Capital Asset Pricing Model with Interval Data

Sutthiporn Piamsuwannakit[1,2], Kittawit Autchariyapanitkul[3](✉), Songsak Sriboonchitta[1], and Rujira Ouncharoen[4]

[1] Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand
[2] Faculty of Management Science, Chiang Rai Rajabhat University, Chiang Rai, Thailand
[3] Faculty of Economics, Maejo University, Chiang Mai, Thailand
kittawit_autchariya@cmu.ac.th
[4] Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand

Abstract. We used interval-valued data to predict stock returns rather than just point valued data. Specifically, we used these interval values in the classical capital asset pricing model to estimate the beta coefficient that represents the risk in the portfolios management analysis. We also use the method to obtain a point valued of asset returns from the interval-valued data to measure the sensitivity of the asset return and the market return. Finally, AIC criterion indicated that this approach can provide us better results than use the close price for prediction.

Keywords: CAPM · Interval-valued data · Least squares method · Linear regression

1 Introduction

Capital asset pricing model provides a piece of information of asset return related to the market return via its systematic risk. In general, asset returns of any interested asset and market returns are calculated from a single-valued data. Most of the papers in financial econometrics use only closed price taking into account for calculation but in the real world stock price is moving up and down within the range of highest price and lowest price. So, in this paper we intend to use all the points in the range of high and low to improve the results in our calculations. We also put an assumption of a normal distribution on these interval-valued data.

An enormous number of research on CAPM model with single-valued data could be found in much financial research topic, the reader is referred to, e.g., William F. Sharpe [1] and John Lintner [2] only a single-valued of interest was considered. Many various technics were applied to the original CAPM model that we can found in the work from Autchariyapanitkul et al. [3], the authors used quantile regression under asymmetric Laplace distribution (ALD) to quantify

V.-N. Huynh et al. (Eds.): IUKM 2015, LNAI 9376, pp. 163–170, 2015.
DOI: 10.1007/978-3-319-25135-6_16

the beta of the asset returns in CAPM model. The results showed that this method can capture the stylized facts in financial data to explain the return of stocks under quantile, especially under the middle quantile levels. In Barnes and Hughes [4], the beta risk is significant in both tails of the conditional distribution of returns. In Chen et al. [5], the authors used a couple of methods to obtain the time-varying market betas in CAPM to analyze stock in the Dow Jones Industrial for several quantiles. The results indicated that smooth transition quantile method performed better than others methods.

Interval-valued data has become popular in many research fields especially in the context of financial portfolio analysis. Most of the financial data are usually affected by imprecision, uncertainty, inaccuracy and incompleteness, etc. The uncertainty in the data may be captured with interval-valued data. There are several existing research in the literature for investigating this issue. see Billard [6], Carvalho [7], Cattaneo [8], Diamond [9], Gil [10], K*ö*rner [11], Manski [12], Neto [14]. However, In these research papers are lacking in a foundation and theoretical background to support this idea.

The connection between the classical linear regression and the interval-valued data that share the important properties could be found for the work by Sun and Li [15]. In their paper, they provided a theoretical support framework between the classical one and the interval-valued linear regression such as least squares estimation, asymptotic properties, variances estimation, etc. However, in their paper only one of an explanatory variable can use to described the responding variable. In this paper, we intend to apply the concept of the interval-valued data to the CAPM model. We replace a single value of market returns and asset returns with the range of high and low historical data into the model.

The rest of the paper is organized as follows. Section 2 gives a basics knowledge of a linear regression model for interval-valued data. In Section 3 discusses the empirical discovering and the solutions of the forecasting problem. The last section gives the conclusion and extension of the paper.

2 A Review of Real Interval-Valued Data

Now, take a close look at financial data (D_i). Suppose, we have a range of any numbers between a minimum and maximum prices given by $D_i = [\min, \ldots, \max] = [\text{Low}, ..., \text{High}]$, where the minimum price is the "lowest price", and the maximum price is the "highest price". Certainly, this range contains the point that we called "close price". In many research papers, they are usually using the close price for calculations. A close price is a number that takes any values in the range of D_i between the lowest and the highest prices, $D_i = [\text{Low}, ...\text{Close}, ..., \text{High}]$. The close price could be either the lowest price or the highest price.

In this paper, we try to find the better value for calculations rather than a close price that is the best-represented point in the range of D_i to improving our predictions. We considered a normal distribution on this interval-valued data.

3 An Interval-Valued Data in a Linear Regression Model

Suppose we can observe an i.i.d random paired intervals variables $x_i = [\underline{x_i}, \overline{x_i}]$ and $y_i = [\underline{y_i}, \overline{y_i}],\ i = 1, 2, \ldots, n$ where $\overline{x_i}, \overline{y_i}$ are the maximum values of x_i and $\underline{x_i}, \underline{y_i}$ are the minimum values of y_i. Additionally, we can rewrite the value of x_i, y_i in the form of intervals as

$$x_i = [x_i^m - x_i^r, x_i^m + x_i^r], \tag{1a}$$

$$y_i = [y_i^m - y_i^r, y_i^m + y_i^r],\ i = 1, 2, \ldots, n, \tag{1b}$$

where x_i^m, y_i^m is the mid-points of x_i and y_i and x_i^r, y_i^r is the radii of x_i and y_i, satisfying $x_i^r, y_i^r \geq 0$. Suppose, we consider the following linear regression model given by

$$y_i = ax_i + b + \varepsilon_i,\ i = 1, 2, \cdots, n. \tag{2}$$

Analogously, it is easy to interpret the meaning of x_i, y_i by the distance of centers and radii as the following equations

$$x_i = x_i^m + \delta_{x_i},\ \delta_{x_i} \in N(0, (k_0 \Delta x_i)^2) \tag{3a}$$

$$y_i = y_i^m + \delta_{y_i},\ \delta_{y_i} \in N(0, (k_0 \Delta y_i)^2), \tag{3b}$$

where x_i^m, y_i^m are the centers of x_i and y_i, respectively. Then, $\Delta x_i = \frac{\overline{x_i} - \underline{x_i}}{2}$, $\Delta y_i = \frac{\overline{y_i} - \underline{y_i}}{2}$ are the radii of x_i and y_i, respectively and $x_i^m = \frac{\overline{x_i} + \underline{x_i}}{2}$, $y_i^m = \frac{\overline{y_i} + \underline{y_i}}{2}$ are the mid-point of x_i and y_i, respectively.Thus, given the linear regression for the interval valued data we have

$$y_i^m + \delta_{y_i} = ax_i^m + a\delta_{x_i} + b \tag{4a}$$

$$y_i^m = ax_i^m + b + (a\delta_{x_i} - \delta_{y_i}), \tag{4b}$$

where $(a\delta_{x_i} - \delta_{y_i}) \sim N(0, \sigma^2) \equiv N(0, k_0^2 a^2 \Delta x_i^2 + \Delta y_i^2)$. Assume that $a\delta_{x_i} - \delta_{y_i}$ is an independence. Thus, we can estimate parameters a, b, k_0 by the maximum likelihood function given by

$$\begin{aligned} &\max_{a,b,k_0} L(a, b, k_0 | ([\underline{x_i}, \overline{x_i}], [\underline{y_i}, \overline{y_i}]), i = 1, \ldots, n) \\ &= \max_{a,b,k_0} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi k_0^2 (a^2 \Delta x_i^2 + \Delta y_i^2)}} \exp\left[-\frac{1}{2} \frac{(y_i^m - ax_i^m - b)^2}{k_0^2 (a^2 \Delta x_i^2 + \Delta y_i^2)} \right] \right) \end{aligned} \tag{5}$$

This approach was already developed in Sun and Li [15]. And soften the criticisms of lack of theory, Manski has a whole book (see, Manski [12],[13]), this is finance not pure mathematics here. The proof of success is better fit not theorems.

3.1 Goodness of Fit in Linear Regression Model for an Interval-valued Data

In the deterministic linear regression model, we use variance to describe variation of the variable interested and so that as we knew the ratio $\frac{a^2 Var(X)}{Var(Y)} \in [0, 1]$ can be explained as an indication of goodness-of-fit. In this paper, we used the concept of the chi-squared test (χ^2) of the goodness of fit. Recall that $\sigma_{x_i} = k_0 \Delta x_i$ and $\sigma_{y_i} = k_0 \Delta y_i$ given the simple linear regression we have

$$y_i = ax_i + b \tag{6a}$$

$$y_i^m + \delta_{y_i} = ax_i^m + a\delta_{x_i} \tag{6b}$$

$$y_i^m - ax_i^m - b = a\delta_{x_i} - \delta_{y_i}, \tag{6c}$$

where $\delta_{x_i}, \delta_{y_i} \sim N(0, \sigma^2)$. Thus, we have $a^2\sigma_{x_i}^2 + \sigma_{y_i}^2$, by replacing $k_0^2(a^2\Delta x_i^2 + \Delta y_i^2)$ to above equation 6. The empirical χ^2-test is obtained by estimated this following equation

$$\chi_{cal}^2 = \sum_{i=1}^{n} \frac{(y_i^m - ax_i^m - b)^2}{k_0^2(a^2\Delta x_i^2 + \Delta y_i^2)}, \tag{7}$$

where the degree of freedom is $n - 2$.

4 An Application to the Stock Market

We consider the following financial model that is so called Capital Asset Pricing Model (CAPM). Only two sets of interval-valued data are used to explain the relationship of the asset. The fitted model is based on the least square estimation.

4.1 Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) is a linear relationship that was created by William F. Sharpe [1] and John Lintner [2]. The CAPM use to calculate a sensitivity of the expected return on the asset to expected return on the market. The combination of a linear function of the security market line:

$$E(R_A) - R_F = \beta_0 + \beta_1 E(R_M - R_F), \tag{8}$$

where $E(R_A)$ explains the expected return of the asset, R_M represents the expected market portfolio return, β_0 is the intercept and R_F is the risk-free rate. $E(R_M - R_F)$ is the expected risk premium, and β_1 is the equity beta, denoting market risk. To measure the systematic risk of each stock via the beta takes form:

$$\beta_1 = \frac{cov(R_A, R_M)}{\sigma_M^2}, \tag{9}$$

where σ_M^2 represents the variance of the expected market return. Given that, the CAPM predicts portfolio's expected return should be about its risk and the market returns.

4.2 Beta Estimation with Interval Data

From the deterministic model in equation (8), we calculate the β coefficient through the likelihood by equation (5) instead. Suppose we have observed the realization interval stock return $[\overline{R_A i}, \underline{R_A i}] = [(\overline{r_{a1}}, \underline{r_{a1}}), \ldots, [\overline{r_{an}}, \underline{r_{an}})], i = 1, 2, \ldots, n$ and return from market $[\overline{R_M i}, \underline{R_M i}] = [(\overline{r_{m1}}, \underline{r_{m1}}), \ldots, (\overline{r_{mn}}, \underline{r_{mn}})], i = 1, 2, \ldots, n$ over the past N years. These observations will be assumed an independent random. From likelihood for an interval values we have

$$\max_{a,b,k_0} L(a, b, k_0 | ([\underline{R_{Mi}}, \overline{R_{Mi}}], [\underline{R_{Ai}}, \overline{R_{Ai}}]), i = 1, \ldots, n)$$
$$= \max_{a,b,k_0} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi k_0^2 (a^2 \Delta Rm_i^2 + \Delta Ra_i^2)}} \exp \left[-\frac{1}{2} \frac{(Ra_i^m - aRm_i^m - b)^2}{k_0^2 (a^2 \Delta Rm_i^2 + \Delta Ra_i^2)} \right] \right) \tag{10}$$

4.3 Empirical Results

Our data contains 259 weekly interval-valued returns in total during 2010-2015 are obtained from Yahoo. We compute the log returns on the following stock, namely, Chesapeake Energy Corporation (CHK)and Microsoft Corporation (MSFT). Due to significant capitalization and high turnover volume.

In this paper, we use Treasury bills as a proxy. From Autchariyapanitkul et al. [3] and Mukherji [16] suggested that Treasury bills are better proxies for the risk-free rate, only related to the U.S. market.

Table 1 and Table 2 report the estimated results from equation (5). For example, the simple linear regression model for the asset returns (Y) and the market returns (X) for interval valued data for CHK is written to be

$$R_A = -0.0021 + 0.9873 R_M. \tag{11}$$

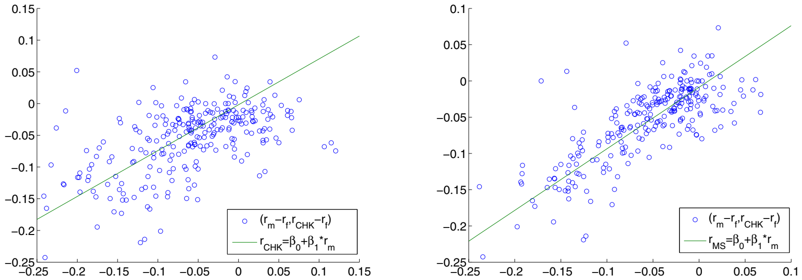
From the above linear equation, the return of a stock is likely to increase less than the return from the market. A non-parametric chi-square test is used

Table 1. Estimated parameter results for CHK

	Interval-Valued data		Point-Valued data	
parameters	values	std. Dev.	values	std. Dev.
β_0	–0.0021	0.0233	–0.0191	0.0055
β_1	0.9873	0.0914	0.7226	0.0713
k	0.4472	0.0845	-	-
MSE	-	-	0.036	
LL	525.7021	-	361.1400	-
χ^2	259.00	-	-	-
AIC	–1045.04	-	–716.28	-

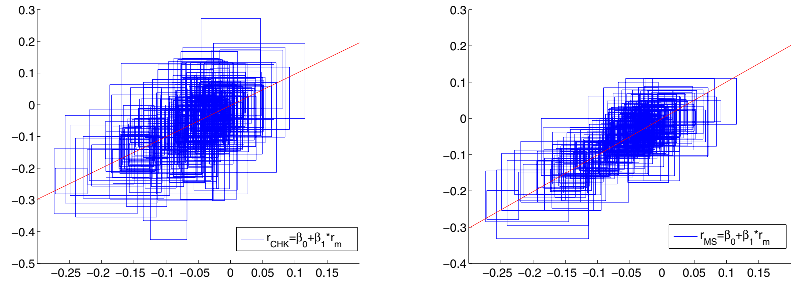
Table 2. Estimated parameter results for MSFT

	Interval-Valued data		Point-Valued data	
parameters	values	std. Dev.	values	std. Dev.
β_0	–0.0004	0.0015	–0.0088	0.0035
β_1	1.0086	0.0220	0.8489	0.0005
k	0.4017	0.0170	-	-
MSE	-	-	0.0025	-
LL	692.3808	-	478.9365	-
χ^2	259.00	-	-	-
AIC	–1378.76	-	–951.87	-



(a) CHK (b) MSFT

Fig. 1. Securities characteristic line for point valued data



(a) CHK (b) MSFT

Fig. 2. Securities characteristic line for interval valued data

to validate the method of interval-valued data. The theoretical χ^2_{n-2} gives the value of CHK, $\chi^2_{n-2} = 303.2984$ compare with the empirical value $\chi^2_{emp} = 259.00$ confirm that the market returns can be used to explain the asset returns. The model selection criteria Akaike information criterion (AIC) was employed to compare these two techniques. The AIC of interval-valued data gives a value of –1051.4402 is smaller than the AIC of pointed-valued data, which indicate that the results from the interval-valued method is more prefer than the deterministic one.

The relationship between market return and asset return are plotted in Figure 1 and Figure 2 for pointed-valued data and interval-valued data, respectively.

The rectangular are the high and low interval-valued data, and the straight line is the securities characteristic line, the slope of this straight line represent the systematic risk beta. All investments and portfolio of investments must lie along a straight line in the return beta space.

5 Conclusions and Extension

The systematic risk has played as the critical role of financial measurement in capital asset pricing model. Academic and practitioners attempt to estimate its underlying value accurately. Fortunately, there have been the novel approaches to evaluating the beta with interval-valued data. We used every price range of real world data to obtained the single value of the systematic risk same as the results from the conventional CAMP model.

In this paper, we use our approach to an interval-valued data in CAPM for only one stock in $S\&P500$ for a demonstration. With this, a method can be used to investigate the linear relationship between the expected asset returns and its asymmetric market risk by including all of the levels of prices in the range of an interval-valued data. The results clearly show that the beta can measure the responsiveness to the asset returns and market returns. However, only a systematic risk is calculated through the model, and we neglect the unsystematic risk under CAPM assumption. CAPM concludes that the expected return of a security or a portfolio equals the rate on a risk-free security plus a risk premium.

By AIC criterion, it should be noticed that the estimation by using interval-valued data more reasonable than just used the single valued in the calculations. Not only one explanatory variable can be used to explain the outcome variable but with this method also allowed us to use more than one covariate in the model.

For future research, we are interested to use this method to the time series models such as ARMA, GARCH model. Additionally, we can use this method to the model with more than one explanatory variables such as Fama and French (1993). A three-factor model can be extended the CAPM by putting size and value factors in the classical one.

Acknowledgments. The authors thank Prof. Dr. Vladik Kreinovich for his suggestions for an interval-valued data computation method. We would like to thank the referee for giving comments on manuscript.

References

1. William, F.: Capital Asset Prices A Theory of Market Equilibrium Under Conditions of Risk. Journal of Finance **19**(3), 425–442 (1964)
2. John, L.: The Valuation of Risk Assets and the Selection of Risky Investments in Stock. The Review of Economics and Statistics **47**(1), 13–37 (1965)
3. Autchariyapanitkul, K., Chainam, S., Sriboonchitta, S.: Quantile regression under asymmetric Laplace distribution in capital asset pricing model. In: Huynh, V.-N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Econometrics of Risk, vol. 583, pp. 219–231. Springer, Heidelberg (2015)
4. Barnes, L.M., Hughes, W.A.: A Quantile Regression Analysis of the Cross Section of Stock Market Returns. Federal Reserve Bank of Boston, working paper (2002)
5. Chen, W.S.C., Lin, S., Yu, L.H.P.: Smooth Transition Quantile Capital Asset Pricing Models with Heteroscedasticity. Computational Economics **40**, 19–48 (2012)
6. Billard, L.: Dependencies and variation components of symbolic interval-valued data. In: Brito, P., Cucumel, G., Bertrand, P., de Carvalho, F. (eds.) Selected Contributions in Data Analysis and Classification. Studies in Classification, Data Analysis, and Knowledge Organization, pp. 3–12. Springer, Heidelberg (2007)
7. de Carvalho, F.A.T., Lima Neto, E.A., Tenorio, C.P.: A new method to fit a linear regression model for interval-valued data. In: Biundo, S., Frühwirth, T., Palm, G. (eds.) KI 2004. LNCS (LNAI), vol. 3238, pp. 295–306. Springer, Heidelberg (2004)
8. Cattaneo, M.E.G.V., Wiencierz, A.: Likelihood-based imprecise regression. International Journal of Approximate Reasoning **53**, 1137–1154 (2012)
9. Diamond, P.: Least square fitting of compact set-valued data. J. Math. Anal. Appl. **147**, 531–544 (1990)
10. Gil, M.A., Lubiano, M.A., Montenegro, M., Lopez, M.T.: Least squares fitting of an affine function and strength of association for interval-valued data. Metrika **56**, 97–101 (2002)
11. Körner, R., Näther, W.: Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. Information Sciences **109**, 95–118 (1998)
12. Manski, C.F., Tamer, T.: Inference on regressions with interval data on a regressor or outcome. Econometrica **70**, 519–546 (2002)
13. Manski, C.F.: Partial Identification of Probability Distributions. Springer, New York (2003)
14. Neto, E.A.L., Carvalho, F.A.T.: Centre and range method for fitting a linear regression model to symbolic interval data. Computational Statistics and Data Analysis **52**, 1500–1515 (2008)
15. Sun, Y., Li, C.: Linear regression for interval-valued data: a new and comprehensive model, under review (2015). arXiv:1401.1831
16. Mukherji, S.: The capital asset pricing model's risk-free rate. International Journal of Business and Finance Research **5**, 793–808 (2011)