Bayesian Inference for the Global Minimum Variance Portfolio

Author: **Muneeb Asif**

Abstract

In this study work we evaluate the weights of global minimum portfolio by using the posterior distributions which are given in Bordnar et al. (2017). The priors which are used to derive these posterior densities for the weights of global minimum portfolio are diffuse, conjugate, Jeffrey, and informative. To achieve this purpose, the weekly logarithmic return data of 5 international stocks have been used. For the performance measure mean, variance, 95 and 99 percent credible intervals have been computed. These posterior densities are then plotted for all the stock indices used in the study.

Keywords: Global minimum variance portfolio, Portfolio weights , Posterior distribution.

Acknowledgements

First of all, I am thankful to my supervisor **Stepan Mazur** whose expertise, guidance and support made it possible for me to work on this thesis.

I would also like to extend my thanks for my examiner, **Nicklas Pettersson** for his thoughts and feedback which helped me to improve this thesis more.

I would also like to express my gratitude towards my **Parents** for their encouragement which facilitated me to complete this thesis on time.

Contents

1 Introduction

For the investment purpose, an investor should not just rely on the expected profit rate. He/She should be able to tolerate a higher risk to gain a higher return. The potential return or profit increases with an increase in risk. Low potential returns are related with low levels of risk whereas, high potential returns are related with high levels of risk. Consequently, analysis of stock prices, measures the investment risk in the capital market, and the selected combination of stocks in an asset portfolio plays a significant role. The composition of investments in a portfolio depends on many factors such as the risk tolerance of the investor, the amount invested etc. For a beginner investor with limited funds, mutual funds or exchange-traded funds may be the right portfolio investments. For a professional investor with high net worth, portfolio investments may consist of stocks, bonds, commodities and rental properties.

1.1 Portfolio

Portfolio is combination of financial assets such as stocks, bonds, currencies, commodities and cash. These are held directly by investors as well as supervise by financial professionals and money managers. The selection of portfolio is concerned with the division of the investor's assets. This held between the different kinds of financial safeties so that the total return of the portfolio can be optimized. The fundamentals of modern portfolio theory introduced by Markowitz (1952). He presented a fundamental base for building of portfolio in a single period, where the risk of a portfolio was measured by the variance of its return and the profit was measured by the expected return. In his paper, the procedure of mean-variance portfolio optimization was presented which became much popular in the literature. This theory enables us to find the optimal portfolio weights which ensure the minimal risk for the given expected portfolio return.

Further Markowitz (1952) expresses, how to minimize the variance of portfolio which subject to the constraint where the expectation of portfolio equals to the suggested level of expected returns. This type of optimal portfolio is called the variance minimizing. In the similar way, if it also attains the maximum expected return between all the portfolios having the same variance of return, then it called an efficient portfolio. This tool is still using of both beginners and researchers in the field of finance nowadays.

1.2 Portfolio Weights

The weight of portfolios is the percentage of an investment portfolio that is held by a single asset. It can be used as influential investment tool. The portfolio weight helps the investor to tell how much money should be invested in a particular financial asset. There are different approaches for allocation of the portfolio weights. The most used and basic method of weight calculation is the value calculation, in which the calculation for portfolio weight is computed by dividing the value of the entire portfolio. Other methods include sectors, unit and cost, and type of securities.

1.3 Expected Returns

Expected return of portfolio is the amount that investor receives on an investment. By utilizing the expected return, investor can determine that whether an investment portfolio has positive

or negative average net income. To calculate the expected return, weights of portfolio and rate of return should be known. Rate of return is the percentage of profit from an investment over a certain time period. Expected return is then calculated by taking average of weights of all possible returns and multiply it by the rate of return.

1.4 Simple Returns and Log Returns

In time series, we use the log return of the prices instead of taking simple return. Usually simple returns are denoted with R whereas, log returns with r. These are defined as:

$$R_t = (P_t - P_{t-1})/P_{t-1} = P_t/P_{t-1} - 1$$

$$r_t = \ln[1 + R_t] = \log\left[\frac{P_t}{P_t - 1}\right] = \log(P_t) - \log(P_{t1})$$

$$V_t = \log(1 + R_t) \approx R_t$$

where P_t is the price of the asset at time t. We are defining the return from time t-1 to time t. The log function used in the equation is the natural logarithm. The measures also are equal when $P(t) = P(t - 1)$ since $R(t) = r(t) = 0$.

1.5 Minimum Variance Theory

It is well known theory presented by Markowitz (1952) for the portfolio optimization. For implementation of this theory, estimates of expected asset returns along with the corresponding variances and covariances needs to be estimated. If these parameter estimates are based only on the information of time series, then the resulted portfolio may likely to far away from the optimal. The portfolio which minimizes the portfolio return variance only with respect to the budget constraint is called the global minimum-variance (GMV) portfolio (Frahm & Memmel (2010)).

1.6 Efficient Frontier Curve

The efficient frontier is the set of optimal portfolios which offers the highest expected return for the lowest risk for a given level of expected return. It signifies the relationship between standard deviation and expected return. Portfolios which are located to the right of the efficient frontier are sub-optimal because they have a higher risk for the specific rate of return. Similarly, Portfolios that are located below the efficient frontier are also sub-optimal since they do not provide appropriate return for the specific level of risk. The tangent line on the efficient frontier curve represents the most efficient portfolio. The concept of efficient frontier is cornerstone of modern portfolio theory and was introduced by Markowitz (1952). Figure 1 illustrates the efficient frontier curve. R_f represents the risk free asset.
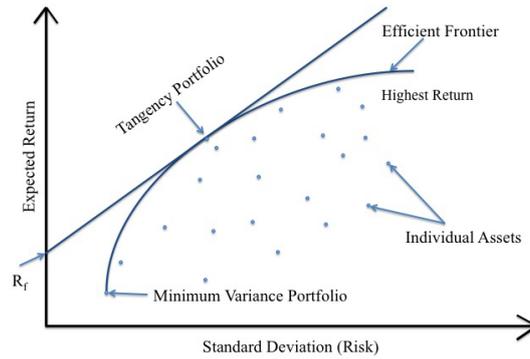
Figure 1: Efficient Frontier Curve

1.7 Purpose

The purpose of this study is to evaluate the weights of GMV portfolio by using the posterior distributions which are given in Bodnar et al. (2017). To achieve this objective, the logarithm data of 5 stock indices have been applied on the given posterior distributions.

1.8 Outline

Section 2 is about the literature review which explains the previous researches about constructing the portfolios and assign them the weights. In section 3, method is defined which describes the whole procedure of this study work. Section 4 is about the data, which explains the source and structure of data. Section 5 explains the theory and models which clarifies the prior and posterior distributions which are utilized in this study work. In section 6, results and outputs are presented for the posterior distributions which are utilized in section 5. Finally, section 7 explains the discussion and conclusion of the results and whole study.

2 Literature Review

The tangency and other portfolios often lead to investment approaches with unsure profits and high risk. Bauder et al. (2017) consider the estimation of the weights of tangent portfolios from Bayesian perspective. By using mean vector and covariance matrix for diffuse and conjugate prior, they formulate the stochastic representation for posterior distributions of the weights of tangent portfolios and their linear combinations. The results were estimated within the numerical study where the coverage probabilities of credible intervals have been measured.

The implementation of the (Markowitz (1952)) theory involves the knowledge of the distribution parameters of the stock returns. These parameters should be estimated, not to be observed. Usually time series data has been used to find out these distribution parameters. Generally expected return are difficult to estimate, as these returns mainly control the composition of optimal portfolio and errors in the expected returns. Then these errors may cause to portfolios that are distant from the real optimal portfolio. Kempf & Memmel (2006a) suggests a two-step approach to determine the optimal portfolio weights. In the first step one estimates the

parameters for return distribution. In the second stage, portfolio weights using the estimated parameters can be optimized.

In (Bodnar et al. (2017)), the estimation of the weights of the optimal portfolio and global minimum variance of the portfolio have been considered by Bayesian perspective. This system enable us to integrate prior beliefs of the investors and to integrate these into the decisions of portfolio. The standard priors for the mean vector and the covariance matrix have also been utilized to find out the posterior distribution. Furthermore, the posterior distributions of the portfolio weights are derived in explicit form for all models. By using the coverage probabilities of credible intervals, these models then compared with each other. According to model transformation, a prior directly for the portfolio weights has been suggested. The results of empirical study presented good results for the suggested priors.

Modern portfolio theory suggests a positive relationship between risk and expected returns. Maillet et al. (2015) has conducted a research to compute the global minimum variance portfolio by using sample covariance matrix. Monthly data of 49 industry portfolios represents the US stock market, has been used for this purpose. In this way, they suggest a robust approach to temperate the impact of parameter uncertainty. Further, using Monte Carlo simulations in the presence of parameter uncertainty has also been used which controls the stability of portfolio weight, portfolio variance and risk-adjusted returns. They also provide a data-adaptive method to calibrate the uncertainty coefficient.

Standard portfolio theory suggests that the tangency portfolio is the only efficient stock portfolio. Whereas, many empirical studies proves that an investment in the global minimum variance portfolio often yields towards the better results than does an investment in the tangency portfolio. Kempf & Memmel (2006b) conducted a study to show that the weights of the global minimum variance portfolio are equal to regression coefficients. For this purpose, they use the ordinary least square method. In this way, they derive the conditional distributions of the estimated portfolio weights and estimated return parameters. Conditional distributions are essential for analyzing the global minimum variance portfolio and also to estimate the portfolio weights.

3 Method

This section defines the method and statistical techniques which is used in this study work.

This study work is based on (Bodnar et al. (2017)) in which the global minimum variance portfolio within a Bayesian framework have been analyzed. The logarithm data of 5 stocks indices have been used for this purpose. The data is splits into 2 equal parts for the prior and working data. This data is then applied on the posterior distributions which are derived in (Bodnar et al. (2017). The priors which are used to compute these posterior distributions are diffuse, conjugate, Jeffrey, and informative. These distributions are then plotted. For the performance measure mean, variance, 95, and 99 percent credible intervals of these posterior distributions are then compared. R statistical package is used for the analysis purpose (R Core Team (2013)).

4 Data

Data which is used in this for the analysis purpose is downloaded from the yahoo finance website site (www.finance.yahoo.com) which is open and free source for the financial data.Weekly data for 5 stock indices has been used. The reason for take the weekly data is to achieve normality. The period of data is from 01 January 2007 to 31 march 2018. The total number of observations are 586. The data is then splitting into 2 parts. The first 293 observations from the period 01/01/2007 to 12/08/2012 are used as prior data whereas, the remaining 293 observations from the period 19/08/2012 to 31/01/2018 are considered as the working data. The 5 stock indices are DAX, S&P 500, OMX 30, CAC40, and Nikkei 225. The details of these stock indices are given below.

Stock indices	Description
DAX	German stock index which is consists of 30 major German companies trading on the Frankfurt stock exchange.
S&P 500	American stock market index based on the market capitalization of 500 large companies.
OMX 30	Swedish stock market index traded for the Stockholm stock exchange which consists of 30 most-traded stock classes.
CAC 40	French stock market index represents a capitalization-weighted measure of the 40 most significant values among the 100 highest market caps on the Euronext Paris.
Nikkei 225	Stock market index for the Tokyo stock exchange. It is most widely quoted average of Japanese equities.

Table 1: Description of Stock Indices

5 Theory and Models

This section explains the theory and concepts of statistical approaches which have been used in this study.

5.1 Bayesian Portfolio Selection

There are two school of thoughts in statistical inference i.e. frequentist and Bayesian. Both approaches allow one to evaluate evidence about competing hypotheses. The Bayesian school makes inferences which depends on prior and likelihood of observed data whereas, frequentist school does not depend on prior and may vary from one investigator to another.

In Bayesian, for a model parameter θ, first we have to state the prior knowledge about θ as a probability distribution $p(\theta)$. Then collect the data and form the likelihood function $p(Data|\theta)$. Following is the Bayes theorem for a model parameter θ

$$p(\theta|Data) = \frac{p(Data|\theta)p(\theta)}{p(Data)}. \tag{1}$$

The prior $p(\theta)$ is the function that converts the likelihood function $p(Data|\theta)$ into a posterior probability density $p(\theta|Data)$. p(Data) is just a constant that integrates $p(\theta|Data)$ to one. In this way, eq (1) becomes as following

$$p(\theta|Data) \propto p(Data|\theta)p(\theta)$$

or

$$Posterior \propto Likelihood \times Prior.$$

Let $\mathbf{X}_i = (X_{1i}, X_{2i}, ..., X_{ki})^T$ be the k-dimensional random vector of log returns at time $i = 1, ..., n$. Let $\mathbf{w} = (w_1, w_2, ..., w_k)^T$ be the vector of portfolio weights, where w_j denotes the weight of jth asset and $\mathbf{1}$ be the vector of ones. Let the mean vector of asset returns is denoted by $\boldsymbol{\mu}$. According to Bodnar et al. (2017), GMV portfolio is the unique solution of optimization problem which is given as

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^T\boldsymbol{\Sigma}^{-1}\mathbf{1}}, \tag{2}$$

where $\boldsymbol{\Sigma}$ is the positive definite covariance matrix. The diagonal elements of this matrix are the variances of returns. $\boldsymbol{\Sigma}$ matrix is given as following

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k1} & \cdots & \sigma_{kk} \end{pmatrix},$$

where $\sigma_{ij} = cov(x_i, x_j)$.

Short selling allows trader to borrow and sell a stock without actually owning it. Generally, short sales are allowed therefore equation (2) also deals with the negative weights. In (2) $\boldsymbol{\Sigma}$ is unknown parameter therefore, equation (2) becomes infeasible to use for the practical solution. Therefore, the equation (2) can be written as

$$\widehat{\mathbf{w}}_{GMV} = \frac{\mathbf{S}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1}}. \tag{3}$$

In equation (3) $\mathbf{S}$ represents the sample covariance matrix which can be compute as $\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x_i} - \bar{\mathbf{x}})(\mathbf{x_i} - \bar{\mathbf{x}})^T$ and $\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i$. In (Bodnar et al. (2017)) arbitrary linear combinations of the GMV portfolio weights have been considered. Let $\mathbf{L}$ is an arbitrary $p \times k$ matrix of constants where rank is $p < k$.

$$\boldsymbol{\theta} = \mathbf{L}\mathbf{w}_{GMV} = \frac{\mathbf{L}\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^T\boldsymbol{\Sigma}^{-1}\mathbf{1}} \tag{4}$$

Whereas, sample estimator of $\boldsymbol{\theta}$ is following

$$\widehat{\boldsymbol{\theta}} = \mathbf{L}\widehat{\mathbf{w}}_{GMV} = \frac{\mathbf{L}\mathbf{S}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1}}. \tag{5}$$

According to Bodnar et al. (2017) $\overline{\mathbf{x}}$ and $\mathbf{S}$ diverges from their true parameter i.e. $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ therefore, $\mathbf{w}_{GMV}$ and $\boldsymbol{\theta}$ should be considered as random quantity.

5.2 Prior and Posterior Distribution for GMV Portfolio Weights

This section is based on Bodnar et al. (2017). In Bodnar et al. (2017), three priors have been utilized such as diffuse prior, conjugate prior, and hierarchical prior. These priors are then used to derive the posterior distribution of linear combination of portfolio weights. Then the reparametrized model for the asset returns have been measured to derive the informative and non-informative prior. For the corresponding priors, the posterior distributions for portfolio weights have also been derived. In this way, they provide the whole distribution of the optimal portfolio weights and their point estimator. Usually, only the point estimator is given for the optimal portfolio weights not the distributions. In this study work diffuse, conjugate, Jeffreys non-informative, and informative prior have been used for the empirical analysis.

5.2.1 Diffuse Prior

Diffuse prior is generally a non-informative prior which implies that prior has minimal impact on the posterior distribution of θ. In this way, stochastic nature of unknown parameters has no additional information and Bayesian estimator for the portfolio weights affects by this uncertainty. Here, we take the prior for mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ as we get the posterior for weights. The prior densities of diffuse prior is following

$$p_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto | \boldsymbol{\Sigma} |^{-\frac{k+1}{2}}.$$

Let $\mathbf{X}_1, ..., \mathbf{X}_n \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}$ be independently and identically distributed with $\mathbf{X}_1 \mid \boldsymbol{\mu}, \boldsymbol{\Sigma} \sim \boldsymbol{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\mathbf{L}$ be a $p \times k$ matrix of constants with rank is $p < k$ and $\mathbf{1}$ denotes the vector of ones. According to Bodnar et al. (2017), the posterior for $\boldsymbol{\theta}$ under the diffuse prior $p_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is then given by

$$\boldsymbol{\theta} \mid \mathbf{X}_1, ..., \mathbf{X}_n \sim t_p\left(n - 1; \widehat{\boldsymbol{\theta}}; \frac{1}{n-1}\frac{\mathbf{L}\mathbf{R}_d\mathbf{L}^T}{\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1}}\right) \tag{6}$$

where $\mathbf{R}_d = \mathbf{S}^{-1} - \mathbf{S}^{-1}\mathbf{1}\mathbf{1}^T\mathbf{S}^{-1}/\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1}$.

Equation (6) shows that posterior distribution for the linear combinations of GMV portfolio weights under the diffuse prior follows multivariate t-distribution.

5.2.2 Conjugate Prior

In Bayesian statistical inference, if the prior probability distribution is belong to same probability distribution family as posterior distribution, then prior and posterior are said to be conjugate distributions and the prior is called the conjugate prior. It is an informative prior and which

reflects a normal prior for $\boldsymbol{\mu}$ (conditional on $\boldsymbol{\Sigma}$) and inverse Wishart prior for $\boldsymbol{\Sigma}$. In this way, investor has prior information regarding the targeted values of expected returns and covariance matrix. Usually, this is defined as

$$p_c(\boldsymbol{\mu} \mid \boldsymbol{\Sigma}) \propto \mid \boldsymbol{\Sigma} \mid^{-\frac{1}{2}} \exp\left\{ -\frac{\kappa_c}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_c)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_c) \right\}$$

and

$$p_c(\boldsymbol{\Sigma}) \propto \mid \boldsymbol{\Sigma} \mid^{-\frac{\nu_c}{2}} \exp\left\{ -\frac{1}{2}tr[\mathbf{S}_c\boldsymbol{\Sigma}^{-1}] \right\},$$

where $\boldsymbol{\mu}_c$ is the prior mean, κ_c is the parameter which reflects the prior precision of $\boldsymbol{\mu}_c$. ν_c is also prior precision on $\boldsymbol{\Sigma}$ and $\mathbf{S}_c$ is a known prior matrix of $\boldsymbol{\Sigma}$. Therefore, the joint prior for both parameters is expressed as following

$$p_c(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto \mid \boldsymbol{\Sigma} \mid^{-\frac{\nu_c+1}{2}} \exp\left\{ -\frac{\kappa_c}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_c)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_{c)} - \frac{1}{2}tr[\mathbf{S}_c\boldsymbol{\Sigma}^{-1}] \right\}.$$

Let $\mathbf{X}_1, ..., \mathbf{X}_n \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}$ be independently and identically distributed with $\mathbf{X}_1 \mid \boldsymbol{\mu}, \boldsymbol{\Sigma} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\mathbf{L}$ be a $p \times k$ matrix of constants with rank is $p < k$ and $\mathbf{1}$ denotes the vector of ones. According to Bodnar et al. (2017), the posterior for $\boldsymbol{\theta}$ under the conjugate prior $p_c(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is following

$$\boldsymbol{\theta} \mid \mathbf{X}_1, ..., \mathbf{X}_n \sim t_p\left(v_c + n - k - 1; \frac{\mathbf{L}\mathbf{V}_c^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{V}_c^{-1}\mathbf{1}}; \frac{1}{V_c + n - k - 1}\frac{\mathbf{L}\mathbf{R}_c\mathbf{L}^T}{\mathbf{1}^T\mathbf{V}_c^{-1}\mathbf{1}} \right) \tag{7}$$

where
$\mathbf{r}_c = \frac{n\overline{\mathbf{X}} + k_c\boldsymbol{\mu}_c}{n + k_c}$,

$\mathbf{V}_c = (n-1)\mathbf{S} + \mathbf{S}_c + (n+k_c)\mathbf{r}_c\mathbf{r}_c^T + n\overline{\mathbf{X}\mathbf{X}}^T + k_c\boldsymbol{\mu}_c\boldsymbol{\mu}_c^T$,

$\mathbf{R}_c = \mathbf{V}_c^{-1} - \mathbf{V}_c^{-1}\mathbf{1}\mathbf{1}^T\mathbf{V}_\mathbf{c}^{-1}/\mathbf{1}^T\mathbf{V}_\mathbf{c}^{-1}\mathbf{1}$.

Equation (7) represents that posterior distribution for the linear combinations of GMV portfolio weights under the conjugate prior follows multivariate t-distribution.

5.2.3 Jeffrey's Prior

Jeffreys prior is non-informative prior distribution for a parameter space. For Jeffreys prior it has been assumed that there is no initial information about the weights is present. Its construction is based on the Fisher information function of a model. We have provided the prior for the weights directly. The Jeffreys prior for $(\boldsymbol{\theta}, \boldsymbol{\Psi}, \zeta)$ is given by

$$p_n(\boldsymbol{\theta}, \boldsymbol{\Psi}, \zeta) \propto \zeta^{\frac{p}{2}-1} \mid \boldsymbol{\Psi} \mid^{-p/2-1},$$

where $\zeta = \mathbf{1}^T\mathbf{\Sigma}^{-1}\mathbf{1}$ and $\mathbf{\Psi} = \mathbf{L}\mathbf{\Sigma}^{-1}\mathbf{L}^T - \frac{\mathbf{L}\mathbf{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\mathbf{\Sigma}^{-1}\mathbf{L}^T}{\mathbf{1}^T\mathbf{\Sigma}^{-1}\mathbf{1}}$.

Let $\mathbf{X}_1, ..., \mathbf{X}_n \mid \boldsymbol{\mu}, \mathbf{\Sigma}$ be independently and identically distributed with $\mathbf{X}_i \mid \boldsymbol{\mu}, \mathbf{\Sigma} \sim N_k(\boldsymbol{\mu}, \mathbf{\Sigma})$. Let $\mathbf{L}$ be a $p \times k$ matrix of constants with rank is $p < k$. Then posterior for the GMV portfoilio weights $\boldsymbol{\theta}$ under the Jeffreys non-informative prior $p_n(\boldsymbol{\theta}, \mathbf{\Psi}, \zeta)$ is the following

$$\boldsymbol{\theta} \mid \mathbf{X}_1, ..., \mathbf{X}_n \sim t_p\left(n - k + p; \widehat{\boldsymbol{\theta}}; \frac{1}{n - k + p}\frac{\mathbf{L}\mathbf{R}_d\mathbf{L}^T}{\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1}}\right). \tag{8}$$

Equation (8) represents that posterior for the GMV portfolio weights under Jeffreys non-informative prior $p_n(\boldsymbol{\theta}, \mathbf{\Psi}, \zeta)$ follows p-variate t-distribution with $n - k + p$ degrees of freedom, location vector $\widehat{\boldsymbol{\theta}}$ and dispersion matrix $\frac{1}{n-k+p}\frac{\mathbf{L}\mathbf{R}_d\mathbf{L}^T}{\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1}}$.

5.2.4 Informative Prior

In Bodnar et al. (2017), it has been attained under a hierarchical Bayesian model for the GMV weights. The model which is used is the multiple response model for counts, specified hierarchically, and belongs to the fully Bayesian family. This model is developed by Tunaru (2002). The proposed informative prior is following

$$\boldsymbol{\theta} \sim \mathbf{N}_p\left(\mathbf{w}_I, \tfrac{1}{\zeta}\mathbf{\Psi}^{-1}\right),$$

$$\mathbf{\Psi} \sim W_p(\nu_I, \mathbf{S}_I),$$

$$\boldsymbol{\zeta} \sim Gamma(\delta_1, 2\delta_2)$$

where $\mathbf{w}_I$ is the prior mean, ν_I is a prior precision parameter on $\mathbf{\Psi}$; $\mathbf{S}_I$ is the known matrix; δ_1 and δ_2 are prior constants. In this way, the joint prior is expressed as following

$$\begin{aligned}p_I(\boldsymbol{\theta}, \mathbf{\Psi}, \zeta) \propto{} & \left|\frac{1}{\zeta}\mathbf{\Psi}^{-1}\right|^{-\frac{1}{2}}\exp\left\{\frac{-\zeta}{2}(\boldsymbol{\theta} - \mathbf{w}_I)^T\mathbf{\Psi}(\boldsymbol{\theta} - \mathbf{w}_I)\right\} \\ & \times \zeta^{\delta_1 - 1}|\mathbf{\Psi}|^{(\nu_I - p - 1)/2}\exp\left\{-\frac{1}{2}tr[\mathbf{S}_I^{-1}\mathbf{\Psi}] - \frac{\zeta}{2\delta_2}\right\}.\end{aligned} \tag{9}$$

Integrating this prior for $\mathbf{\Psi}$ and ζ the posterior distribution of the portfolio weights can be obtained. The proof of the equation (9) can be found in Bodnar et al. (2017).

Let $\mathbf{X}_1, ..., \mathbf{X}_n \mid \boldsymbol{\mu}, \mathbf{\Sigma}$ be independently and identically distributed with $\mathbf{X}_i \mid \boldsymbol{\mu}, \mathbf{\Sigma} \sim N_k(\boldsymbol{\mu}, \mathbf{\Sigma})$. Let $\mathbf{L}$ be a $p \times k$ matrix of constants with rank is $p < k$. Then posterior for the GMV portfoilio weights $\boldsymbol{\theta}$ under the informative prior $p_n(\boldsymbol{\theta}, \mathbf{\Psi}, \zeta))$ is following

$$p_I(\boldsymbol{\theta} \mid (\mathbf{X}_1, ..., \mathbf{X}_n)) \propto [(\boldsymbol{\theta} - \mathbf{w}_I)^T (\mathbf{S}_I^{-1} + (n-1)(\mathbf{LR}_d\mathbf{L}^T)^{-1})^{-1}(\boldsymbol{\theta} - \mathbf{w}_I)]^{(n-k+2p+2\delta_1)/2}$$
$$\times U((n-k+2p+2\delta_1)/2; (p+2\delta_1 - \nu_1 + 1)/2; g(\boldsymbol{\theta})), \tag{10}$$

where $g(\boldsymbol{\theta}) = \frac{n-1}{2} \frac{((\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}})^T(\mathbf{LR}_d\mathbf{L}^T)^{-1}(\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}})+(\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1})^{-1}+\frac{\delta_2^{-1}}{n-1}}{(\boldsymbol{\theta}-\mathbf{w}_I)^T(\mathbf{S}_I^{-1}+(n-1)(\mathbf{LR}_d\mathbf{L}^T)^{-1})^{-1}(\boldsymbol{\theta}-\mathbf{w}_I)}$

and

$U(\cdot)$ stands for the confluent hypergeometric function (Abramowitz et al. (1972)).

In finance, stochastic models are use to represent the apparently random behavior of stocks. This is important tool to represent and analysis of the time series data. The stochastic representation for $\boldsymbol{\theta}$ under the informative prior expressed as

$$\boldsymbol{\theta} = \mathbf{r}_I(\tau) + \zeta^{-1/2}(\mathbf{V}_I(\tau))^{1/2}\mathbf{z}_0, \tag{11}$$

where
$\mathbf{z}_0 \sim N_p(\mathbf{0}_p, \mathbf{I}_p),$

$\zeta \mid \tau \sim Gamma\left((n-k+2p+2\delta_1)/2, \frac{2}{h_I(\tau)}\right),$

$\tau \sim Gamma((n-k+p+v_1-1)/2, 2),$

$\mathbf{V}_I(\tau) = (\tau\mathbf{P}_1 + \mathbf{P}_2)^{-1},$

$\mathbf{r}_I(\tau) = (\tau\mathbf{P}_1 + \mathbf{P}_2)^{-1}(\tau\mathbf{P}_1\mathbf{w}_I + \mathbf{P}_2\widehat{\boldsymbol{\theta}}),$

$\mathbf{h}_I(\tau) = r + \tau\mathbf{w}_I^T\mathbf{P}_1\mathbf{w}_I + \widehat{\boldsymbol{\theta}}^T\mathbf{P}_2\widehat{\boldsymbol{\theta}} - \mathbf{r}_I(\tau)^T(\mathbf{V}_I(\tau))^{-1}\mathbf{r}_I(\tau),$

$\mathbf{P}_1 = (\mathbf{S}_I^{-1} + (n-1)(\mathbf{LR}_d\mathbf{L}^T)^{-1})^{-1},$

$\mathbf{P}_2 = (n-1)(\mathbf{LR}_d\mathbf{L}^T)^{-1},$

$r = \delta_2^{-1} + (n-1)(\mathbf{1}^T\mathbf{S}^{-1}\mathbf{1})^{-1}.$

6 Results

This section consists of outputs in the form of table and plots which have been discussed in the section 5.

6.1 Empirical Results

There are 5 stock indices which are used in this study work i.e. DAX, S&P500, OMX, CAC, and Nikkei. Following matrices illustrates the global minimum variance portfolio weights for the prior and working data. The comparison is done for $p = 1$, $k = 5$, $\mathbf{L} = \mathbf{e}_i^T$ which is equal to basis vectors for i = 1, ... , 5. For informative prior we considered the $\delta_1 = 1$ and $\delta_2 = 0.5$, $\nu = n$, $w_I = 1/k$, $S_I = 1$. The values for $\delta_1 = 1$ and $\delta_2 = 0.5$ are suggested in Bodnar et al. (2017). Further mean, variance, 95 and 99 percent credible interval bounds have also been computed. Following we are presenting the mean and covariance matrix for the prior and working data.

$$\boldsymbol{\mu}_{prior} = \begin{pmatrix} 0.0224 & 0.0020 & -0.0170 & -0.1564 & -0.2127 \end{pmatrix} \tag{12}$$

$$\mathbf{S}_{prior} = \begin{pmatrix} 14.2818 & 10.5011 & 11.8746 & 13.3609 & 9.8050 \\ 10.5011 & 9.9739 & 9.2975 & 10.3961 & 8.0134 \\ 11.8746 & 9.2975 & 13.1195 & 12.0288 & 8.7128 \\ 13.3609 & 10.3609 & 12.0288 & 14.1629 & 9.7064 \\ 9.8050 & 8.0134 & 8.7128 & 9.7064 & 12.4069 \end{pmatrix} \tag{13}$$

$$\overline{\mathbf{X}} = \begin{pmatrix} 0.1847 & 0.2122 & 0.1164 & 0.1341 & 0.2903 \end{pmatrix} \tag{14}$$

$$\mathbf{S} = \begin{pmatrix} 5.9096 & 2.8303 & 4.1720 & 5.1366 & 3.9202 \\ 2.8303 & 2.6901 & 2.4360 & 2.7167 & 2.5517 \\ 4.1720 & 2.4360 & 4.2307 & 3.955 & 3.2909 \\ 5.1366 & 2.7167 & 3.9553 & 5.2338 & 3.9551 \\ 3.9202 & 2.5514 & 3.2909 & 3.9551 & 7.9849 \end{pmatrix} \tag{15}$$

For performance measure, we estimate the mean, variance, 95 and 99 percent credible interval of the 5 stock indices using all the priors. The results are summarized in the Table (2). There are 5 columns in the table. First column represents the index number and their respective estimates. Second column shows the estimates of diffuse prior. Third column illustrates the estimates under the conjugate prior. Similarly, fourth and fifth columns explains the estimates using Jeffrey's and informative priors respectively. For all the stock indices, the minimum variance is given by the informative prior.

6.2 Plots

We plot these densities for portfolio weights of stock indices. Following are the table and plots of these posterior densities.

.

	Diffuse	**Conjugate**	**Jeffrey**	**Informative**
DAX				
Mean	-0.2084	-0.2393	-0.2084	0.1992
Variance $\times 10^{-2}$	0.1007	0.5403	0.1192	0.0020
95% Lower Credible Interval	-0.4155	-0.3839	-0.4166	0.1905
95% Upper Credible Interval	-0.0012	-0.0946	-0.0001	0.2080
99% Lower Credible Interval	-0.4812	-0.4299	-0.4827	0.2012
99% Upper Credible Interval	0.0644	-0.0487	0.0658	0.2107
S&P 500				
Mean	0.9098	0.8965	0.9098	0.2012
Variance $\times 10^{-2}$	0.5109	0.2602	0.5163	0.0011
95% Lower Credible Interval	0.7691	0.7961	0.7684	0.1944
95% Upper Credible Interval	1.0505	0.9969	1.0521	0.2079
99% Lower Credible Interval	0.7245	0.7642	0.7235	0.1923
99% Upper Credible Interval	1.0951	1.0288	1.0961	0.2101
OMX				
Mean	0.2723	0.2868	0.2723	0.2001
Variance $\times 10^{-2}$	0.8221	0.4199	0.8307	0.0014
95% Lower Credible Interval	0.1316	0.1593	0.0929	0.1927
95% Upper Credible Interval	0.4130	0.4144	0.4517	0.2075
99% Lower Credible Interval	0.0870	0.1188	0.0360	0.1903
99% Upper Credible Interval	0.4577	0.4548	0.5086	0.2098
CAC				
Mean	0.0052	0.0394	0.0052	0.2012
Variance $\times 10^{-2}$	0.3902	0.6687	0.4047	0.0033
95% Lower Credible Interval	-0.1354	-0.1215	-0.2279	0.1907
95% Upper Credible Interval	0.1459	0.2003	0.2385	0.2117
99% Lower Credible Interval	-0.1800	-0.1726	-0.3020	0.1873
99% Upper Credible Interval	0.1906	0.2514	0.3125	0.2151
Nikkei				
Mean	0.0209	0.0165	0.0209	0.2012
Variance $\times 10^{-2}$	0.1867	0.0956	0.1886	0.0012
95% Lower Credible Interval	-0.1197	-0.0443	-0.0645	0.1946
95% Upper Credible Interval	0.1616	0.0773	0.1064	0.2078
99% Lower Credible Interval	-0.1644	-0.0637	-0.0917	0.1926
99% Upper Credible Interval	0.2062	0.0967	0.1335	0.2099

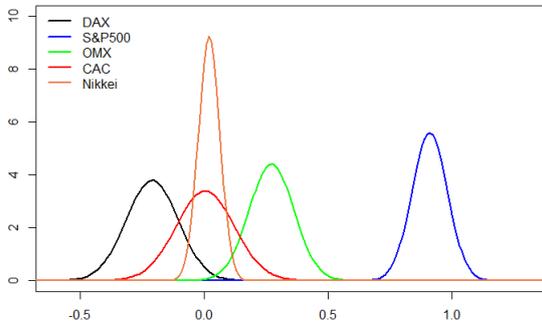Table 2: Mean, Variance, Credible Intervals of indices

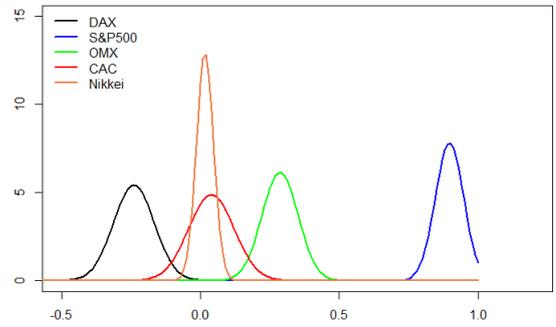Figure 2: Posterior densities: Diffuse Prior
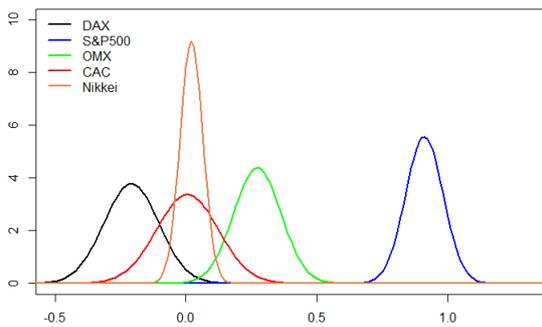

Figure 3: Posterior densities: Conjugate Prior

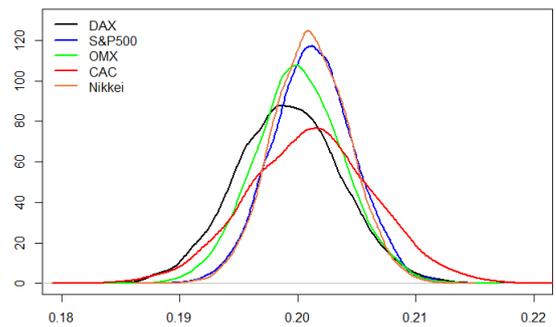Figure 4: Posterior densities: Jeffrey Prior

Figure 5: Posterior densities: Informative Prior

14

7   Discussion and Conclusion

This section explains the discussion of results for different priors within empirical study. we consider the weekly logarithmic returns of five international stock indices DAX, S&P 500, OMX, CAC, Nikkei for the period from 01.01.2007 till 31.01.2018 with 586 total number of observations. We consider the prior for asset returns and prior for portfolio weights. For asset returns, diffuse and conjugate priors have been used whereas for portfolio weights Jeffreys non-informative prior and informative prior have been used.

Mean, covariance matrix, and their corresponding global minimum variance portfolio weights are shown in (12) and (13). The input parameters in the prior distributions are $\boldsymbol{\mu}_c = \boldsymbol{\mu}_{prior}$, $\mathbf{S}_c = \mathbf{S}_{prior}$, $\mathbf{w}_I = \mathbf{w}_{prior}$. These parameters for the working sample can be observed in (14) and (15). As prior data covers the period of global financial crisis which was starting from 2010. This reflected in the estimated parameters and the average return in this crisis period are much lower and even negative for the three indices.

Table (2) summarizes the all mean, variance, 95 and 99 percent credible intervals of all the stock indices which are used in this study. In the case of DAX, except the informative prior, the all priors shows the negative weights. These are short sales. For short sales it is belief that a security's price will decline, enabling it to be bought back at a lower price to make a profit. However, short selling should only be used by experienced traders, who are familiar with the risks. In case of $S\&P\ 500$, there are no short sales given by any of the prior. Similarly, for the OMX, CAC, and $Nikkei$, no prior shows the negative weights. Outputs obtained diffuse and conjugate prior have almost same mean as both priors belong to non-informative. Overall, the best returns are given by the informative prior as informative prior has utilizes a lot of prior information. The weights obtained from the informative prior are very close to the equally weighted portfolio. One can use the sharp ratio to measure the risk of these portfolio.

Further, Figure (2), (3), (4) and (5) represents the plots of the posterior densities. In all the plots, the highest curve is obtained by the textitNikkei index. For $S\&P\ 500$, the curve is shifted to the right as compare to the other stock indices. In contrast to other posterior densities, all the stock indices curves for informative prior are located at the posterior centered of their sample weights.

Summary

In this study work we estimate the weights of global minimum portfolio by using the posterior distributions which are given in Bodnar et al. (2017). For this purpose, the weekly logarithm data of 5 international stocks indices have been used and applied on the posterior densities. To assess the posterior distribution, we divide the data into two equal parts, that is 293 observations for each. The first 293 observations are used for the prior data, whereas the remaining are considered as working data. Further mean, variance, 95 and 99 percent credible intervals have been computed to check the performance measure. The best results are obtained by the informative prior where the weights are close to the equally weighted portfolio. In the end, the posterior densities for portfolio weights are then plotted for all the stock indices.

References

Abramowitz, M., Stegun, I. A. et al. (1972), *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Vol. 55, Dover publications New York.

Bauder, D., Bodnar, T., Mazur, S. & Okhrin, Y. (2017), 'Bayesian inference for the tangent portfolio'.

Bodnar, T., Mazur, S. & Okhrin, Y. (2017), 'Bayesian estimation of the global minimum variance portfolio', *European Journal of Operational Research* **256**(1), 292–307.

Frahm, G. & Memmel, C. (2010), 'Dominating estimators for minimum-variance portfolios', *Journal of Econometrics* **159**(2), 289–302.

Kempf, A. & Memmel, C. (2006*a*), 'Estimating the global minimum variance portfolio', *Schmalenbach Business Review* **58**(4), 332–348.

Kempf, A. & Memmel, C. (2006*b*), 'Estimating the global minimum variance portfolio', *Schmalenbach Business Review* **58**(4), 332–348.

Maillet, B., Tokpavi, S. & Vaucher, B. (2015), 'Global minimum variance portfolio optimisation under some model risk: A robust regression-based approach', *European Journal of Operational Research* **244**(1), 289–299.

Markowitz, H. (1952), 'Portfolio selection', *The journal of finance* **7**(1), 77–91.

R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *http://www.R-project.org/*

Tunaru, R. (2002), 'Hierarchical bayesian models for multiple count data', *Austrian Journal of statistics* **31**(3), 221–229.

Appendix

Following is the R code which is used in this thesis

```
#Load Data Set
full_data=Data
#Divide data set into prior and working data
prior = full_data[c(1:293),c(1:5)];
working_data = full_data[c(294:586),c(1:5)];
#################### For Prior Data ##########################
#Mean vector and covariance matrix of prior data
mu_prior <- colMeans(x=prior, na.rm = TRUE)
s_prior <- var(x=prior, y = NULL, na.rm = FALSE, use="all.obs")
#Calculation of Wprior
inv_s_prior <- solve(s_prior)
vector_1 <- rep(1, 5)
trnsp_vector_1 <- t(vector_1)
numerator_wprior <- inv_s_prior%*%vector_1
denomi1 <- trnsp_vector_1%*%inv_s_prior
denominator_wprior <- as.vector(denomi1%*%vector_1)
wprior <- numerator_wprior/denominator_wprior
wprior
sum (wprior[,1])
#################### For Working Data #########################
#Mean vector and covariance matrix of working data
Xbar_working <- colMeans(x=working_data, na.rm = TRUE)
s_working <- var(x=working_data, y = NULL, na.rm = FALSE, use="all.obs")
w_working <- ginv(s_working) %*% transpose(vector_1)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
w_working
sum (w_working[,1])
#################### Diffuse Prior #########################
k <- 5
p <- 1
n <- 293
## For index1
e1 <- c(1,0,0,0,0)
e1
theta1_hat <- e1 %*% w_working
R_d <- ginv(s_working)-ginv(s_working) %*% transpose(vector_1) %*% vector_1 %*%
ginv(s_working)/sum(diag(vector_1 %*% ginv(s_working) %*% transpose(vector_1)))
R1_d <- 1/(n-1) * e1%*%R_d%*%transpose(e1)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
## For index2
e2 <- c(0,1,0,0,0)
e2
theta2_hat <- e2 %*% w_working
R2_d <- 1/(n-1) * e2%*%R_d%*%transpose(e2)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
## For index3
e3 <- c(0,0,1,0,0)
theta3_hat <- e3 %*% w_working
R3_d <- 1/(n-1) * e3%*%R_d%*%transpose(e3)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
## For index4
```

```
e4 <- c(0,0,0,1,0)
theta4_hat <- e4 %*% w_working
R4_d <- 1/(n-1) * e4%*%R_d%*%transpose(e4)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
## For index5
e5 <- c(0,0,0,0,1)
theta5_hat <- e5 %*% w_working
R5_d <- 1/(n-1) * e5%*%R_d%*%transpose(e5)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
## PLOTS
x<- seq(-2,2,0.01)
plot(x,dt.scaled(x,n-1,theta1_hat,sqrt(R1_d)),        xlim=c(-0.6,1.3),       ylim=c(0,10),       col="black",
type="l",lwd=2, lty=1, xlab="", ylab="")
lines(x,dt.scaled(x,n-1,theta2_hat,sqrt(R2_d)) ,type="l",col="blue",lwd=2, lty=1 )
lines(x,dt.scaled(x,n-1,theta3_hat,sqrt(R3_d)) ,type="l",col="green",lwd=2, lty=1)
lines(x,dt.scaled(x,n-1,theta4_hat,sqrt(R4_d)) ,type="l",col="red",lwd=2, lty=1)
lines(x,dt.scaled(x,n-1,theta5_hat,sqrt(R5_d)) ,type="l",col="sienna2",lwd=2, lty=1)
legend("topleft", c("DAX ","S&P500","OMX ","CAC","Nikkei"), col=c("black", "blue", "green", "red",
"sienna2"), lwd=c(2,2,2,2,2), lty=c(1,1,1,1,1), bty = "n")


################### Conjugate Prior #########################
v_c <- 293
k_c <- v_c
r_c <- (n*Xbar_working+k_c*mu_prior)/(n+k_c)
V_c       <-      (n-1)*s_working      +      s_prior      +      (n+k_c)*transpose(r_c)%*%r_c      +
n*transpose(Xbar_working)%*%Xbar_working + k_c*transpose(mu_prior)%*%mu_prior
R_c <- ginv(V_c)-ginv(V_c) %*% transpose(vector_1) %*% vector_1%*%ginv(V_c)/sum(diag(vector_1
%*% ginv(V_c) %*% transpose(vector_1)))
V1_c   <-   e1%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1   %*%   ginv(V_c)   %*%
transpose(vector_1)))
V2_c   <-   e2%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1   %*%   ginv(V_c)   %*%
transpose(vector_1)))
V3_c   <-   e3%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1   %*%   ginv(V_c)   %*%
transpose(vector_1)))
V4_c   <-   e4%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1   %*%   ginv(V_c)   %*%
transpose(vector_1)))
V5_c   <-   e5%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1   %*%   ginv(V_c)   %*%
transpose(vector_1)))
R1_c   <-   1/(v_c+n-k-1)  *  e1%*%R_c%*%transpose(e1)/sum(diag(vector_1  %*%  ginv(V_c)  %*%
transpose(vector_1)))
R2_c   <-   1/(v_c+n-k-1)  *  e2%*%R_c%*%transpose(e2)/sum(diag(vector_1  %*%  ginv(V_c)  %*%
transpose(vector_1)))
R3_c   <-   1/(v_c+n-k-1)  *  e3%*%R_c%*%transpose(e3)/sum(diag(vector_1  %*%  ginv(V_c)  %*%
transpose(vector_1)))
R4_c   <-   1/(v_c+n-k-1)  *  e4%*%R_c%*%transpose(e4)/sum(diag(vector_1  %*%  ginv(V_c)  %*%
transpose(vector_1)))
R5_c   <-   1/(v_c+n-k-1)  *  e5%*%R_c%*%transpose(e5)/sum(diag(vector_1  %*%  ginv(V_c)  %*%
transpose(vector_1)))
# PLOTS
x<- seq(-1,1,0.01)
plot(x,dt.scaled(x,v_c+n-k-1,V1_c,sqrt(R1_c))    ,    xlim=c(-0.5,1.2),    ylim=c(0,15),    col="black",
type="l",lwd=2, lty=1, xlab="", ylab="")
lines(x,dt.scaled(x,v_c+n-k-1,V2_c,sqrt(R2_c)) ,type="l",col="blue",lwd=2, lty=1 )
lines(x,dt.scaled(x,v_c+n-k-1,V3_c,sqrt(R3_c)) ,type="l",col="green",lwd=2, lty=1 )
lines(x,dt.scaled(x,v_c+n-k-1,V4_c,sqrt(R4_c)) ,type="l",col="red",lwd=2, lty=1)
lines(x,dt.scaled(x,v_c+n-k-1,V5_c,sqrt(R5_c)) ,type="l",col="sienna2",lwd=2, lty=1)
```

```
legend("topleft", c("DAX ","S&P500","OMX ","CAC","Nikkei"), col=c("black", "blue", "green", "red",
"sienna2"), lwd=c(2,2,2,2,2), lty=c(1,1,1,1,1), bty = "n")
################### Jeffreys non-informative Prior ########################
R1_d_j <- 1/(n-k+p) * e1%*%R_d%*%transpose(e1)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
R2_d_j <- 1/(n-k+p) * e2%*%R_d%*%transpose(e2)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
R3_d_j <- 1/(n-k+p) * e3%*%R_d%*%transpose(e3)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
R4_d_j <- 1/(n-k+p) * e4%*%R_d%*%transpose(e4)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
R5_d_j <- 1/(n-k+p) * e5%*%R_d%*%transpose(e5)/sum(diag(vector_1 %*% ginv(s_working) %*%
transpose(vector_1)))
## PLOTS
x<- seq(-2,2,0.01)
plot(x,dt.scaled(x,n-k+p,theta1_hat,sqrt(R1_d_j)),    xlim=c(-0.5,1.3),    ylim=c(0,10),    col="black",
type="l",lwd=2, lty=1, xlab="", ylab="")
lines(x,dt.scaled(x,n-k+p,theta2_hat,sqrt(R2_d_j)) ,type="l",col="blue",lwd=2, lty=1 )
lines(x,dt.scaled(x,n-k+p,theta3_hat,sqrt(R3_d_j)) ,type="l",col="green",lwd=2, lty=1 )
lines(x,dt.scaled(x,n-k+p,theta4_hat,sqrt(R4_d_j)) ,type="l",col="red",lwd=2, lty=1 )
lines(x,dt.scaled(x,n-k+p,theta5_hat,sqrt(R5_d_j)) ,type="l",col="sienna2",lwd=2, lty=1 )
legend("topleft", c("DAX ","S&P500","OMX ","CAC","Nikkei"), col=c("black", "blue", "green", "red",
"sienna2"), lwd=c(2,2,2,2,2), lty=c(1,1,1,1,1), bty = "n")
################### Informative Prior ########################
################# For index 1 ####################
S_I <-1
P1_1 <- 1/S_I+(n-1)*(ginv(e1%*%R_d%*%transpose(e1)))
P2_1 <- (n-1)*ginv(e1%*%R_d%*%transpose(e1))
P1_1 <- apply(P1_1,1,as.numeric)
P2_1 <- apply(P2_1,1,as.numeric)
delta1 <- 1
delta2 <- 0.5
r <- ginv(delta2) +(n-1)*ginv(sum(diag(vector_1 %*% ginv(s_working) %*% transpose(vector_1))))
vI <- n
r <- apply(r,1,as.numeric)
wI <- 1/k
NR <- 10^4
theta1 <- rep(1,NR)
for (i in 1:NR){
z0 <- rnorm(1,0,1)
tau <- rgamma(1,shape=(n-k+p+vI-1)/2,scale=2)
v1_tau <- ginv((tau*P1_1)+P2_1)
v1_tau <- apply(v1_tau,1,as.numeric)
r1_tau <- ginv((tau*P1_1)+P2_1)*((tau*P1_1*wI)+(P2_1*theta1_hat))
r1_tau <- apply(r1_tau,1,as.numeric)
h1_tau  <-  r+  (tau*transpose(wI)*P1_1*wI)  +  (transpose(theta1_hat)*P2_1*theta1_hat)  -
(transpose(r1_tau))*(ginv(v1_tau))*(r1_tau)
h1_tau <- apply(h1_tau,1,as.numeric)
sai1_t <- rgamma(1,shape=(n-k+(2*p)+(2*delta1))/2,scale=2/(h1_tau))
################Stochastic Representation ################
theta1[i] <- r1_tau+(sai1_t)^(-1/2)*((v1_tau))^(1/2)*z0
}
mean(theta1)
quantile(theta1, 0.05)
var(theta1)
plot(density(theta1))
```

```
################ For index 2 ###################
P1_2 <- 1/S_I+(n-1)*(ginv(e2%*%R_d%*%transpose(e2)))
P2_2 <- (n-1)*ginv(e2%*%R_d%*%transpose(e2))
P1_2 <- apply(P1_2,1,as.numeric)
P2_2 <- apply(P2_2,1,as.numeric)
NR <- 10^4
theta2 <- rep(1,NR)
for (i in 1:NR){
 z0 <- rnorm(1,0,1)
 tau <- rgamma(1,shape=(n-k+p+vI-1)/2,scale=2)
 v2_tau <- ginv((tau*P1_2)+P2_2)
 v2_tau <- apply(v2_tau,1,as.numeric)
 r2_tau <- ginv((tau*P1_2)+P2_2)*((tau*P1_2*wI)+(P2_2*theta2_hat))
 r2_tau <- apply(r2_tau,1,as.numeric)
 h2_tau  <-  r+  (tau*transpose(wI)*P1_2*wI)  +  (transpose(theta2_hat)*P2_2*theta2_hat)  -
(transpose(r2_tau))*(ginv(v2_tau))*(r2_tau)
 h2_tau <- apply(h2_tau,1,as.numeric)
 sai2_t <- rgamma(1,shape=(n-k+(2*p)+(2*delta1))/2,scale=2/(h2_tau))
 #################Stochastic Representation ##################
 theta2[i] <- r2_tau+(sai2_t)^(-1/2)*((v2_tau))^(1/2)*z0
}
mean(theta2)
quantile(theta2, 0.05)
var(theta2)
plot(density(theta2))
################ For index 3 ###################
P1_3 <- 1/S_I+(n-1)*(ginv(e3%*%R_d%*%transpose(e3)))
P2_3 <- (n-1)*ginv(e3%*%R_d%*%transpose(e3))
P1_3 <- apply(P1_3,1,as.numeric)
P2_3 <- apply(P2_3,1,as.numeric)

NR <- 10^4
theta3 <- rep(1,NR)
for (i in 1:NR){
 z0 <- rnorm(1,0,1)
 tau <- rgamma(1,shape=(n-k+p+vI-1)/2,scale=2)
 v3_tau <- ginv((tau*P1_3)+P2_3)
 v3_tau <- apply(v3_tau,1,as.numeric)
 r3_tau <- ginv((tau*P1_3)+P2_3)*((tau*P1_3*wI)+(P2_3*theta3_hat))
 r3_tau <- apply(r3_tau,1,as.numeric)
 h3_tau  <-  r+  (tau*transpose(wI)*P1_3*wI)  +  (transpose(theta3_hat)*P2_3*theta3_hat)  -
(transpose(r3_tau))*(ginv(v3_tau))*(r3_tau)
 h3_tau <- apply(h3_tau,1,as.numeric)
 sai3_t <- rgamma(1,shape=(n-k+(2*p)+(2*delta1))/2,scale=2/(h3_tau))
 #################Stochastic Representation ##################
 theta3[i] <- r3_tau+(sai3_t)^(-1/2)*((v3_tau))^(1/2)*z0
}
mean(theta3)
quantile(theta3, 0.05)
var(theta3)
plot(density(theta3))

################ For index 4 ###################

P1_4 <- 1/S_I+(n-1)*(ginv(e4%*%R_d%*%transpose(e4)))
P2_4 <- (n-1)*ginv(e4%*%R_d%*%transpose(e4))
```

```
P1_4 <- apply(P1_4,1,as.numeric)
P2_4 <- apply(P2_4,1,as.numeric)
NR <- 10^4
theta4 <- rep(1,NR)
for (i in 1:NR){
 z0 <- rnorm(1,0,1)
 tau <- rgamma(1,shape=(n-k+p+vI-1)/2,scale=2)
 v4_tau <- ginv((tau*P1_4)+P2_4)
 v4_tau <- apply(v4_tau,1,as.numeric)
 r4_tau <- ginv((tau*P1_4)+P2_4)*((tau*P1_4*wI)+(P2_4*theta4_hat))
 r4_tau <- apply(r4_tau,1,as.numeric)
 h4_tau <- r+ (tau*transpose(wI)*P1_4*wI) + (transpose(theta4_hat)*P2_4*theta4_hat) -
(transpose(r4_tau))*(ginv(v4_tau))*(r4_tau)
 h4_tau <- apply(h4_tau,1,as.numeric)
 sai4_t <- rgamma(1,shape=(n-k+(2*p)+(2*delta1))/2,scale=2/(h4_tau))
 #################Stochastic Representation #################
 theta4[i] <- r2_tau+(sai2_t)^(-1/2)*((v4_tau))^(1/2)*z0
}
mean(theta4)
quantile(theta4, 0.05)
var(theta4)
plot(density(theta4))
################# For index 5 ####################
P1_5 <- 1/S_I+(n-1)*(ginv(e5%*%R_d%*%transpose(e5)))
P2_5 <- (n-1)*ginv(e5%*%R_d%*%transpose(e5))
P1_5 <- apply(P1_5,1,as.numeric)
P2_5 <- apply(P2_5,1,as.numeric)
NR <- 10^4
theta5 <- rep(1,NR)
for (i in 1:NR){
 z0 <- rnorm(1,0,1)
 tau <- rgamma(1,shape=(n-k+p+vI-1)/2,scale=2)
 v5_tau <- ginv((tau*P1_5)+P2_5)
 v5_tau <- apply(v5_tau,1,as.numeric)
 r5_tau <- ginv((tau*P1_5)+P2_5)*((tau*P1_5*wI)+(P2_5*theta5_hat))
 r5_tau <- apply(r5_tau,1,as.numeric)
 h5_tau <- r+ (tau*transpose(wI)*P1_5*wI) + (transpose(theta5_hat)*P2_5*theta5_hat) -
(transpose(r5_tau))*(ginv(v5_tau))*(r5_tau)
 h5_tau <- apply(h5_tau,1,as.numeric)
 sai5_t <- rgamma(1,shape=(n-k+(2*p)+(2*delta1))/2,scale=2/(h5_tau))

 #################Stochastic Representation #################
 theta5[i] <- r2_tau+(sai2_t)^(-1/2)*((v2_tau))^(1/2)*z0
}
mean(theta5)
quantile(theta5, 0.05)
var(theta5)
plot(density(theta5))
## PLOTS
x<- seq(-2,2,0.01)
plot(density(theta1), xlim=c(0.18,0.22), ylim=c(0,130), col="black", type="l",lwd=2, lty=1, xlab="",
ylab="")
lines(density(theta2,kernel="epanechnikov"), type="l",col="blue",lwd=2, lty=1 )
lines(density(theta3,kernel="epanechnikov"), type="l",col="green",lwd=2, lty=1 )
lines(density(theta4,kernel="epanechnikov") ,type="l",col="red",lwd=2, lty=1 )
lines(density(theta5,kernel="epanechnikov"), type="l",col="sienna2",lwd=2, lty=1 )
```

```
legend("topleft", c("DAX ","S&P500","OMX ","CAC","Nikkei"), col=c("black", "blue", "green", "red",
"sienna2"), lwd=c(2,2,2,2,2), lty=c(1,1,1,1,1), bty = "n")


######################## Performance Measure Diffuse Prior
############################
#Mean
mean_df_1 <- theta1_hat
mean_df_2 <- theta2_hat
mean_df_3 <- theta3_hat
mean_df_4 <- theta4_hat
mean_df_5 <- theta5_hat


#Variance
var_df_1 <- 1/((n-1)-2) * e1%*%R_d%*%transpose(e1)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_df_2 <- 1/((n-1)-2) * e2%*%R_d%*%transpose(e2)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_df_3 <- 1/((n-1)-2) * e3%*%R_d%*%transpose(e3)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_df_4 <- 1/((n-1)-2) * e4%*%R_d%*%transpose(e4)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_df_5 <- 1/((n-1)-2) * e5%*%R_d%*%transpose(e5)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
# 95% Confidence Interval for diffuse prior
SE_df_1 <- sqrt(var_df_1)
CI_95_df_1_lower <- mean_df_1 - SE_df_1* qt(0.975,df=n-1)
CI_95_df_1_upper <- mean_df_1 + SE_df_1* qt(0.975,df=n-1)
SE_df_2 <- sqrt(var_df_2)
CI_95_df_2_lower <- mean_df_2-SE_df_2* qt(0.975,df=n-1)
CI_95_df_2_upper <- mean_df_2+SE_df_2* qt(0.975,df=n-1)
SE_df_3 <- sqrt(var_df_2)
CI_95_df_3_lower <- mean_df_3-SE_df_3* qt(0.975,df=n-1)
CI_95_df_3_upper <- mean_df_3+SE_df_3* qt(0.975,df=n-1)
SE_df_4 <- sqrt(var_df_2)
CI_95_df_4_lower <- mean_df_4-SE_df_4* qt(0.975,df=n-1)
CI_95_df_4_upper <- mean_df_4+SE_df_4* qt(0.975,df=n-1)
SE_df_5 <- sqrt(var_df_2)
CI_95_df_5_lower <- mean_df_5-SE_df_5* qt(0.975,df=n-1)
CI_95_df_5_upper <- mean_df_5+SE_df_5* qt(0.975,df=n-1)


# 99% Confidence Interval for diffuse prior

CI_99_df_1_lower <- mean_df_1-SE_df_1* qt(0.995,df=n-1)
CI_99_df_1_upper <- mean_df_1+SE_df_1* qt(0.995,df=n-1)
CI_99_df_2_lower <- mean_df_2-SE_df_2* qt(0.995,df=n-1)
CI_99_df_2_upper <- mean_df_2+SE_df_2* qt(0.995,df=n-1)
CI_99_df_3_lower <- mean_df_3-SE_df_3* qt(0.995,df=n-1)
CI_99_df_3_upper <- mean_df_3+SE_df_3* qt(0.995,df=n-1)
CI_99_df_4_lower <- mean_df_4-SE_df_4* qt(0.995,df=n-1)
CI_99_df_4_upper <- mean_df_4+SE_df_4* qt(0.995,df=n-1)
CI_99_df_5_lower <- mean_df_5-SE_df_5* qt(0.995,df=n-1)
CI_99_df_5_upper <- mean_df_5+SE_df_5* qt(0.995,df=n-1)


#################### Performance Measure Conjugate Prior
########################
#Mean
```

```r
mean_cn_1 <- e1%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1 %*% ginv(V_c) %*%
transpose(vector_1)))
mean_cn_2 <- e2%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1 %*% ginv(V_c) %*%
transpose(vector_1)))
mean_cn_3 <- e3%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1 %*% ginv(V_c) %*%
transpose(vector_1)))
mean_cn_4 <- e4%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1 %*% ginv(V_c) %*%
transpose(vector_1)))
mean_cn_5 <- e5%*%ginv(V_c)%*%transpose(vector_1)/sum(diag(vector_1 %*% ginv(V_c) %*%
transpose(vector_1)))
#Variance
var_cn_1 <- 1/((v_c+n-k-1)-2) * e1%*%R_c%*%transpose(e1)/sum(diag(vector_1 %*% ginv(V_c)
%*% transpose(vector_1)))
var_cn_2 <- 1/((v_c+n-k-1)-2) * e2%*%R_c%*%transpose(e2)/sum(diag(vector_1 %*% ginv(V_c)
%*% transpose(vector_1)))
var_cn_3 <- 1/((v_c+n-k-1)-2) * e3%*%R_c%*%transpose(e3)/sum(diag(vector_1 %*% ginv(V_c)
%*% transpose(vector_1)))
var_cn_4 <- 1/((v_c+n-k-1)-2) * e4%*%R_c%*%transpose(e4)/sum(diag(vector_1 %*% ginv(V_c)
%*% transpose(vector_1)))
var_cn_5 <- 1/((v_c+n-k-1)-2) * e5%*%R_c%*%transpose(e5)/sum(diag(vector_1 %*% ginv(V_c)
%*% transpose(vector_1)))
# 95% Confidence Interval for conjugate prior
SE_cn_1 <- sqrt(var_cn_1)
CI_95_cn_1_lower <- mean_cn_1-SE_cn_1* qt(0.975,df=n-1)
CI_95_cn_1_upper <- mean_cn_1+SE_cn_1* qt(0.975,df=n-1)
SE_cn_2 <- sqrt(var_cn_2)
CI_95_cn_2_lower <- mean_cn_2-SE_cn_2* qt(0.975,df=n-1)
CI_95_cn_2_upper <- mean_cn_2+SE_cn_2* qt(0.975,df=n-1)
SE_cn_3 <- sqrt(var_cn_3)
CI_95_cn_3_lower <- mean_cn_3-SE_cn_3* qt(0.975,df=n-1)
CI_95_cn_3_upper <- mean_cn_3+SE_cn_3* qt(0.975,df=n-1)
SE_cn_4 <- sqrt(var_cn_4)
CI_95_cn_4_lower <- mean_cn_4-SE_cn_4* qt(0.975,df=n-1)
CI_95_cn_4_upper <- mean_cn_4+SE_cn_4* qt(0.975,df=n-1)
SE_cn_5 <- sqrt(var_cn_5)
CI_95_cn_5_lower <- mean_cn_5-SE_cn_5* qt(0.975,df=n-1)
CI_95_cn_5_upper <- mean_cn_5+SE_cn_5* qt(0.975,df=n-1)
# 99% Confidence Interval for conjugate prior
CI_99_cn_1_lower <- mean_cn_1-SE_cn_1* qt(0.995,df=n-1)
CI_99_cn_1_upper <- mean_cn_1+SE_cn_1* qt(0.995,df=n-1)
CI_99_cn_2_lower <- mean_cn_2-SE_cn_2* qt(0.995,df=n-1)
CI_99_cn_2_upper <- mean_cn_2+SE_cn_2* qt(0.995,df=n-1)
CI_99_cn_3_lower <- mean_cn_3-SE_cn_3* qt(0.995,df=n-1)
CI_99_cn_3_upper <- mean_cn_3+SE_cn_3* qt(0.995,df=n-1)
CI_99_cn_4_lower <- mean_cn_4-SE_cn_4* qt(0.995,df=n-1)
CI_99_cn_4_upper <- mean_cn_4+SE_cn_4* qt(0.995,df=n-1)
CI_99_cn_5_lower <- mean_cn_5-SE_cn_5* qt(0.995,df=n-1)
CI_99_cn_5_upper <- mean_cn_5+SE_cn_5* qt(0.995,df=n-1)
####################### Performance Measure Jeffrey Prior ###########
#Mean
mean_jf_1 <- theta1_hat
mean_jf_2 <- theta2_hat
mean_jf_3 <- theta3_hat
mean_jf_4 <- theta4_hat
mean_jf_5 <- theta5_hat
#Variance
```

```
var_jf_1 <- 1/((n-k+p)-2) * e1%*%R_d%*%transpose(e1)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_jf_2 <- 1/((n-k+p)-2) * e2%*%R_d%*%transpose(e2)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_jf_3 <- 1/((n-k+p)-2) * e3%*%R_d%*%transpose(e3)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_jf_4 <- 1/((n-k+p)-2) * e4%*%R_d%*%transpose(e4)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
var_jf_5 <- 1/((n-k+p)-2) * e5%*%R_d%*%transpose(e5)/sum(diag(vector_1 %*% ginv(s_working)
%*% transpose(vector_1)))
# 95% Confidence Interval for Jeffreys non-informative
SE_jf_1 <- sqrt(var_jf_1)
CI_95_jf_1_lower <- mean_jf_1-SE_jf_1* qt(0.975,df=n-1)
CI_95_jf_1_upper <- mean_jf_1+SE_jf_1* qt(0.975,df=n-1)
SE_jf_2 <- sqrt(var_jf_2)
CI_95_jf_2_lower <- mean_jf_2-SE_jf_2* qt(0.975,df=n-1)
CI_95_jf_2_upper <- mean_jf_2+SE_jf_2* qt(0.975,df=n-1)
SE_jf_3 <- sqrt(var_jf_3)
CI_95_jf_3_lower <- mean_jf_3-SE_jf_3* qt(0.975,df=n-1)
CI_95_jf_3_upper <- mean_jf_3+SE_jf_3* qt(0.975,df=n-1)
SE_jf_4 <- sqrt(var_jf_4)
CI_95_jf_4_lower <- mean_jf_4-SE_jf_4* qt(0.975,df=n-1)
CI_95_jf_4_upper <- mean_jf_4+SE_jf_4* qt(0.975,df=n-1)
SE_jf_5 <- sqrt(var_jf_5)
CI_95_jf_5_lower <- mean_jf_5-SE_jf_5* qt(0.975,df=n-1)
CI_95_jf_5_upper <- mean_jf_5+SE_jf_5* qt(0.975,df=n-1)
# 99% Confidence Interval for Jeffreys non-informative
CI_99_jf_1_lower <- mean_jf_1-SE_jf_1* qt(0.995,df=n-1)
CI_99_jf_1_upper <- mean_jf_1+SE_jf_1* qt(0.995,df=n-1)
CI_99_jf_2_lower <- mean_jf_2-SE_jf_2* qt(0.995,df=n-1)
CI_99_jf_2_upper <- mean_jf_2+SE_jf_2* qt(0.995,df=n-1)
CI_99_jf_3_lower <- mean_jf_3-SE_jf_3* qt(0.995,df=n-1)
CI_99_jf_3_upper <- mean_jf_3+SE_jf_3* qt(0.995,df=n-1)
CI_99_jf_4_lower <- mean_jf_4-SE_jf_4* qt(0.995,df=n-1)
CI_99_jf_4_upper <- mean_jf_4+SE_jf_4* qt(0.995,df=n-1)
CI_99_jf_5_lower <- mean_jf_5-SE_jf_5* qt(0.995,df=n-1)
CI_99_jf_5_upper <- mean_jf_5+SE_jf_5* qt(0.995,df=n-1)


###################### Performance Measure Informative Prior  ################
#Mean
mean_in_1 <- mean(theta1)
mean_in_2 <- mean(theta2)
mean_in_3 <- mean(theta3)
mean_in_4 <- mean(theta4)
mean_in_5 <- mean(theta5)
#Variance
var_in_1 <- var(theta1)
var_in_2 <- var(theta2)
var_in_3 <- var(theta3)
var_in_4 <- var(theta4)
var_in_5 <- var(theta5)
# 95% Confidence Interval for informative prior
SE_in_1 <- sqrt(var_in_1)
CI_95_in_1_lower <- mean_in_1-SE_in_1* qt(0.975,df=n-1)
CI_95_in_1_upper <- mean_in_1+SE_in_1* qt(0.975,df=n-1)
SE_in_2 <- sqrt(var_in_2)
```

```
CI_95_in_2_lower <- mean_in_2-SE_in_2* qt(0.975,df=n-1)
CI_95_in_2_upper <- mean_in_2+SE_in_2* qt(0.975,df=n-1)
SE_in_3 <- sqrt(var_in_3)
CI_95_in_3_lower <- mean_in_3-SE_in_3* qt(0.975,df=n-1)
CI_95_in_3_upper <- mean_in_3+SE_in_3* qt(0.975,df=n-1)
SE_in_4 <- sqrt(var_in_4)
CI_95_in_4_lower <- mean_in_4-SE_in_4* qt(0.975,df=n-1)
CI_95_in_4_upper <- mean_in_4+SE_in_4* qt(0.975,df=n-1)
SE_in_5 <- sqrt(var_in_5)
CI_95_in_5_lower <- mean_in_5-SE_in_5* qt(0.975,df=n-1)
CI_95_in_5_upper <- mean_in_5+SE_in_5* qt(0.975,df=n-1)
# 99% Confidence Interval for informative prior
CI_99_in_1_lower <- mean_in_1-SE_in_1* qt(0.995,df=n-1)
CI_99_in_1_upper <- mean_in_1+SE_in_1* qt(0.995,df=n-1)
CI_99_in_2_lower <- mean_in_2-SE_in_2* qt(0.995,df=n-1)
CI_99_in_2_upper <- mean_in_2+SE_in_2* qt(0.995,df=n-1)
CI_99_in_3_lower <- mean_in_3-SE_in_3* qt(0.995,df=n-1)
CI_99_in_3_upper <- mean_in_3+SE_in_3* qt(0.995,df=n-1)
CI_99_in_4_lower <- mean_in_4-SE_in_4* qt(0.995,df=n-1)
CI_99_in_4_upper <- mean_in_4+SE_in_4* qt(0.995,df=n-1)
CI_99_in_5_lower <- mean_in_5-SE_in_5* qt(0.995,df=n-1)
CI_99_in_5_upper <- mean_in_5+SE_in_5* qt(0.995,df=n-1).
```