GARCH-ASGT, GARCH-SGT, GARCH-GT and GARCH-ASGN models. The other models perform poorly and produce underestimated VaR forecasts. Based on the results given in Table 20, GACH-AST model is the best performed model according to results of AQLF and UL loss functions. However, GARCH-GAST model is the best performed model according to results of FLF and FABL loss functions. FLF and FABL loss functions evaluate the models according to theirs forecasting error and excess capital value. Hence, it is easy to conclude that GARCH-AST model has higher excess capital value than GARCH-GAST model. Therefore, GARCH-GAST model is the best model at $p = 0.01$ level and produce the most accurate VaR forecasts among others.

Table 19: Out-of-sample performances of VaR models at $p = 0.01$ for Nasdaq-100 index.

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc
Riskmetrics^{TM}	-1.960146	0.030000	31.588 (<0.001)	32.261 (<0.001)
GARCH-N	-1.942391	0.028333	27.228 (<0.001)	27.230 (<0.001)
GARCH-SN	-2.116057	0.027500	25.139 (<0.001)	25.148 (<0.001)
GARCH-T	-2.192240	0.025833	21.148 (<0.001)	21.196 (<0.001)
GARCH-ST	-2.201780	0.024167	17.423 (<0.001)	17.541 (<0.001)
GARCH-AST	-2.860846	0.009167	0.086 (0.768)	3.028 (0.220)
GARCH-GHST	-2.410733	0.021667	12.371(<0.001)	12.660 (0.002)
GARCH-NIG	-2.428040	0.021667	12.371 (<0.001)	12.660 (0.002)
GARCH-GED	-2.184546	0.025833	21.148 (<0.001)	21.196 (<0.001)
GARCH-SGED	-2.329230	0.021667	12.371 (<0.001)	12.660 (0.002)
GARCH-GT	-2.509783	0.012500	0.702 (0.402)	2.514 (0.284)
GARCH-SGT	-2.610881	0.011667	0.319 (0.572)	2.372 (0.305)
GARCH-ASGT	-2.501769	0.015833	3.503 (0.061)	4.559 (0.102)
GARCH-ASGN	-2.484593	0.015833	3.503 (0.061)	4.559 (0.102)
GARCH-GAST	-2.592540	0.012500	0.702 (0.402)	2.514 (0.284)
GARCH-EVT	-2.638549	0.013333	1.219 (0.269)	2.814 (0.245)
FHS	-2.428463	0.020000	9.392 (0.002)	9.843 (0.007)

p values of LR-uc and LR-cc tests are presented in parentheses.

Table 20: Results of loss functions at $p = 0.01$ for Nasdaq-100 index.

Models	AQLF	UL	FLF	FABL
Riskmetrics^{TM}	0.020345860	-0.018695380	0.039432347	0.040833526
GARCH-N	0.021392290	-0.020026770	0.04033098	0.041717642
GARCH-SN	0.013084760	-0.014929730	0.033721836	0.035093884
GARCH-T	0.015661650	-0.014891530	0.037097299	0.038432264
GARCH-ST	0.012181410	-0.013111300	0.033726415	0.035029578
GARCH-AST	0.003212329	-0.004289289	0.031593947	0.032562934
GARCH-GHST	0.010767410	-0.010795400	0.034419478	0.035681948
GARCH-NIG	0.007171372	-0.008976530	0.030989685	0.032240791
GARCH-GED	0.014824280	-0.014672030	0.036175832	0.037515717
GARCH-SGED	0.008400208	-0.010472960	0.03123078	0.032496493
GARCH-GT	0.006229179	-0.006807509	0.031079236	0.03209434
GARCH-SGT	0.006080951	-0.006272239	0.031959438	0.032951736
GARCH-ASGT	0.005965562	-0.007212417	0.030622813	0.031754626
GARCH-ASGN	0.006693712	-0.007928160	0.031191294	0.03231818
GARCH-GAST	0.004596994	-0.005252798	0.030480216	0.031512287
GARCH-EVT	0.004830963	-0.005666951	0.030654447	0.031734266
FHS	0.006574809	-0.008351412	0.030410468	0.03164221

As seen in Table 21, ten VaR models produce accurate VaR forecasts at $p = 0.025$ level on the basis of the results of LR_{uc} and LR_{cc} tests. These are GARCH-GAST, GARCH-AST, GARCH-EVT, GARCH-SGT, GARCH-GT, GARCH-ASGN, GARCH-ASGT, FHS, GARCH-SGED and GARCH-GHST models. Other models perform poorly and produce underestimated VaR forecasts. Based on the result given in Table 22, GARCH-GAST model is the best performed model according to results of AQLF, UL, FLF and FABL loss functions. Therefore, GARCH-GAST model is the best model at $p = 0.025$ level and produce the most accurate VaR forecasts with minumum excess capital value and minimum forecast error among others.

Table 21: Out-of-sample performances of VaR models at $p = 0.025$ for Nasdaq-100 index.

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc
Riskmetrics^{TM}	-1.642622	0.041667	11.426 (<0.001)	11.430 (0.003)
GARCH-N	-1.624122	0.043333	13.621 (<0.001)	14.569 (<0.001)
GARCH-SN	-1.752156	0.036667	5.871 (0.015)	6.160 (0.045)
GARCH-T	-1.670122	0.041667	11.426 (<0.001)	12.176 (0.002)
GARCH-ST	-1.677082	0.041667	11.426 (<0.001)	12.176 (<0.001)
GARCH-AST	-2.083726	0.027500	0.298 (0.585)	0.307 (0.857)
GARCH-GHST	-1.824539	0.034167	3.718 (0.054)	3.854 (0.145)
GARCH-NIG	-1.860182	0.033333	3.100 (0.078)	3.198 (0.202)
GARCH-GED	-1.721147	0.040833	10.391 (0.001)	11.050 (0.004)
GARCH-SGED	-1.838163	0.034167	3.718 (0.054)	3.854 (0.145)
GARCH-GT	-1.962228	0.030000	1.157 (0.282)	1.164 (0.558)
GARCH-SGT	-2.040309	0.026667	0.133 (0.714)	0.1589 (0.924)
GARCH-ASGT	-1.946904	0.029167	0.811 (0.367)	0.812 (0.666)
GARCH-ASGN	-1.989565	0.027500	0.298 (0.585)	0.307 (0.857)
GARCH-GAST	-2.023920	0.026667	0.133 (0.714)	0.158 (0.924)
GARCH-EVT	-2.127702	0.025833	0.033 (0.854)	0.082 (0.959)
FHS	-1.948280	0.030000	1.157 (0.282)	1.164 (0.558)

p values of LR-uc and LR-cc tests are presented in parentheses.

Table 22: Results of loss functions at $p = 0.025$ for Nasdaq-100 index.

Models	AQLF	UL	FLF	FABL
Riskmetrics^{TM}	0.033577970	-0.029027140	0.049371029	0.050993715
GARCH-N	0.035747310	-0.030260360	0.051344501	0.052992808
GARCH-SN	0.026222320	-0.024835230	0.043167573	0.044691478
GARCH-T	0.034428950	-0.028885570	0.050520928	0.052118272
GARCH-ST	0.030180660	-0.027145940	0.046324846	0.047922191
GARCH-AST	0.014739540	-0.015521790	0.035078128	0.036431272
GARCH-GHST	0.027234830	-0.024185790	0.044936613	0.046421332
GARCH-NIG	0.021956850	-0.022006770	0.04000585	0.041477983
GARCH-GED	0.031465350	-0.027139360	0.048060747	0.049647465
GARCH-SGED	0.022622990	-0.022482440	0.04044397	0.041928692
GARCH-GT	0.016312230	-0.016531960	0.035391121	0.036799085
GARCH-SGT	0.015059830	-0.014760070	0.034956487	0.036309777
GARCH-ASGT	0.018004430	-0.017790340	0.036947161	0.038350627
GARCH-ASGN	0.017400890	-0.016874730	0.036791211	0.038164541
GARCH-GAST	0.013270350	-0.014064900	0.034044061	0.035387275
GARCH-EVT	0.014530520	-0.015165520	0.034277377	0.035620635
FHS	0.018719140	-0.019175030	0.037679664	0.039092947

As seen in Table 23, fifteen VaR models produce accurate VaR forecasts at $p = 0.05$ level on the basis of the results of LR_{uc} and LR_{cc} tests. These are GARCH-GAST, GARCH-AST, GARCH-EVT, GARCH-SGT, GARCH-GT, GARCH-ASGN, GARCH-ASGT, FHS, GARCH-SGED, GARCH-SN, GARCH-NIG, Riskmetris, GARCH-N, GARCH-GED and GARCH-GHST models. The other two models, GARCH-T and GARCH-ST, perform poorly and produce underestimated VaR forecasts. Based on the results given in Table 24, GACH-GAST model is the best performed model according to results of AQLF, UL, FLF and FABL loss functions. Therefore, GARCH-GAST model is the best model at $p = 0.005$ level and produce the most accurate VaR forecasts with minumum excess capital value and minimum forecast error among others.

Table 23: Out-of-sample performances of VaR models at $p = 0.05$ for Nasdaq-100 index.

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc
Riskmetrics^{TM}	-1.369533	0.060833	2.781 (0.095)	2.833 (0.242)
GARCH-N	-1.350269	0.059167	2.010 (0.156)	2.171 (0.337)
GARCH-SN	-1.440973	0.055000	0.612 (0.689)	0.743 (0.689)
GARCH-T	-1.298487	0.068333	7.656 (0.006)	7.735 (0.021)
GARCH-ST	-1.303731	0.068333	7.656 (0.006)	7.735 (0.021)
GARCH-AST	-1.590137	0.045833	0.450(0.502)	1.735 (0.420)
GARCH-GHST	-1.407899	0.057500	1.358 (0.244)	1.647 (0.439)
GARCH-NIG	-1.434775	0.055833	0.829 (0.362)	1.006 (0.605)
GARCH-GED	-1.353171	0.059167	2.010 (0.156)	2.021 (0.364)
GARCH-SGED	-1.445667	0.055000	0.612 (0.689)	0.743 (0.689)
GARCH-GT	-1.562823	0.049167	0.017 (0.894)	0.363 (0.834)
GARCH-SGT	-1.618168	0.045833	0.450 (0.502)	0.578 (0.749)
GARCH-ASGT	-1.541274	0.050833	0.017 (0.895)	0.021 (0.989)
GARCH-ASGN	-1.594829	0.045833	0.450 (0.502)	0.578 (0.749)
GARCH-GAST	-1.610245	0.044167	0.893 (0.344)	0.950 (0.622)
GARCH-EVT	-1.660786	0.040000	2.704 (0.100)	2.707 (0.258)
FHS	-1.486718	0.051667	0.069 (0.792)	0.644 (0.724)

p values of LR-uc and LR-cc tests are presented in parentheses.

Table 24: Results of loss functions at $p = 0.05$ for Nasdaq-100 index.

Models	AQLF	UL	FLF	FABL
Riskmetrics^{TM}	0.051714460	-0.042597270	0.064633671	0.066535023
GARCH-N	0.054406510	-0.043585290	0.067125896	0.069044298
GARCH-SN	0.044201390	-0.037780300	0.057878396	0.059688175
GARCH-T	0.059085840	-0.047246340	0.071220892	0.073242426
GARCH-ST	0.054725880	-0.045827160	0.066898858	0.068920715
GARCH-AST	0.034135240	-0.030448930	0.049354839	0.051040353
GARCH-GHST	0.050131950	-0.040499630	0.063467753	0.065315185
GARCH-NIG	0.044451950	-0.038446760	0.058054324	0.059883421
GARCH-GED	0.054629490	-0.043957970	0.067388092	0.069300025
GARCH-SGED	0.043738080	-0.037573400	0.05745321	0.059269733
GARCH-GT	0.033080060	-0.030071930	0.047958774	0.049708266
GARCH-SGT	0.030816020	-0.027888770	0.046289608	0.04797584
GARCH-ASGT	0.036661480	-0.032052220	0.05129764	0.053093995
GARCH-ASGN	0.034116350	-0.029891950	0.049337629	0.051062805
GARCH-GAST	0.030303420	-0.027313910	0.046168563	0.047808845
GARCH-EVT	0.030725940	-0.028119700	0.046286443	0.047883673
FHS	0.040783630	-0.035337260	0.054942533	0.056703435

Table 25 shows the ranking of VaR models according to results of loss functions for long position. The models that passed the first stage are ranked according to results of loss functions. GARCH-AST could be evaluated as the best model according results of AQLF and UL loss functions. However, it is easy to see that GARCH-AST model has significantly higher excess capital value than GARCH-GAST model. Based on the FLF and FABL results, GARCH-AST is dropped to sixth rank. The VaR forecasts of GARCH-GAST is more consistent than other models. It is clear that GARCH-GAST is the best performed model based on the results of four loss functions for left-tail modeling.

Table 25: Ranking of VaR models according to results of loss functions for the long position

Long position			
$p = 0.01$			
AQLF	UL	FLF	FABL
GARCH-AST	GARCH-AST	GARCH-GAST	GARCH-GAST
GARCH-GAST	GARCH-GAST	GARCH-ASGT	GARCH-EVT
GARCH-EVT	GARCH-EVT	GARCH-EVT	GARCH-ASGT
GARCH-ASGT	GARCH-SGT	GARCH-GT	GARCH-GT
GARCH-SGT	GARCH-GT	GARCH-ASGN	GARCH-ASGN
GARCH-GT	GARCH-ASGT	GARCH-AST	GARCH-AST
GARCH-ASGN	GARCH-ASGN	GARCH-SGT	GARCH-SGT
$p = 0.025$			
AQLF	UL	FLF	FABL
GARCH-GAST	GARCH-GAST	GARCH-GAST	GARCH-GAST
GARCH-EVT	GARCH-SGT	GARCH-EVT	GARCH-EVT
GARCH-AST	GARCH-EVT	GARCH-SGT	GARCH-SGT
GARCH-SGT	GARCH-AST	GARCH-AST	GARCH-AST
GARCH-GT	GARCH-GT	GARCH-GT	GARCH-GT
GARCH-ASGN	GARCH-ASGN	GARCH-ASGN	GARCH-ASGN
GARCH-ASGT	GARCH-ASGT	GARCH-ASGT	GARCH-ASGT
FHS	FHS	FHS	FHS
GARCH-SGED	GARCH-SGED	GARCH-SGED	GARCH-SGED
GARCH-GHST	GARCH-GHST	GARCH-GHST	GARCH-GHST
$p = 0.05$			
AQLF	UL	FLF	FABL
GARCH-GAST	GARCH-GAST	GARCH-GAST	GARCH-GAST
GARCH-EVT	GARCH-SGT	GARCH-EVT	GARCH-EVT
GARCH-SGT	GARCH-EVT	GARCH-SGT	GARCH-SGT
GARCH-GT	GARCH-ASGN	GARCH-GT	GARCH-GT
GARCH-ASGN	GARCH-GT	GARCH-ASGN	GARCH-AST
GARCH-AST	GARCH-AST	GARCH-AST	GARCH-ASGN
GARCH-ASGT	GARCH-ASGT	GARCH-ASGT	GARCH-ASGT
FHS	FHS	FHS	FHS
GARCH-SGED	GARCH-SGED	GARCH-SGED	GARCH-SGED
GARCH-SN	GARCH-SN	GARCH-SN	GARCH-SN
GARCH-NIG	GARCH-NIG	GARCH-NIG	GARCH-NIG
GARCH-GHST	GARCH-GHST	GARCH-GHST	GARCH-GHST
Riskmetrics^{TM}	Riskmetrics^{TM}	Riskmetrics^{TM}	Riskmetrics^{TM}
GARCH-N	GARCH-N	GARCH-N	GARCH-N
GARCH-GED	GARCH-GED	GARCH-GED	GARCH-GED

Empirical Results for Right-Tail (Short-Position) Modeling

Tables 26, 28 and 30 show the average value of VaR forecasts, failure rates, LR_{uc} and LR_{cc} results at $p = 0.99$, $p = 0.975$ and $p = 0.95$, respectively, for seventeen VaR models. Tables 27, 29 and 31 show the results of ARLF, UL, FLF and FABL loss functions at $p = 0.99$, $p = 0.975$ and $p = 0.95$, respectively, for seventeen VaR models. As seen in Table 26, ten of VaR models produce accurate VaR forecasts at $p = 0.99$ level on the basis of the results of LR_{uc} and LR_{cc} tests. These are GARCH-GAST,GARCH-SGED, GARCH-AST, GARCH-NIG, GARCH-GHST, GARCH-SN, Riskmetrics, GARCH-N, GARCH-EVT and FHS models. The other models perform poorly and produce overestimated VaR forecasts. Based on the results given in Table 27, GACH-GAST model is the best performed model according to results of AQLF, UL, FLF and FABL loss functions. Therefore, GARCH-GAST model is the best performed model at $p = 0.99$ level for short position.

Table 26: Out-of-sample performances of VaR models at $p = 0.99$ for Nasdaq-100 index.

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc
Riskmetrics^{TM}	2.072083	0.990833	0.086 (0.768)	0.290 (0.865)
GARCH-N	2.099291	0.992500	0.829 (0.362)	0.965 (0.617)
GARCH-SN	1.926702	0.987500	0.701 (0.402)	2.514 (0.284)
GARCH-T	2.395289	0.995833	5.286 (0.021)	5.328 (0.069)
GARCH-ST	2.405007	0.997500	9.750 (0.002)	9.765 (0.007)
GARCH-AST	2.222486	0.992500	0.829 (0.362)	0.965 (0.617)
GARCH-GHST	2.032741	0.991667	0.356 (0.550)	3.667 (0.160)
GARCH-NIG	2.097494	0.992500	0.829 (0.362)	4.556 (0.102)
GARCH-GED	2.367327	0.995833	5.286 (0.021)	5.328 (0.069)
GARCH-SGED	2.134111	0.992500	0.829 (0.362)	0.965 (0.617)
GARCH-GT	2.509783	0.998333	12.916 (<0.001)	12.923 (0.002)
GARCH-SGT	2.288419	0.997500	9.750 (0.002)	9.765(0.007)
GARCH-ASGT	2.260199	0.996667	7.264 (0.007)	7.291 (0.026)
GARCH-ASGN	2.312384	0.995833	5.286 (0.022)	5.328 (0.070)
GARCH-GAST	2.180479	0.993333	1.526 (0.216)	5.729 (0.057)
GARCH-EVT	1.990751	0.990833	0.086 (0.768)	3.028 (0.219)
FHS	1.947894	0.990833	0.086 (0.768)	3.028 (0.219)

p values of LR-uc and LR-cc tests are presented in parentheses.

Table 27: Results of loss functions at $p = 0.99$ for Nasdaq-100 index.

Models	AQLF	UL	FLF	FABL
Riskmetrics^{TM}	0.004252293	0.004055139	0.024805878	0.02431442
GARCH-N	0.004649519	0.004004361	0.025496785	0.024981613
GARCH-SN	0.003797671	0.004676088	0.023961078	0.023411586
GARCH-T	0.001701321	0.001594453	0.02556641	0.024970904
GARCH-ST	0.001657597	0.001364904	0.02565154	0.025022064
GARCH-AST	0.002614729	0.003204915	0.024672555	0.024172384
GARCH-GHST	0.003354847	0.003763992	0.023499365	0.023020648
GARCH-NIG	0.002796637	0.003112300	0.023597543	0.023103448
GARCH-GED	0.001517913	0.001721999	0.025079565	0.024509155
GARCH-SGED	0.002370973	0.002617218	0.023539791	0.023039003
GARCH-GT	0.001184021	0.001186422	0.026249816	0.025594465
GARCH-SGT	0.001492774	0.001131291	0.024322304	0.023689025
GARCH-ASGT	0.002501239	0.002464591	0.025016958	0.024428627
GARCH-ASGN	0.002230654	0.002353767	0.025249193	0.024678782
GARCH-GAST	0.002282065	0.002398051	0.022806035	0.022412196
GARCH-EVT	0.007820032	0.005525755	0.027541194	0.027083546
FHS	0.008397837	0.006040184	0.027694321	0.027237925

As seen in Table 28, twelve of VaR models produce accurate VaR forecasts at $p = 0.975$ level on the basis of the results of LR_{uc} and LR_{cc} tests. These are GARCH-GAST, GARCH-SGT, GARCH-SGED, GARCH-AST, GARCH-ASGT, GARCH-NIG, GARCH-GHST, GARCH-SN, Riskmetrics, GARCH-N, GARCH-EVT and FHS models. Other models perform poorly and produce overestimated VaR forecasts. Based on the results given in Table 29, GARCH-GAST model is the best performed model according to results of AQLF, UL, FLF and FABL loss functions. Therefore, GARCH-GAST model is the best performed model at $p = 0.975$ level for short position.

Table 28: Out-of-sample performances of VaR models at $p = 0.975$ for Nasdaq-100 index.

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc
Riskmetrics^{TM}	1.754559	0.980000	1.319 (0.251)	1.770262 (0.413)
GARCH-N	1.781022	0.982500	3.088 (0.078)	3.864205 (0.144)
GARCH-SN	1.656783	0.973333	0.133 (0.714)	0.1589077 (0.923)
GARCH-T	1.873165	0.988333	10.877 (<0.001)	12.93094 (0.001)
GARCH-ST	1.880309	0.988333	10.877 (<0.001)	12.93094 (0.001)
GARCH-AST	1.778105	0.976667	0.139 (0.708)	0.306 (0.858)
GARCH-GHST	1.674642	0.973333	0.133 (0.714)	0.158 (0.923)
GARCH-NIG	1.705943	0.974167	0.033 (0.854)	0.082 (0.956)
GARCH-GED	1.903928	0.987500	11.210 (<0.001)	9.397 (0.002)
GARCH-SGED	1.741966	0.976667	0.139 (0.708)	0.306 (0.858)
GARCH-GT	1.962228	0.989167	12.503 (<0.001)	14.822 (<0.001)
GARCH-SGT	1.827286	0.980833	1.819 (0.177)	2.366 (0.306)
GARCH-ASGT	1.793682	0.981667	2.407 (0.121)	3.062 (0.216)
GARCH-ASGN	1.849769	0.987500	9.397 (0.002)	11.210 (0.004)
GARCH-GAST	1.778716	0.980833	1.819 (0.177)	2.366 (0.306)
GARCH-EVT	1.699757	0.975000	0 (1)	0.079 (0.961)
FHS	1.682527	0.971667	0.524 (0.468)	0.526 (0.768)

p values of LR-uc and LR-cc tests are presented in parentheses.

Table 29: Results of loss functions at $p = 0.975$ for Nasdaq-100 index.

Models	AQLF	UL	FLF	FABL
Riskmetrics^{TM}	0.007884943	0.008150897	0.025090493	0.024812786
GARCH-N	0.008577423	0.007943489	0.026076413	0.025766484
GARCH-SN	0.007732292	0.009905903	0.023857085	0.023699931
GARCH-T	0.004555089	0.005305897	0.023055753	0.02264055
GARCH-ST	0.004189176	0.004759324	0.022755811	0.022340608
GARCH-AST	0.006404154	0.008215765	0.023777037	0.023568108
GARCH-GHST	0.007471018	0.009407202	0.023771972	0.023612254
GARCH-NIG	0.006826866	0.008759346	0.023445044	0.023274638
GARCH-GED	0.004713092	0.005237554	0.023499166	0.023105496
GARCH-SGED	0.006115850	0.007868273	0.023127944	0.022914937
GARCH-GT	0.003208314	0.003708784	0.022614225	0.022168431
GARCH-SGT	0.005532482	0.006401832	0.022969104	0.022684411
GARCH-ASGT	0.006509263	0.006435837	0.024117854	0.023811191
GARCH-ASGN	0.005706436	0.005818383	0.023964661	0.023563058
GARCH-GAST	0.004148469	0.005478655	0.022075434	0.021776858
GARCH-EVT	0.011918590	0.010434830	0.028486538	0.028321251
FHS	0.012378080	0.010964740	0.028726007	0.028613738

As seen in Table 30, sixteen VaR models produce accurate VaR forecasts at $p =$

0.95 level on the basis of the results of LR_{uc} and LR_{cc} tests. These are GARCH-GAST, GARCH-SGT, GARCH-SGED, GARCH-AST, GARCH-ASGT, GARCH-NIG, GARCH-GHST, GARCH-SN, Riskmetrics, GARCH-N, GARCH-EVT and FHS models. GARCH-GT model performs poorly and produce overestimated VaR forecasts at $p = 0.95$ level. Based on the results given in Table 31, GARCH-ST and GARCH-GED models are the best performed models according to results of AQLF, UL loss functions, respectively. However, GARCH-GAST model is the best performed model according to results of FLF and FABL loss functions at $p = 0.95$ level. Hence, it is easy to conclude that GARCH-ST and GARCH-GED models have higher excess capital value than GARCH-GAST model. Therefore, GARCH-GAST model is the best model at $p = 0.95$ level and produce the most accurate VaR forecasts among others.

Table 30: Out-of-sample performances of VaR models at $p = 0.95$ for Nasdaq-100 index.

Models	Mean VaR (%)	Failure rate	LR-uc	LR-cc
Riskmetrics™	1.481470	0.957500	1.493 (0.222)	1.508 (0.470)
GARCH-N	1.507182	0.960000	2.704 (0.100)	2.707(0.258)
GARCH-SN	1.422886	0.950833	0.017 (0.894)	0.432 (0.806)
GARCH-T	1.501531	0.960833	3.193 (0.074)	3.687 (0.158)
GARCH-ST	1.506958	0.960833	3.193 (0.074)	3.687 (0.158)
GARCH-AST	1.458960	0.953333	0.286 (0.592)	0.459 (0.794)
GARCH-GHST	1.395976	0.946667	0.274 (0.600)	0.381 (0.826)
GARCH-NIG	1.404186	0.947500	0.155 (0.693)	0.306 (0.858)
GARCH-GED	1.535952	0.965833	3.718 (0.054)	3.854 (0.145)
GARCH-SGED	1.427695	0.951667	0.070 (0.790)	0.085 (0.958)
GARCH-GT	1.562823	0.969167	10.687 (0.001)	10.706 (0.005)
GARCH-SGT	1.483608	0.958333	1.855 (0.173)	1.859 (0.395)
GARCH-ASGT	1.452726	0.954167	0.450 (0.502)	0.544 (0.761)
GARCH-ASGN	1.492163	0.959167	2.258 (0.132)	2.918 (0.232)
GARCH-GAST	1.459766	0.954167	0.450 (0.502)	0.544 (0.761)
GARCH-EVT	1.434281	0.954167	0.450 (0.502)	1.269 (0.530)
FHS	1.417881	0.950000	0 (1)	0.335 (0.845)

p values of LR-uc and LR-cc tests are presented in parentheses.

Table 31: Results of loss functions at $p = 0.95$ for Nasdaq-100 index.

Models	AQLF	UL	FLF	FABL
Riskmetrics^{TM}	0.014294340	0.015564790	0.028515107	0.028565441
GARCH-N	0.015229000	0.015478590	0.029697791	0.029756358
GARCH-SN	0.014614350	0.018464380	0.028122044	0.028328598
GARCH-T	0.011583800	0.014040810	0.026006734	0.026040265
GARCH-ST	0.011027390	0.013558310	0.025499768	0.025533299
GARCH-AST	0.014239840	0.017923760	0.028149398	0.028309431
GARCH-GHST	0.015168890	0.019308990	0.028372363	0.028622489
GARCH-NIG	0.014862170	0.019039440	0.028155348	0.028395172
GARCH-GED	0.011721170	0.013210680	0.026543181	0.026513542
GARCH-SGED	0.014228970	0.017868030	0.027808789	0.027985345
GARCH-GT	0.009077096	0.011298330	0.024226811	0.024119057
GARCH-SGT	0.012254920	0.015852990	0.02617971	0.026311865
GARCH-ASGT	0.014795240	0.016888860	0.028639413	0.028792134
GARCH-ASGN	0.013199090	0.015236180	0.027504856	0.027573826
GARCH-GAST	0.011066100	0.014684410	0.025304608	0.025349766
GARCH-EVT	0.019202930	0.019792740	0.032885255	0.033044412
FHS	0.019745450	0.020492980	0.033214988	0.033429941

Table 32 shows the ranking of VaR models according to results of loss functions for long position. It is clear that GARCH-GAST is the best performed model based on the results of four loss functions for right-tail modeling.

Table 32: Ranking of VaR models according to results of loss functions for short position

Short Position			
$p = 0.99$			
AQLF	UL	FLF	FABL
GARCH-GAST	GARCH-GAST	GARCH-GAST	GARCH-GAST
GARCH-SGED	GARCH-SGED	GARCH-GHST	GARCH-GHST
GARCH-AST	GARCH-NIG	GARCH-SGED	GARCH-SGED
GARCH-NIG	GARCH-AST	GARCH-NIG	GARCH-NIG
GARCH-GHST	GARCH-GHST	GARCH-SN	GARCH-SN
GARCH-SN	GARCH-N	GARCH-AST	GARCH-AST
Riskmetrics^{TM}	Riskmetrics^{TM}	Riskmetrics^{TM}	Riskmetrics^{TM}
GARCH-N	GARCH-SN	GARCH-N	GARCH-N
GARCH-EVT	GARCH-EVT	GARCH-EVT	GARCH-EVT
FHS	FHS	FHS	FHS
$p = 0.975$			
AQLF	UL	FLF	FABL
GARCH-GAST	GARCH-GAST	GARCH-GAST	GARCH-GAST
GARCH-SGT	GARCH-SGT	GARCH-SGT	GARCH-SGT
GARCH-SGED	GARCH-ASGT	GARCH-SGED	GARCH-SGED
GARCH-AST	GARCH-SGED	GARCH-NIG	GARCH-NIG
GARCH-ASGT	GARCH-N	GARCH-GHST	GARCH-AST
GARCH-NIG	Riskmetrics^{TM}	GARCH-AST	GARCH-GHST
GARCH-GHST	GARCH-AST	GARCH-SN	GARCH-SN
GARCH-SN	GARCH-NIG	GARCH-ASGT	GARCH-ASGT
Riskmetrics^{TM}	GARCH-GHST	Riskmetrics^{TM}	Riskmetrics^{TM}
GARCH-N	GARCH-SN	GARCH-N	GARCH-N
GARCH-EVT	GARCH-EVT	GARCH-EVT	GARCH-EVT
FHS	FHS	FHS	FHS
$p = 0.95$			
AQLF	UL	FLF	FABL
GARCH-ST	GARCH-GED	GARCH-GAST	GARCH-GAST
GARCH-GAST	GARCH-ST	GARCH-ST	GARCH-ST
GARCH-T	GARCH-T	GARCH-T	GARCH-T
GARCH-GED	GARCH-GAST	GARCH-SGT	GARCH-SGT
GARCH-SGT	GARCH-ASGN	GARCH-GED	GARCH-GED
GARCH-ASGN	GARCH-N	GARCH-ASGN	GARCH-ASGN
GARCH-SGED	Riskmetrics^{TM}	GARCH-SGED	GARCH-SGED
GARCH-AST	GARCH-SGT	GARCH-SN	GARCH-AST
Riskmetrics^{TM}	GARCH-ASGT	GARCH-AST	GARCH-SN
GARCH-SN	GARCH-SGED	GARCH-NIG	GARCH-NIG
GARCH-ASGT	GARCH-AST	GARCH-GHST	Riskmetrics^{TM}
GARCH-NIG	GARCH-SN	Riskmetrics^{TM}	GARCH-GHST
GARCH-GHST	GARCH-NIG	GARCH-ASGT	GARCH-ASGT
GARCH-N	GARCH-GHST	GARCH-N	GARCH-N
GARCH-EVT	GARCH-EVT	GARCH-EVT	GARCH-EVT
FHS	FHS	FHS	FHS

Figure 29 displays the VaR forecasts of GARCH-GAST model for $p = 0.01$, $p = 0.025$, $p = 0.05$, $p = 0.99$, $p = 0.975$, $p = 0.95$ levels. This figure reveals that GARCH model with GAST innovation distribution exhibits great consistency for estimating the true quantile value of innovation distribution.

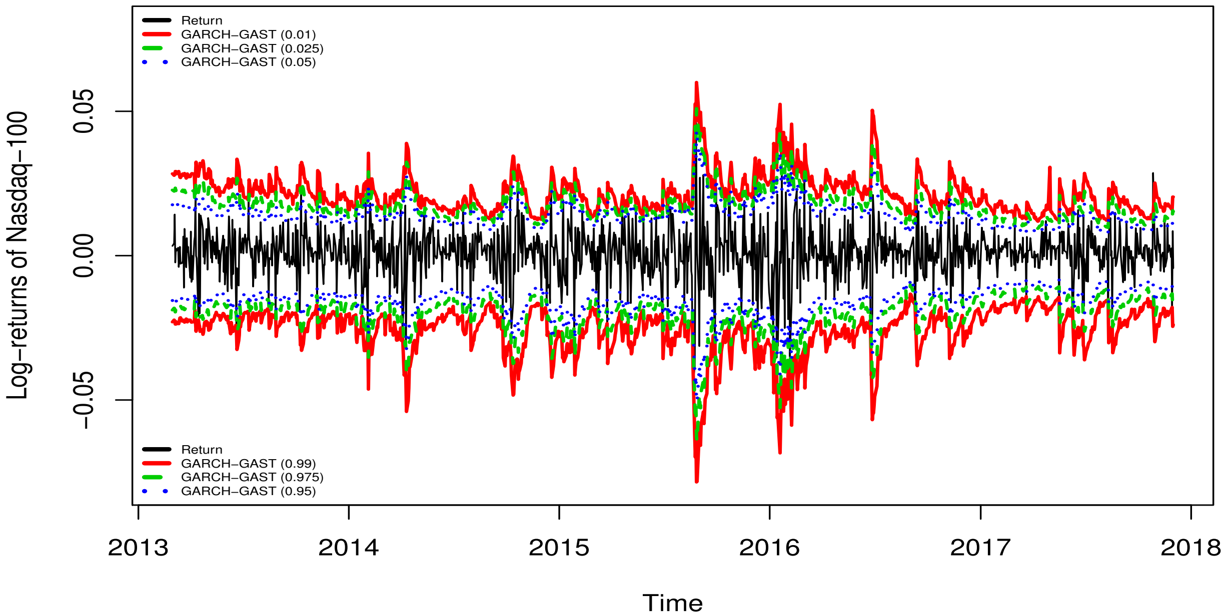


Figure 29: Daily VaR forecasts of GARCH-GAST model for long and short positions

7. CONCLUSION

Forecasting VaR with misspecified model causes to underestimation or overestimation of the real market risk. Accurately forecasting VaR with GARCH models requires a flexible innovation distribution. The flexible distributions increase the modeling accuracy of empirical distribution of financial returns. The selection of innovation distribution for ε_t is important task since it effects the quality of the VaR forecasts directly.

The contributions of this study can be summarized as follows: ASGT distribution is applied to GARCH models for the first time. A new skew generalization of generalized normal distribution, called ASGN, is proposed and applied to GARCH models. A new generalized Skew-T distribution, called GAST, is proposed and applied to GARCH models for symmetric and asymmetric cases. The importance of distributional assumption on innovation process for FHS model is investigated.

The empirical results of this study can be summarized as follows: Five applications to real data sets are provided to demonstrate the out-of-sample performance of GARCH models with fifteen innovation distribution, GARCH-EVT and FHS models in terms of accuracy of VaR forecasts. The first four applications are related to research papers produced by the thesis. The last application is the comprehensive application study consists of all VaR models given in Section 4.

Based on the results given in comprehensive empirical study, GARCH model with GAST, AST, ASGT, SGT, GT and ASGN innovation distributions produce accurate VaR forecasts at $p = 0.01$ level for the long position. Moreover, GARCH-EVT model produces also accurate VaR forecasts at $p = 0.01$ level for long position. According to results of loss functions, GARCH model with GAST innovation distribution is the best model and provides superior fits to empirical distribution of financial returns of Nasdaq-100 index. The similar results are also obtained for $p = 0.025$ and $p = 0.05$ levels. Consequently, GARCH model with GAST innovation distribution produce the most realistic and accurate VaR forecasts among others for long position. When the empirical results are examined for short position, GARCH model with GAST, SGED, AST, NIG, GHST, SN and normal innovation distributions produce accurate VaR forecasts at $p = 0.99$ level for short position. Moreover, GARCH-EVT and FHS models produce also accurate VaR forecasts at $p = 0.99$ level for short position. According to results of loss functions, GARCH model with GAST innovation distribution is the best model and provides superior fits to empirical distribution of financial returns of Nasdaq-100 index. The similar results are also obtained for $p = 0.975$ and $p = 0.95$ levels. Consequently, GARCH model with GAST innovation distribution produces the most realistic and accurate VaR forecasts among others for short position.

As a result of this study, it is clear that the effects of skewness and fat-tails are more important than only the effect of fat-tail and skewness on VaR forecast in stock markets. Accurately modeling of the financial returns is an essential task in estimating the real market risk. This study proves that the skewed and fat-tailed distributions produce better VaR forecasts based on the backtesting and loss function results. Therefore, new financial risk systems should be developed taking into account both effects of skewness and kurtosis. The flexible distributions should be considered in financial risk management. We hope that the results given in this study will be useful for researchers studying in financial risk management, regulators and risk managers.

REFERENCES

[1] Morgan, J.P., *Riskmetrics Technical Document*, 4th ed. J.P. Morgan, New York, **1996**.

[2] Hull, J. and White, A., Incorporating volatility updating into the historical simulation method for value-at-risk, *Journal of Risk*, 1(1), 5-19 **1998**.

[3] Barone-Adesi, G., Giannopoulos, K. and Vosper, L., VaR without correlations for nonlinear portfolios, *Journal of Futures Markets*, 19, 583-602, **1999**.

[4] Angelidis, T., Benos, A. and Degiannakis, S., The use of GARCH models in VaR estimation, *Statistical Methodology*, 1(1), 105-128, **2004**.

[5] Lee, M. C., Su, J. B. and Liu, H. C., Value-at-risk in US stock indices with skewed generalized error distribution, *Applied Financial Economics Letters*, 4(6), 425-431,**2008**.

[6] Bali, T. G. and Theodossiou, P., A conditional-SGT-VaR approach with alternative GARCH models, *Annals of Operations Research*, 151(1), 241-267, **2007**.

[7] Lee, C. F. and Su, J. B., Alternative statistical distributions for estimating value-at-risk: theory and evidence, *Review of Quantitative Finance and Accounting*, 39(3), 309-331, **2012**.

[8] Venkataraman, S., Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques, *Economic Perspectives-Federal Reserve Bank of Chicago*, 21, 2-13, **1997**.

[9] Zangari, P., An improved methodology for measuring VaR. *RiskMetrics Monitor*, 2(1), 7-25, **1996**.

[10] Giot, P. and Laurent, S., Market risk in commodity markets: a VaR approach, *Energy Economics*, 25(5), 435-457, **2003**.

[11] Brooks, C. and Persand, G., The effect of asymmetries on stock index return Value-at-Risk estimates, *The Journal of Risk Finance*, 4(2), 29-42, **2003**.

[12] Hung, J. C., Lee, M. C. and Liu, H. C., Estimation of value-at-risk for energy commodities via fat-tailed GARCH models, *Energy Economics*, 30(3), 1173-1191, **2008**.

[13] Politis, D. N., A heavy-tailed distribution for ARCH residuals with application to volatility prediction, *Annals of Economics and Finance*, 5, 283-298, **2004**.

[14] Dendramis, Y., Spungin, G. E. and Tzavalis, E., Forecasting VaR models under different volatility processes and distributions of return innovations. *Journal of Forecasting*, 33(7), 515-531, **2014**.

[15] Braione, M. and Scholtes, N. K., Forecasting Value-at-Risk under Different Distributional Assumptions, *Econometrics*, 4(3), 1-27, **2016**.

[16] Lyu, Y., Wang, P., Wei, Y. and Ke, R., Forecasting the VaR of crude oil market: Do alternative distributions help?. *Energy Economics*, 66, 523-534, **2017**.

[17] McNeil, A. J. and Frey, R., Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach, *Journal of Empirical Finance*, 7, 271-300, **2000**.

[18] Gilli, M. and Kellezi, E., An Application of Extreme Value Theory for Measuring Financial Risk, *Computational Economics*, 27(1), 1-23, **2006**.

[19] Chan, K. F. and Gray, P., Using extreme value theory to measure value-at-risk for daily electricity spot prices, *International Journal of Forecasting*, 22, 283-300, **2006**.

[20] Onour, I. A., Extreme risk and fat-tails distribution model: empirical analysis, *Journal of Money, Investment and Banking*, 13, **2010**.

[21] Singh, A. K., Allen, D. E. and Robert, P. J., Extreme market risk and extreme value theory, *Mathematics and Computers in Simulation*, 94, 310-328, **2013**.

[22] Karmakar, M., Estimation of tail-related risk measures in the Indian stock market: An extreme value approach, *Review of Financial Economics*, 22, 79-85, **2013**.

[23] Furió, D. and Climent, F. J., Extreme value theory versus traditional GARCH approaches applied to financial data: a comparative evaluation. *Quantitative Finance*, 13(1), 45-63, **2013**.

[24] Harmantzis, F. C., Miao, L. and Chien, Y., Empirical study of value-at-risk and expected shortfall models with heavy tails. *The Journal of Risk Finance*, 7(2), 117-135, **2006**.

[25] Altun, E. and Tatlidil, H., A Comparison of Extreme Value Theory with Heavy-tailed Distributions in Modeling Daily VAR. *Journal of Finance and Investment Analysis*, 4(2), 69-83, **2015**.

[26] Barone-Adesi, G., and Giannopoulos, K., Non parametric var techniques, myths and realities. *Economic Notes*, 30(2), 167-181, **2001**.

[27] Kuester, K., Mittnik, S., and Paolella, M. S., Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1), 53-89, **2006**.

[28] Roy, I., Estimating Value at Risk (VaR) using Filtered Historical Simulation in the Indian capital market. *Reserve Bank of India Occasional Papers*, 32(2), 81-98, **2011**.

[29] Omari, C. O., A Comparative Performance of Conventional Methods for Estimating Market Risk Using Value at Risk. *International Journal of Econometrics and Financial Management*, 5(2), 22-32, **2017.**

[30] Artzner, P., Delbaen, F., Eber, J. M. and Heath, D., Coherent measures of risk. *Mathematical Finance*, 9(3), 203-228, **1999**.

[31] Hull, J. (2007). VAR vs Expected Shortfall, Risk, Available at: `http://www.risk.net/risk-magazine/technical-paper/1506669/var-versusexpected-shortfall>` (February, 2018).

[32] Yamai, Y. and Yoshiba, T., Comparative analyses of expected shortfall and value-at-risk: their estimation error, decomposition, and optimization, *Monetary and Economic Studies*, 20(1), 87-121, **2002**.

[33] Cont, R., Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance*, 1, 223-236, **2001**.

[34] Engle, R. F., Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica: Journal of the Econometric Society*, 987-1007, **1982**.

[35] Bollerslev, T., Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31(3), 307-327, **1986**.

[36] Nelson, D., Conditional heteroscedasticity in asset returns: a new approach. *Econometrica*, 59, 347-370, **1991**.

[37] Glosten, L. R., Jagannathan, R. and Runkle, D. E., On the relation between the expected value and the volatility of the nominal excess return on stocks, *The Journal of Finance*, 48(5), 1779-1801, **1993**.

[38] Ding, Z., Granger, C. W. and Engle, R. F., A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, 1(1), 83-106, **1993**.

[39] Engle, R. F. and Bollerslev, T., Modelling the persistence of conditional variances, *Econometric reviews*, 5(1), 1-50, **1986**.

[40] Azzalini, A., A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, 171-178, **1985**.

[41] Bollerslev, T., A conditionally heteroskedastic time series model for speculative prices and rates of return, *The Review of Economics and Statistics*, **69**, 542-547, **1987**.

[42] Azzalini, A. and Capitanio, A., Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2), 367-389, **2003**.

[43] Zhu, D. and Galbraith, J. W., A generalized asymmetric Student-t distribution with application to financial econometrics, *Journal of Econometrics*, 157(2), 297-305, **2010**.

[44] Aas, K. and Haff, I. H., The generalized hyperbolic skew student's t-distribution, *Journal of Financial Econometrics*, 4(2), 275-309, **2006**.

[45] Abramowitz, M. and Stegun, I. A., *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, (Vol. 9). Dover, New York, **1972**.

[46] Forsberg, L. and Bollerslev, T., Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model, *Journal of Applied Econometrics*, 17(5), 535-548, **2002**.

[47] Chen, Y. and Lu, J., *Value at risk estimation*, In Handbook of computational finance (pp. 307-333), Springer Berlin Heidelberg, **2012.**

[48] McDonald, J. B. and Newey, W. K., Partially adaptive estimation of regression models via the generalized t distribution, *Econometric Theory*, 4(03), 428-457, **1988**.

[49] Bollerslev T, Engle RF and Nelson, D.B., *ARCH models*, In Handbook of Econometrics, Engle RF, McFadden DL (eds.),Vol. 4, North-Holland: Amsterdam, pp.2959–3038, **1994**.

[50] Theodossiou, P., Financial data and the skewed generalized t distribution, *Management Science*, 44(12-part-1), 1650-1661, **1998**.

[51] Acitas, S., Senoglu, B. and Arslan, O., Alpha-skew generalized t distribution, *Revista Colombiana de Estadistica*, 38(2), 353-370, **2015**.

[52] Elal-Olivero, D., Alpha-skew-normal distribution, *Proyecciones (Antofagasta)*, 29(3), 224-240, **2010**.

[53] Piessens, R., de Doncker-Kapenga, E., Uberhuber, C. W. and Kahaner, D. K., *Quadpack: a subroutine package for automatic integration*, (Vol. 1), Springer Science & Business Media, **2012**.

[54] Henze, N., A probabilistic representation of the skew-normal distribution, *Scandinavian Journal of Statistics*, 13, 271-275, **1986**.

[55] Azzalini, A. and Dalla Valle, A., The multivariate skew-normal distribution, *Biometrika*, 83, 715-26, **1996**.

[56] Branco, M. D., Dey, D. K., A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis*, 79, 99-113, **2001**.

[57] Arnold, B.C. and Beaver, R.J., Skewed multivariate models related to hidden truncation and/or selective reporting, *Test*, 11(1), 7-54, **2002**.

[58] Arellano-Valle, R. B., Cortes, M. A. and Gomez, H. W., An extension of the epsilon-skew-normal distribution, *Communications in Statistics Theory and Methods*, 39(5), 912-922, **2010**.

[59] Nadarajah, S., A generalized normal distribution, *Journal of Applied Statistics*, 32(7), 685-694, **2005**.

[60] Balkema, A. A. and De Haan, L., Residual life time at great age, *The Annals of Probability*, 792-804, **1974**.

[61] Pickands, J., Statistical inference using extreme order statistics, *The Annals of Statistics*, 119-131, **1975**.

[62] Scarrott, C. and MacDonald, A., A review of extreme value threshold es-timation and uncertainty quantification, *REVSTAT-Statistical Journal*, 10(1), 33-60, **2012**.

[63] Hill, B. M., A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, 3(5), 1163-1174, **1975**.

[64] Coles, S., Bawa, J., Trenner, L. and Dorazio, P., *An introduction to statistical modeling of extreme values*, (Vol. 208), London, Springer, **2001**.

[65] DuMouchel, W. H., Estimating the stable index α in order to measure tail thickness: A critique, *The Annals of Statistics*, 1019-1031, **1983**.

[66] Ferreira, A., de Haan, L. and Peng, L., On optimising the estimation of high quantiles of a probability distribution, *Statistics*, 37(5), 401-434, **2003**.

[67] Loretan, M. and Phillips, P. C., Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets, *Journal of Empirical Finance*, 1(2), 211-248, **1994**.

[68] Reiss, R. D. and Thomas, M., *Statistical analysis of extreme values: with applications to insurance, finance, hydrology and other fields*, Birkhauser, Boston, **2007**.

[69] Kupiec, P. H., Techniques for verifying the accuracy of risk measurement models, *The Journal of Derivatives*, 3(2), 73-84, **1995**.

[70] Christoffersen, P. F., Evaluating interval forecasts, *International Economic Review*, 39, 841-862, **1998**.

[71] Lopez, J. A., Methods for evaluating value-at-risk estimates, *Economic Review-Federal Reserve Bank of San Francisco*, 2(2), 3-17, **1999**.

[72] Sarma, M., Thomas, S. and Shah, A., Selection of Value-at-Risk models. *Journal of Forecasting*, 22(4), 337-358, **2003**.

[73] Abad, P., Muela, S. B. and Martín, C. L., The role of the loss function in value-at-risk comparisons, *The Journal of Risk Model Validation*, 9(1), 1-19, **2015**.

[74] D'Agostino, R. B., Transformation to normality of the null distribution of g1, *Biometrika*, 679-681, **1970**.

[75] Anscombe, F. J. and Glynn, W. J., Distribution of the kurtosis statistic b 2 for normal samples, *Biometrika*, 70(1), 227-234, **1983**.

APPENDIX

APPENDIX 1: The Elements of Observed Information Matrix for ASGN and GAST Distributions

The elements of $J\left(\hat{\Theta}\right)$ for ASGN distribution are given by

$$J_{\alpha\alpha}=\sum_{i=1}^{n}\frac{4x_i^2\Gamma\left(1/\kappa\right)+2\alpha x_i\Gamma\left(3/\kappa\right)+2\Gamma\left(3/\kappa\right)\left(\alpha x_i-2\right)}{\left(\alpha^2x_i^2-2\alpha x_i+2\right)\left(\alpha^2\Gamma\left(3/\kappa\right)+2\Gamma\left(1/\kappa\right)\right)}$$
$$-\sum_{i=1}^{n}\frac{\left(2\alpha x_i^2-2x_i\right)\left(4x_i\Gamma\left(1/\kappa\right)\left(\alpha x_i-1\right)+2\alpha\Gamma\left(3/\kappa\right)\left(\alpha x_i-2\right)\right)}{\left(\alpha^2x_i^2-2\alpha x_i+2\right)^2\left(\alpha^2\Gamma\left(3/\kappa\right)+2\Gamma\left(1/\kappa\right)\right)}$$
$$+\sum_{i=1}^{n}\frac{2\alpha\Gamma\left(3/\kappa\right)\left(4x_i\Gamma\left(1/\kappa\right)\left(\alpha x_i-1\right)+2\alpha\Gamma\left(3/\kappa\right)\left(\alpha x_i-2\right)\right)}{\left(\alpha^2x_i^2-2\alpha x_i+2\right)\left(\alpha^2\Gamma\left(3/\kappa\right)+2\Gamma\left(1/\kappa\right)\right)^2}$$

$$J_{\alpha\kappa}=\sum_{i=1}^{n}\frac{-\kappa^{-2}\left[4x\Gamma\left(1/\kappa\right)\psi^{(0)}\left(1/\kappa\right)\left(\alpha x_i-1\right)+6\alpha\Gamma\left(3/\kappa\right)\psi^{(0)}\left(3/\kappa\right)\left(\alpha x_i-2\right)\right]}{\left(\alpha^2x_i^2-2\alpha x+2\right)\left(\alpha^2\Gamma\left(3/\kappa\right)+2\Gamma\left(1/\kappa\right)\right)}$$
$$+\sum_{i=1}^{n}\frac{-\kappa^{-2}\left(3\alpha^2\Gamma\left(3/\kappa\right)\psi^{(0)}\left(3/\kappa\right)+2\Gamma\left(1/\kappa\right)\psi^{(0)}\left(1/\kappa\right)\right)\left(4x_i\Gamma\left(1/\kappa\right)\left(\alpha x_i-1\right)+2\alpha\Gamma\left(3/\kappa\right)\left(\alpha x_i-2\right)\right)}{\left(\alpha^2x_i^2-2\alpha x+2\right)\left(\alpha^2\Gamma\left(3/\kappa\right)+2\Gamma\left(1/\kappa\right)\right)^2}$$

$$J_{\kappa\kappa} = \sum_{i=1}^{n} \frac{1}{\kappa}\Gamma\left(\frac{1}{\kappa}\right)\exp\left(|x_i|^\kappa\right)\left(\frac{\psi^{(0)}\left(\frac{1}{\kappa}\right)^2\exp\left(-|x_i|^\kappa\right)}{2\kappa^3\Gamma\left(\frac{1}{\kappa}\right)} - \frac{\psi^{(1)}\left(\frac{1}{\kappa}\right)^2\exp\left(-|x_i|^\kappa\right)}{2\kappa^3\Gamma\left(\frac{1}{\kappa}\right)} - \frac{\kappa\exp\left(-|x_i|^\kappa\right)|x_i|^\kappa\log^2\left(|x_i|\right)}{2\kappa\Gamma\left(\frac{1}{\kappa}\right)}\right)$$

$$+\sum_{i=1}^{n} \frac{1}{\kappa}\Gamma\left(\frac{1}{\kappa}\right)\exp\left(|x_i|^\kappa\right)\left(\frac{\kappa\exp\left(-|x_i|^\kappa\right)|x_i|^{2\kappa}\log^2\left(|x_i|\right)}{2\kappa\Gamma\left(\frac{1}{\kappa}\right)} - \frac{\exp\left(-|x_i|^\kappa\right)|x_i|^\kappa\log\left(|x_i|\right)}{\Gamma\left(\frac{1}{\kappa}\right)} - \frac{\psi^{(0)}\left(\frac{1}{\kappa}\right)\exp\left(-|x_i|^\kappa\right)|x_i|^\kappa\log\left(|x_i|\right)}{2\kappa^3\Gamma\left(\frac{1}{\kappa}\right)}\right)$$

$$-\sum_{i=1}^{n} \frac{2\Gamma\left(\frac{1}{\kappa}\right)\psi^{(0)}\left(\frac{1}{\kappa}\right)^2\exp\left(|x_i|^\kappa\right)\left(\frac{\exp(-|x_i|^\kappa)}{2\Gamma\left(\frac{1}{\kappa}\right)} - \frac{\kappa\exp(-|x_i|^\kappa)|x_i|^\kappa\log(|x_i|)}{2\Gamma\left(\frac{1}{\kappa}\right)} + \frac{\psi^{(0)}\left(\frac{1}{\kappa}\right)\exp(-|x_i|^\kappa)}{2\kappa\Gamma\left(\frac{1}{\kappa}\right)}\right)}{\kappa^3}$$

$$-\sum_{i=1}^{n} \frac{2\Gamma\left(\frac{1}{\kappa}\right)\exp\left(|x_i|^\kappa\right)\left(\frac{\exp(-|x_i|^\kappa)}{2\Gamma\left(\frac{1}{\kappa}\right)} - \frac{\kappa\exp(-|x_i|^\kappa)|x_i|^\kappa\log(|x_i|)}{2\Gamma\left(\frac{1}{\kappa}\right)} + \frac{\psi^{(0)}\left(\frac{1}{\kappa}\right)\exp(-|x_i|^\kappa)}{2\kappa\Gamma\left(\frac{1}{\kappa}\right)}\right)}{\kappa^2}$$

$$+\sum_{i=1}^{n} \frac{2\Gamma\left(\frac{1}{\kappa}\right)\exp\left(|x_i|^\kappa\right)|x_i|^\kappa\log\left(|x_i|\right)\left(\frac{\exp(-|x_i|^\kappa)}{2\Gamma\left(\frac{1}{\kappa}\right)} - \frac{\kappa\exp(-|x_i|^\kappa)|x_i|^\kappa\log(|x_i|)}{2\Gamma\left(\frac{1}{\kappa}\right)} + \frac{\psi^{(0)}\left(\frac{1}{\kappa}\right)\exp(-|x_i|^\kappa)}{2\kappa\Gamma\left(\frac{1}{\kappa}\right)}\right)}{\kappa}$$

$$+n\frac{\left(\frac{\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(0)}\left(\frac{1}{\kappa}\right)}{\kappa^2\Gamma\left(\frac{1}{\kappa}\right)} - \frac{3\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(0)}\left(\frac{3}{\kappa}\right)}{\kappa^2\Gamma\left(\frac{1}{\kappa}\right)}\right)^2}{\left(\frac{\alpha^2\Gamma\left(\frac{3}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)}+2\right)^2}$$

$$-\sum_{i=1}^{n} \frac{1}{\left(\frac{\alpha^2\Gamma\left(\frac{3}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)}+2\right)}\left(\frac{\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(0)}\left(\frac{1}{\kappa}\right)^2}{\kappa^4\Gamma\left(\frac{1}{\kappa}\right)} + \frac{9\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(0)}\left(\frac{3}{\kappa}\right)^2}{\kappa^4\Gamma\left(\frac{1}{\kappa}\right)} - \frac{6\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(0)}\left(\frac{1}{\kappa}\right)\psi^{(0)}\left(\frac{3}{\kappa}\right)}{\kappa^4\Gamma\left(\frac{1}{\kappa}\right)}\right)$$

$$-\sum_{i=1}^{n} \frac{1}{\left(\frac{\alpha^2\Gamma\left(\frac{3}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right)}+2\right)}\left(\frac{\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(1)}\left(\frac{1}{\kappa}\right)}{\kappa^4\Gamma\left(\frac{1}{\kappa}\right)} - \frac{9\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(1)}\left(\frac{3}{\kappa}\right)}{\kappa^4\Gamma\left(\frac{1}{\kappa}\right)} + \frac{2\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(0)}\left(\frac{1}{\kappa}\right)}{\kappa^4\Gamma\left(\frac{1}{\kappa}\right)} - \frac{6\alpha^2\Gamma\left(\frac{3}{\kappa}\right)\psi^{(0)}\left(\frac{3}{\kappa}\right)}{\kappa^4\Gamma\left(\frac{1}{\kappa}\right)}\right)$$

The elements of $J\left(\hat{\Theta}\right)$ for GAST distribution are given by

$$J_{\alpha\alpha} = \sum_{i=1}^{n} \frac{\left(\alpha x_i^2-2\right)^2 c_\alpha(\alpha,\lambda,\upsilon)^2 - \left(\alpha x_i^2-2\right)^2 c\left(\alpha,\lambda,\upsilon\right)c_{\alpha,\alpha}\left(\alpha,\lambda,\upsilon\right) + x_i^4\left(-c(\alpha,\lambda,\upsilon)^2\right)}{\left(2-\alpha x_i^2\right)^2 c(\alpha,\lambda,\upsilon)^2}$$

$$J_{\alpha\lambda} = n\frac{c_\lambda\left(\alpha,\lambda,\upsilon\right)c_\alpha\left(\alpha,\lambda,\upsilon\right) - c\left(\alpha,\lambda,\upsilon\right)c_{\alpha,\lambda}\left(\alpha,\lambda,\upsilon\right)}{c(\alpha,\lambda,\upsilon)^2}$$

$$J_{\alpha\upsilon} = n\frac{c_\upsilon\left(\alpha,\lambda,\upsilon\right) - c_\alpha\left(\alpha,\lambda,\upsilon\right) - c\left(\alpha,\lambda,\upsilon\right)c_{\alpha,\upsilon}\left(\alpha,\lambda,\upsilon\right)}{c(\alpha,\lambda,\upsilon)^2}$$

$$J_{\lambda\lambda} = n\frac{c_\lambda(\alpha,\lambda,\upsilon)^2 - c\left(\alpha,\lambda,\upsilon\right)c_\lambda\left(\alpha,\lambda,\upsilon\right)}{c(\alpha,\lambda,\upsilon)^2}$$

$$+\sum_{i=1}^{n}\frac{\left(1+\upsilon\right)x_i^2\left(T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)t'\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)-t\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)^2\right)}{\left(\upsilon+x_i^2\right)T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)^2}$$

$$J_{\lambda\upsilon} = \frac{c_\upsilon\left(\alpha,\lambda,\upsilon\right)c_\lambda\left(\alpha,\lambda,\upsilon\right) - c\left(\alpha,\lambda,\upsilon\right)c_{\lambda,\upsilon}\left(\alpha,\lambda,\upsilon\right)}{c(\alpha,\lambda,\upsilon)^2}$$

$$+\sum_{i=1}^{n}\frac{\lambda x_i^2\left(\frac{1}{\upsilon+x_i^2}-\frac{1+k}{\left(\upsilon+x_i^2\right)^2}\right)t'\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}{2T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}$$

$$-\sum_{i=1}^{n}\frac{\lambda x_i^2\left(\frac{1}{\upsilon+x_i^2}-\frac{1+k}{\left(\upsilon+x_i^2\right)^2}\right)t\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)^2}{2T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)^2}$$

$$+\sum_{i=1}^{n}\frac{x_i\left(\frac{1}{\upsilon+x_i^2}-\frac{1+k}{\left(\upsilon+x_i^2\right)^2}\right)t\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}{2\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}$$

$$J_{\upsilon\upsilon} = \frac{c_\upsilon(\alpha,\lambda,\upsilon)^2 - c\left(\alpha,\lambda,\upsilon\right)c_{\upsilon,\upsilon}\left(\alpha,\lambda,\upsilon\right)}{c(\alpha,\lambda,\upsilon)^2} + \frac{t_{\upsilon^2}\left(x_i;\upsilon\right)}{t\left(x_i;\upsilon\right)} - \frac{t_\upsilon(x_i;\upsilon)^2}{t(x_i;\upsilon)^2}$$

$$+\sum_{i=1}^{n}\frac{\lambda^2 x_i^2\left(\upsilon+x_i^2\right)\left(\frac{1}{\upsilon+x_i^2}-\frac{1+k}{\left(\upsilon+x_i^2\right)^2}\right)^2 t'\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}{4\left(1+\upsilon\right)T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}$$

$$-\sum_{i=1}^{m}\frac{\lambda^2 x_i^2\left(\upsilon+x_i^2\right)\left(\frac{1}{\upsilon+x_i^2}-\frac{1+k}{\left(\upsilon+x_i^2\right)^2}\right)^2 t\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)^2}{4\left(1+\upsilon\right)T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)^2}$$

$$-\sum_{i=1}^{n}\frac{\lambda x_i\left(\upsilon+x_i^2\right)\left(\frac{1}{\upsilon+x_i^2}-\frac{1+k}{\left(\upsilon+x_i^2\right)^2}\right)^2 t\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}{4\left(\frac{1+\upsilon}{\upsilon+x_i^2}\right)^{3/2}T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}$$

$$+\sum_{i=1}^{n}\frac{\lambda x_i\left(\frac{2(1+\upsilon)}{\left(\upsilon+x_i^2\right)^3}-\frac{2}{\left(\upsilon+x_i^2\right)^2}\right)^2 t\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}{2\sqrt{\frac{1+\upsilon}{\upsilon+x_i^2}}T\left(\sqrt{\frac{\upsilon+1}{x_i^2+\upsilon}}\lambda x_i;\upsilon+1\right)}$$

APPENDIX 2: The Used Mathematical Functions

1. Gamma function

$$\Gamma\left(z\right) = \int_0^\infty x^{z-1}e^{-x}dx$$

2. Beta function

$$B\left(\alpha,\beta\right) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

3. Modified bessel function of the third kind of order j

$$K_j\left(x\right) = \begin{cases} \pi\cos\left(\pi x\right)\left[I_{-j}\left(x\right) - I_j\left(x\right)\right], \text{if} j \notin \mathrm{Z} \\ \lim\limits_{\mu\to j} K_\mu\left(x\right), \text{if} j \in \mathrm{Z} \end{cases}$$

where $I_j(\cdot)$ denotes the modified Bessel function of the first kind of order j defined by

$$I_j(x) = \sum_{k=0}^{\infty} \frac{1}{\Gamma(k+j+1)\, k!} \left(\frac{x}{2}\right)^{2k+j}$$

4. Polygamma function

$$\psi^n(x) = \frac{\partial^n}{\partial z^n} \frac{\Gamma'(z)}{\Gamma(z)}$$

CURRICULUM VITAE

Credentials

Name–Surname	Emrah ALTUN
Place of Birth	Artvin
Marital Status	Married
E-mail	emrahaltun123@gmail.com

Education

BSc.	Gazi University, BSc. in Statistics
MSc.	Hacettepe University, MSc. in Statistics
PhD.	Hacettepe University, PhD. in Statistics

Foreign Languages

English

Work Experience

2012 September-2012 December	Research Assistant at Department of Econometrics, Trakya University
2012 December-2018 January	Research Assistant at Department of Statistics, Hacettepe University
2018 January-Present	Research Assistant at Department of Econometrics, Trakya University

Areas of Experience

-

Projects and Budgets

-

Publications Produced from Thesis

Altun, E., Tatlidil H., Ozel, G.	Conditional ASGT-GARCH approach to Value-at-Risk, Iranian Journal of Science and Technology, Transactions A: Science, DOI: https://doi.org/10.1007/s40995-018-0484-1.
Altun, E., Tatlidil H., Ozel, G., Nadarajah, S.	Does the assumption on innovation process play an important role for Filtered Historical Simulation model?, Journal of Risk and Financial Management, DOI: https://doi.org/10.3390/jrfm11010007.
Altun, E., Tatlidil H., Ozel, G., Nadarajah, S.	A New Generalization of Skew-T Distribution with Volatility Models, Journal of Statistical Computation and Simulation, DOI: https://doi.org/10.1080/00949655.2018.1427240.

Oral and Poster Presentations Produced from Thesis

Altun, E.	A New Skew Extension of Generalized Normal Distribution with Volatility Model, XVIIth INTERNATIONAL SYMPOSIUM ON ECONOMETRICS, OPERATIONS RESEARCH AND STATISTICS, 2017, Trabzon, Turkey.

HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING
THESIS/DISSERTATION ORIGINALITY REPORT

HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING
TO THE DEPARTMENT OF STATISTICS

Date: 06/03/2018

Thesis Title / Topic: THE IMPORTANCE OF FAT-TAILED AND SKEWED DISTRIBUTIONS IN MODELING VALUE-AT-RISK

According to the originality report obtained by myself/my thesis advisor by using the *Turnitin* plagiarism detection software and by applying the filtering options stated below on 06/03/2018 for the total of 115 pages including the a) Title Page, b) Introduction, c) Main Chapters, d) Conclusion sections of my thesis entitled as above, the similarity index of my thesis is 9%.

Filtering options applied:

1. Bibliography/Works Cited excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Sciene and Engineering Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

06/03/2018

Name Surname: EMRAH ALTUN

Student No: N13243234

Department: STATISTICS

Program: STATISTICS

Status: ☐ Masters ☒ Ph.D. ☐ Integrated Ph.D.

ADVISOR APPROVAL

APPROVED.

PROF. DR. HÜSEYİN TATLIDİL