ISSN 1684-8403
Journal of Statistics
Volume 22, 2015. pp. 92-102

A Bayesian Approach for Estimating Mean-Standard Deviation Ratios of Financial Data

Rashid Mansoor[1]

Abstract

The relation between excess return and risk of financial assets is frequently determined by the mean-standard deviation ratio. Previous research in this matter only derived Point Estimators of this parameter while Inference procedures are currently void. This paper derives a Bayesian procedure for making Inference of this ratio which is easy to apply. Specifically, a method for testing if two ratio coefficients are equal for two independent population segments is derived. This, hence, provides the analyst with a tool for assessing if, e.g. Technique Stocks and Forestry Stocks, have equal risk/return ratio. This paper demonstrates the procedure by an empirical application using data from the Stockholm Stock Exchange.

Keywords

Bayesian regression, Hypotheses testing, Credibility interval, Sharpe ratio

1. Introduction

It is often of interest to compare the standard deviation with the mean value of a Distribution. For example, in financial economic theories it is stated that the higher the risk an investor is willing to take, the higher is the potential return from the investment. The standard deviation can be used to compare risks among investments that have the same expected return. Another way to assess risk/return ratios is to use the reciprocal of the coefficient of variation, which corresponds to the mean value divided by the standard deviation.

[1]Jonkoping International Business School, Box 1026, 55 111, Sweden.

This has been used in many fields such as finance, Osteryoung et al. (1977) and imaging, Lopes et al. (1993),etc. This measure indicates the expected return per unit of risk. Hence, a larger value of this quotient signifies a greater return relative to the risk.

In financial theory, this measure is known as the Sharpe Ratio, (Sharpe (1966)). While this ratio is traditionally calculated for a single Stock, Holgersson et al. (2012) derived a related ratio to describe the common risk/return ratio, say β, for a whole segment of Stocks, e.g. for Large Cap Stocks. Three different Estimators of β were proposed using the Frequentist approach. However, that paper did not provide any theories for Inference of this parameter and it is questionable if such procedures can be derived at all within a Frequentist framework. This paper, therefore, derives procedures for Statistical Inference, in particular Interval Estimation and hence Hypothesis Testing, for assessing if two population segments, e.g. Technique Stocks and Forestry Stocks or Large Cap and Small Cap Stocks, have equal risk/return ratio. This in turn should be a useful tool for investors since they would typically like to have the highest possible financial award for taking a given risk, while Point Estimates alone will be insufficient for important decisions. Specifically, we propose the use of a vague Prior, which is a Non-informative Prior popular in the Bayesian literature (see for example; Box and Tiao (1973), Tiao and Zellner (1964) and Zellner (1971)). By using this Prior, one can connect the Bayesian modeling result with Frequentist results proposed by Holgersson et al. (2012). In order to demonstrate the use and application of the proposed method it is applied on a data set from the Stockholm Stock Exchange where the possible equality between β_{Small} (Small Cap parameter) and β_{Large} (Large Cap parameter) is tested.

The paper is organized as; in Section 2 the underlying model is specified along with the proposed method. Then in Section 3, the empirical application is provided and, finally, in Section 4 a brief summary and a conclusion are given.

2. Model specification

To model the link between the standard deviation and mean of a variable X_{it}, e.g. the return of Stock X_i at time *t*, we consider the following functional form between the mean and standard deviation,

$$\mu_i = \beta\sigma_i \tag{2.1}$$

where

X_{it} is a normally distributed variable such that $\mu_i = E\left[X_{it}\right]$, $\sigma_i = \sqrt{E\left[X_{it} - \mu_i\right]^2}$, $i = 1,...,n,\ t = 1,...,T$ and $-\infty < \mu_i < \infty, 0 < \sigma < \infty$ (Holgersson et al. (2012)).

Hence, the functional form of the relationship between the mean and standard deviation corresponds to a linear one. Such a setting may be justified by looking at Figure 1 (further investigations are supplied in the Appendix). In that figure, the relationship between the sample mean and standard deviation of the S&P 100 Stocks is applied as an illustration. The data used are retrieved from the database Data-Stream involving monthly returns covering the time period 1995- 2010.

The problem of making Inference of the Beta parameter of equation (2.1) should not be confused with traditional Financial Risk Analysis which is typically concerned only with Risk Estimation of a single Stock or of a portfolio of Stocks. In order to make Inference of the β parameter we note that $\bar{X}_i = T^{-1}\sum_{t=1}^{T} X_{it}$ and $S_i = \sqrt{(T-1)^{-1}\sum_{t=1}^{T}\left(X_{it} - \bar{X}_i\right)^2}$ are Sufficient Statistics for μ_i and σ_i respectively. Hence, the model specified in equation (2.1) can be written as:

$$\bar{X}_i = \beta\sigma_i + \delta_{it} \tag{2.2}$$

where

δ_{it} allows the relationship/points to fall away from the line with slope β and $\delta_{it} \sim N\left(0, \frac{\sigma_\delta^2}{n}\right)$, $E(\bar{X}_i \mid \sigma_i, \beta, \sigma_\delta) = \beta\sigma_i$ and $Var(\bar{X}_i \mid \sigma_i, \beta, \sigma_\delta) = \frac{\sigma_\delta^2}{n}$

Hence, it follows that $\bar{X}_i \mid \sigma_i \in N\left(\beta\sigma_i, \frac{\sigma_\delta^2}{n}\right)$ (2.3)

The Likelihood Function as a function of the parameters β and σ_δ can then be written as:

$$p(\bar{X}_i \mid \sigma_i, \beta, \sigma_\delta) \propto \prod_{i=1}^{n}\sqrt{\frac{n}{\sigma_\delta^2}}\exp\left[-\frac{\left(\bar{X}_i - \beta\sigma_i\right)}{\frac{2\sigma_\delta^2}{n}}\right] = \frac{n^{n/2}}{\sigma_\delta^n}\exp\left(-\frac{\sum_{i=1}^{n}\left(\bar{X}_i - \beta\sigma_i\right)^2}{2\sigma_\delta^2/n}\right) \tag{2.4}$$

The expression in equation (2.4) is the Likelihood Function to be combined with a suitable Prior probability density function for the parameters of interest. In recent Bayesian literature, Non-informative Priors are used for fidelity reason and the reason behind the use of a vague Prior is that one can conveniently link the Bayesian modeling result with Frequentist results. Numerically, the Frequentist and the Bayesian approach usually yield similar results in case the Bayesian procedure uses a vague Prior. In our context, however, no Frequentist Inference procedures are currently available due to the complexity of the parameter Point estimate (Holgersson et al. (2012)). For the Prior Distribution of β and σ_δ we here assume that only little Prior information is available about these parameters. A specific Non-informative Prior proposed by Zellner (1971) is defined as:

$$p(\beta,\sigma_\delta) = p(\beta)\,p(\sigma_\delta)\ . \tag{2.5}$$

Using equation (2.5), the Prior probability density function (p.d.f) for β and σ_δ can be written as:

$$p(\beta,\sigma_\delta) \propto \frac{1}{\sigma_\delta} \tag{2.6}$$

where

$-\infty < \beta < \infty$ and $0 < \sigma_\delta < \infty$.

By combining equation (2.4) and (2.6) the Joint Posterior p.d.f for β and σ_δ becomes as follows:

$$p(\beta,\sigma_\delta \mid \bar{X}_i,\sigma_i) \propto p\left(\beta,\sigma_\delta\right) p(\bar{X}_i \mid \sigma_i,\beta,\sigma_\delta) = \frac{n^{\frac{n}{2}}}{\sigma_\delta{}^{n+1}} \exp\left(-\frac{\sum_{i=1}^{n}\left(\bar{X}_i - \beta\sigma_i\right)^2}{2\sigma_\delta^2/n} \right) \tag{2.7}$$

The Joint Posterior p.d.f of β and σ_δ given in equation (2.7) facilitates estimation of the Marginal Posterior p.d.f of β. This is defined as follows:

$$p(\beta \mid \bar{X}_i,\sigma_i) = \int_0^\infty p(\beta,\sigma_\delta \mid \bar{X}_i,\sigma_i)\, d\sigma_\delta = \int_0^\infty \frac{n^{\frac{n}{2}}}{\sigma_\delta{}^{n+1}} \exp\left(-\frac{\sum_{i=1}^{n}\left(\bar{X}_i - \beta\sigma_i\right)^2}{2\sigma_\delta^2/n} \right) d\sigma_\delta \tag{2.8}$$

In this paper, Particular interest is to test if the β coefficients for two independent segments, or sub-samples, are equal. Specifically, we wish to test $H_0 : \beta_d = 0$ vs $H_0 : \beta_d \neq 0$, where β_d is defined as the difference between two

Beta parameters, e.g., $\beta_1 - \beta_2$. Hence, we are dealing with the two-sample problem of comparing slope parameters of two Regression models, i.e. the β coefficients of two independent segments following equation (2.1) of Holgersson et al. (2012). In order to test $H_0 : \beta_1 - \beta_2 = 0$ *vs.* $H_1 : \beta_1 - \beta_2 \neq 0$, we first calculate the Bayes Estimators $\hat{\beta}_1$ and $\hat{\beta}_2$,and their variances $s^2_{\hat{\beta}_1}$, and $s^2_{\hat{\beta}_2}$,and then plug in the values into the equation given as:

$$\hat{\beta}_d \pm t_{\alpha/2}\sqrt{(s_d)^2} \tag{2.9}$$

where

$\hat{\beta}_d = \hat{\beta}_1 - \hat{\beta}_2$ is the difference between the Estimators of β_1 and β_2, and $s^2_d = s^2_{\hat{\beta}_1} + s^2_{\hat{\beta}_2}$ is the corresponding combined variance of the difference between the slope parameters. Equation (2.9) hence gives the appropriate Credible Interval for β_d. The reason for broadening the Credible Interval by using *t*-distribution critical values is to account for the increased uncertainty due to not knowing σ^2_δ. Further details on Credible Intervals are available in Bolstad (2007). By the Credible Interval, defined in equation (2.9) above, we are able to test $H_0 : \beta_1 - \beta_2 = 0$ *vs.* $H_1 : \beta_1 - \beta_2 \neq 0$ at the α level of significance in a convenient way. If the value 0 lies outside the Interval we reject the null hypotheses in favour of the alternative hypotheses.

Whereas, no Frequentistic method presently exists for making Inference of the difference between two Beta parameters of equation (2.1), other than Point Estimates, the above procedure for Interval Estimation of the difference of two β parameters of equation (2.2) provides investors with a fully operational and fairly simple tool for comparing risk premiums of two segments of Stocks. In order to demonstrate the proposed method further, we will in the following section apply it on a real data set.

3. Empirical application

The empirical analysis is performed on Stocks listed on the Stockholm Stock Exchange during the period from June 1995 to June 2010, retrieved from the database Data-Stream. The data set consists of monthly excess returns. The population is split into two segments, namely Large Cap and Small Cap. This division is determined by the market capitalization of each company following the

Stockholm Stock Exchange segmentation. The Large Cap segment consists of 77 Stocks and the Small Cap consists of 131 Stocks. Scatter plots of the two different sub-populations can be found in Figures 2 and 3. By looking at the scatter plots 2 to 3, it is seen that there is an upward trend for the Large Cap while the Small Cap slope is not so steep though both are positive. Hence, as the standard deviation of the excess return increases, so does the mean excess return.

Using the software package WINBUGS (Ntzoufras, 2009), 10000 iterations (updates) have been conducted and the first 1000 iterations are discarded to reduce eventual start-up effects. In Figures 4 and 5, the Posterior density plots of the slope parameter of the Large Cap and Small Cap, respectively, can be found. Similarly, for the Large Cap and Small Cap the iterations from WINBUGS are shown in Figures 6 and 7. These plots show that the parameters traces are stable and there are no upward or downward trends. Besides, the density plots show bell-like Posterior Distributions indicating tight Posteriors.

The separate β parameters, Posterior means, standard deviations and Credible Interval of the parameter obtained from the Posterior Distribution in equation (2.8) are given in Table 1 while the 95% Credible Interval of the difference ($\beta_{Large} - \beta_{Small}$) using equation (2.9) is given in Table 2.From Table 1, it is seen that the Posterior mean of the $\hat{\beta}_{Large}$ is larger than the one for $\hat{\beta}_{Small}$. This is in line with what was expected based on the preliminary inspection of the scatter plots in Figure 2 and 3. The standard deviation is also larger for $\hat{\beta}_{Large}$ than $\hat{\beta}_{Small}$ even though the variation of the observations seems to be larger for Large Cap than Small Cap when looking at Figures 1 and 2. This is also expected since there are more observations for Small Cap than Large Cap, which in turn decreases the standard deviation.

From Table 2, it is seen that the 95% Credible Interval of the difference ($\beta_{Large} - \beta_{Small}$) is (0.0416; 0.0807). The decision rule to use this Credible Interval to test $H_0 : \beta_{Large} - \beta_{Small} = 0$ against the alternative hypothesis $H_1 : \beta_{Large} - \beta_{Small} \neq 0$ is to reject H_0 if the value 0is not included in the Interval. Since 0 is not included in the Interval in Table 2, it is concluded that the Regression coefficient for the Large Cap and Small Cap are significantly different.

4. Summary and conclusion

Financial Analysis frequently involves Risk Estimation of portfolios or estimation of excess return to standard deviation of single Stocks. However, recent research has proposed a method for estimation of a global ratio of excess return to standard deviation with respect to a specific segment of a market. In this view, an investor can compare the risk reward for two separate segments. These methods, however, only provides theories for Point Estimation which in turn limit the relevance of them. In this paper, it is shown that hypothesis testing of the possible difference between return-to-risk ratios of two market segments can easily be conducted by Bayesian Credible Intervals. Moreover, the specific use of a vague Prior facilitates the use of previously proposed Frequentistic result with the important enhancement of Inference methods. Hence, a simple but yet comprehensive tool is available for investors to quantify and compare risk rewards of different markets. The use of the method is demonstrated through an empirical investigation of the Small Cap and Large Cap segments of the Stockholm Stock Exchange.

Acknowledgements

The author is grateful to Thomas Holgersson and two anonymous reviewers for comments and suggestions that improved the paper.

Table 1: Estimates of β_{Large} and β_{Small}

	Posterior mean	Posterior SD	2.5th Percentile	97.5th Percentile
$\hat{\beta}_{Large}$	0.1195	0.00795	0.1042	0.1353
$\hat{\beta}_{Small}$	0.0583	0.0059	0.0470	0.0699

Table 2: Estimates of ($\beta_{Large} - \beta_{Small}$)

	$\hat{\beta}_d$	s_d	95% Credible Interval
$\hat{\beta}_{Large}$ - $\hat{\beta}_{Small}$	0.0612	0.0099	(0.0416; 0.0807)

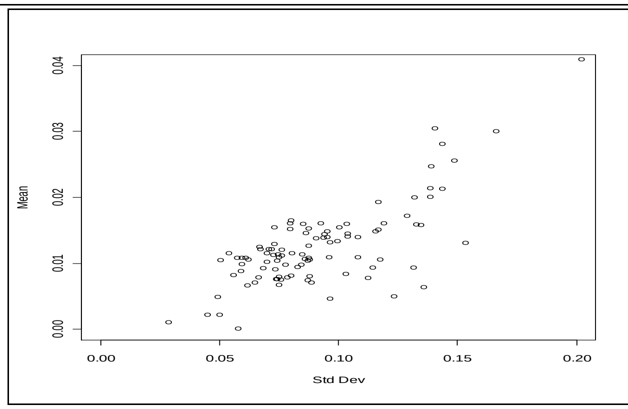

Figure 1: Mean values and standard deviation of the S&P 100

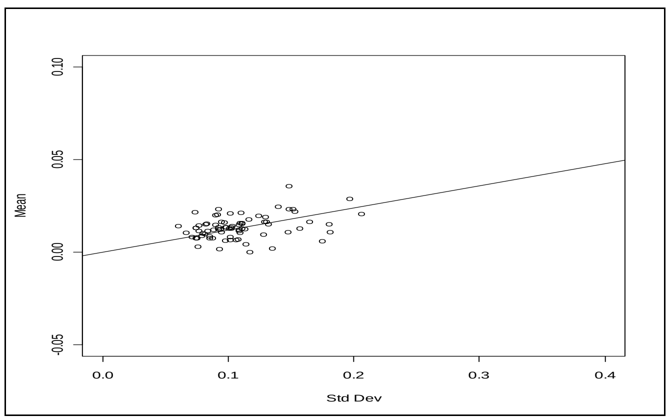

Figure 2: Scatter Plot of Large Cap

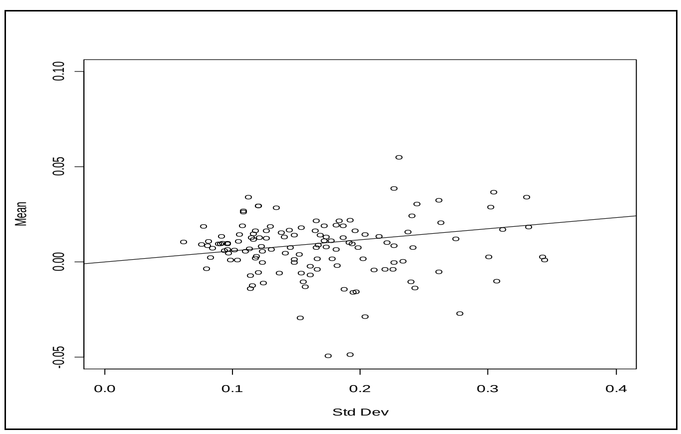

Figure 3: Scatter Plot of Small Cap

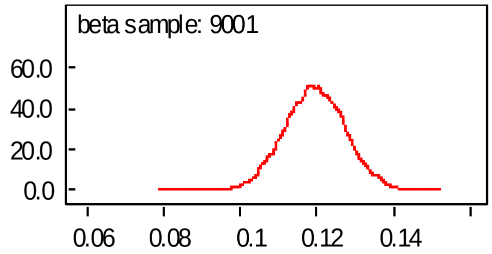

Figure 4: Posterior density plot for $\hat{\beta}_{Large}$

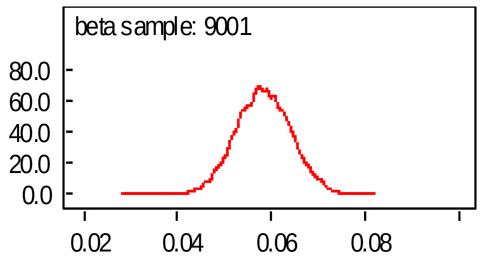

Figure 5: Posterior density plot for $\hat{\beta}_{Small}$

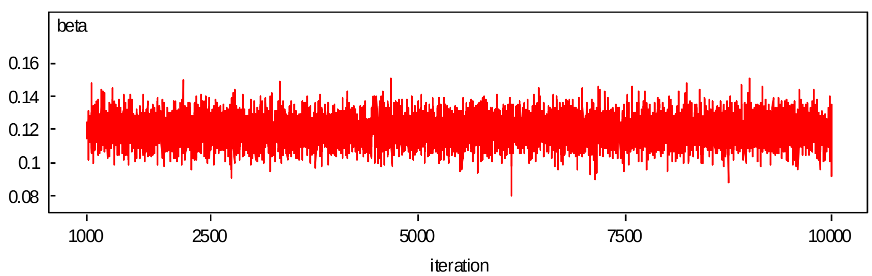

Figure 6: Trend plot for $\hat{\beta}_{Large}$

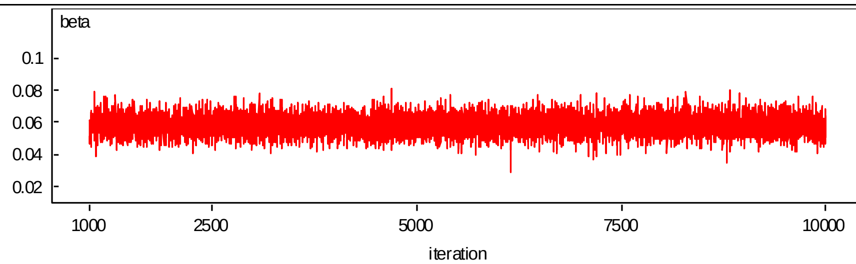

Figure 7: Trend plot for $\hat{\beta}_{Small}$

References

1. Bolstad, W. M. (2007). *Introduction to Bayesian Statistics (Second Edition).* John Wiley and Sons, New Jersey.
2. Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Boston.
3. Holgersson, H. E. T., Karlsson, P. S. and Mansoor, R. (2012). Estimating Mean-Standard deviation Ratio of Financial Data. *Journal of Applied Statistics*, **39(3)**, 657-671.
4. Lopes, A., Nezry, E. and Touzi, R. (1993). Structure detection and statistical adaptive speckle filtering in SAR images. *International Journal of Remote Sensing*, **14(3)**, 1735-1758.
5. Ntzoufras, I. (2009). *Bayesian Modeling using WinBUGS.* John Wiley and Sons, New Jersey.
6. Osteryoung, J. S., Scott, E. and Roberts, G. S. (1977). Selecting Capital projects with the coefficient of variation. *Financial Management*, **6(2)**, 65-70.
7. Sharpe, W. F. (1966). Mutual Fund Performance. *Journal of Business*, **1(2)**, 119-138.
8. Tiao, G. C. and Zellner, A. (1964). On the Bayesian Estimation of Multivariate Regression. *Journal of Royal Statistical Society: Series (B)*, **26(2)**, 277-285.
9. Zellner, A. (1971). *An Introduction to Bayesian Inferences in Econometrics.* John Wiley and Sons, New York.

Appendix

The scatter plot of Figure 1 suggests a linear relationship between the mean values and the standard deviations of the excess returns of S&P 100 Stocks. In order to confirm this relationship more formally, we wish to test

$H_0 : \beta = 0$ vs $H_0 : \beta \neq 0$ of (2.2). This is conveniently done through a 95% Bayesian Credible Interval (ranging between the 2.5th and 97.5th percentiles). From Table A1, below it is seen that this Interval of β is (0.1252; 0.15). The trend plot and Posterior density plot in Figures A1 and A2, respectively, confirm convergence of the Estimate and a tight Posterior. The decision rule to use the Credible Interval to test $H_0 : \beta = 0$ against the alternative hypothesis $H_1 : \beta \neq 0$ is to reject H_0 if the value 0 is not included in the Interval. Since 0 is not included in the Interval i.e. (0.1252; 0.15), we conclude that the Regression coefficient of model (2.2) is significantly different from zero and that a linear Regression relationship exists.

Table A1: Estimates of $\hat{\beta}$

	Posterior mean	Posterior SD	2.5th percentile	97.5th Percentile
$\hat{\beta}$	0.1347	0.0063	0.1252	0.15

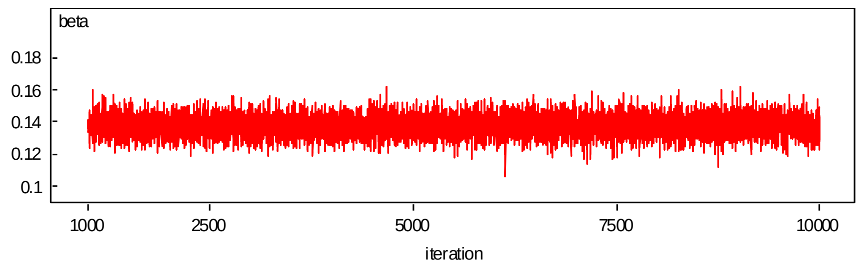

Figure A1: Trend plot of $\hat{\beta}$

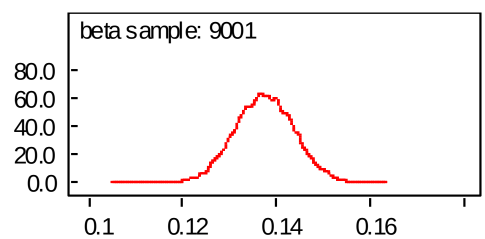

Figure A2: Posterior density plot of $\hat{\beta}$