Predicting Stock Returns in the Capital Asset Pricing Model Using Quantile Regression and Belief Functions

Kittawit Autchariyapanitkul[1], Somsak Chanaim[2], Songsak Sriboonchitta[1], and Thierry Denoeux[3]

[1] Kittawit Autchariyapanitkul, Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 52000, Thailand
kittawit_autchariya@cmu.ac.th,
[2] Somsak Chanaim
Faculty of Science, Chiang Mai University, Chiang Mai 52000, Thailand
[3] Thierry Denoeux
UMR 7253 Heudiasyc, Université de Technologie de Compiègne, CNRS, France

Abstract. We consider an inference method for prediction based on belief functions in quantile regression with an asymmetric Laplace distribution. Specifically, we apply this method to the capital asset pricing model to estimate the beta coefficient and measure volatility under various market conditions at given levels of quantile. Likelihood-based belief functions are calculated from historical data of the securities in the S&P500 market. The results give us evidence on the systematic risk, in the form of a consonant belief function specified from the asymmetric Laplace distribution likelihood function given recorded data. Finally, we use the method to forecast the return of an individual stock.

Keywords: Asymmetric Laplace distribution, Capital Asset Pricing, Dempster-Shafer Theory, Financial data, Likelihood-based belief functions, Quantile regression.

1 Introduction

The Capital Asset Pricing Model (CAPM) is one of the most useful models in investment. In this model, asset returns are usually assumed to be jointly normally distributed random variables. However, this is not always the case. The CAPM assumes that the variance of returns adequately measures risk. This may be true if returns are normally distributed. In this paper, we propose to use quantile regression with an asymmetric Laplace distribution (ALD), coupled with an inference method based on belief functions, to estimate the parameters of the model and predict stock returns.

Quantile regression can characterize the entire conditional distribution of the outcome variable and is more robust to outliers and misspecification of the error distribution. It can also handle heteroscedasticity, as shown by Koenger. For the application of quantile regression to the CAPM, the reader is referred

to, e.g., [1]. It was found that the market price of beta risk is significant in both tails of the conditional distribution of returns.

We present a likelihood-based approach to the estimation of regression quantiles based on the asymmetric Laplace distribution. In [8], this distribution is used to model the distribution of currency exchange rates and is shown to capture the peakedness, leptokurticity (Fat tails) and skewness inherent in such data. Similarly, it is shown in [6] that the Laplace distribution has a geometric stability to represent the weekly and monthly distributions of stock returns and also models the high peak, fat tails and skewness of the returns.

Here, we use the Dempster-Shafer theory of belief functions introduced by Dempster [3] and Shafer [7]. In this approach, a piece of evidence is modeled by a belief function, which can be viewed as the distribution of a random set. This method is applied to estimation using the likelihood-based approach introduced in [7] and recently justified in [4], and to prediction using the method introduced in [5]. The main contribution of this paper is thus to propose an alternative method for drawing inferences about conditional quantiles via a likelihood-based belief function approach.

The remainder of the paper is organized as follows. Section 2 provides the background on quantile regression with asymmetric Laplace distribution and Section 3 introduces the prediction machinery using belief functions. Section 4 discusses the empirical solutions to the forecasting problem. The last section summarizes the paper.

2   Quantile Regression with an Asymmetric Laplace Distribution

Let Y be a response variable and X a vector of explanatory variables. In linear quantile regression, the conditional α-quantile $q_\alpha(Y|X)$ of Y given X is assumed to be linearly related to X through the equation $q_\alpha(Y|X) = X'\beta_\alpha$, where β_α is a vector of unknown parameters and X' is the transpose of X. Denoting the error by ε_α, we can write the quantile regression model as

$$Y = X'\beta_\alpha + \varepsilon_\alpha. \tag{1}$$

We have

$$q_\alpha(Y|X) = q_\alpha[(X'\beta_\alpha + \varepsilon_\alpha)|X] = q_\alpha(X'\beta_\alpha + \varepsilon_\alpha|X) = X'\beta_\alpha + q_\alpha(\varepsilon_\alpha|X), \tag{2}$$

since given X, $X'\beta_\alpha$ is a constant. Thus, $q_\alpha(\varepsilon_\alpha|X) = 0$, which is the counterpart of the standard condition $E(\varepsilon|X) = 0$ in the mean linear regression model. If ε_α is independent of X, the α-quantile of the noise ε_α is zero, that is, $\int_{-\infty}^{0} dF_{\varepsilon_\alpha}(u) = \alpha$. For $q_\alpha(Y|X) = X'\beta_\alpha$, we see that β_α minimize $E[\rho_\alpha(Y - X'\beta)]$ over β, where $\rho_\alpha(\cdot)$ is the so-called check (or loss) function defined by

$$\rho_\alpha(u) = u(\alpha - 1_{(u<0)}), \tag{3}$$

with $1_{(u<0)}$ denoting the usual indicator function. Thus, given i.i.d (X_i, Y_i), a plausible estimator of β_α is

$$\widehat{\beta_\alpha} = \arg\min \frac{1}{n}\{\sum_{i=1}^{n} \rho_\alpha(Y_i - X_i'\beta)\}. \tag{4}$$

This estimator is called the Least Absolute Deviation (LAD) estimator.

Suppose that the error term ε_α has an ALD with mean 0 and standard deviation σ_α:

$$f_{\sigma_\alpha}(\varepsilon_\alpha) = \frac{\alpha(1-\alpha)}{\sigma_\alpha} \exp\left\{-\rho_\alpha\left(\frac{\varepsilon_\alpha}{\sigma_\alpha}\right)\right\}. \tag{5}$$

Then, minimizing the absolute deviation is equivalent to maximizing the likelihood and the LAD estimator of β_α is a maximum likelihood estimator (MLE). The likelihood function for $\beta_\alpha, \sigma_\alpha$ after observing the data $D = (X_1, Y_1), \ldots, (X_n, Y_n)$ is

$$L_D(\beta_\alpha, \sigma_\alpha) = \frac{\alpha^n(1-\alpha)^n}{\sigma_\alpha^n} \exp\left\{-\sum_{i=1}^{n} \rho_\alpha\left(\frac{Y_i - X_i'\beta_\alpha}{\sigma}\right)\right\}. \tag{6}$$

3 Statistical Inference and Prediction Using Belief Functions

3.1 Likelihood-Based Belief Functions

Suppose we observe a realization x of the random vector X with probability density function (pdf) $p_\theta(x)$, where $\theta \in \Theta$ is an unknown parameter. In this paper, we use the method proposed by Shafer [7], which can be derived from the Likelihood Principle (LP) and the Least Commitment Principle (LCP) [4]. According to the LP, all the information about Θ is represented by the likelihood function defined by $L_x(\theta) = p_\theta(x)$ for all $\theta \in \Theta$. In statistics, the likelihood ratio has the meaning of a "relative plausibility", which can be written as:

$$\frac{pl_x(\theta_1)}{pl_x(\theta_2)} = \frac{L_x(\theta_1)}{L_x(\theta_2)}, \tag{7}$$

for all $(\theta_1, \theta_2) \in \Theta^2$ or, equivalently, $pl_x(\theta) = cL_x(\theta)$, for all $\theta \in \Theta$ and some positive constant c. The LCP then implies that the highest possible value should be given to constant c [4], which leads us to equating the contour function pl_x with the relative likelihood:

$$pl_x(\theta) = \frac{L_x(\theta)}{\sup_{\theta\in\Theta} L_x(\theta)}. \tag{8}$$

The information about θ is represented by the consonant belief function Bel_x^Θ with contour function pl_x, i.e., with corresponding plausibility function

$Pl_x^{\Theta}(A) = \sup_{\theta \in A} pl_x(\theta)$, for all $A \subseteq \Theta$. The focal sets of Bel_x^{Θ} are the levels sets of pl_x defined as follows:

$$\Gamma_x(\omega) = \{\theta \in \Theta | pl_x(\theta) \geq \omega\}, \tag{9}$$

for $\theta \in [0, 1]$. These sets are also called *plausibility regions*. The consonant belief function Bel_x^{Θ} is equivalent to the random set induced by the Lebesgue measure λ on [0,1] and the multi-valued mapping Γ_x from $[0, 1] \to 2^{\Theta}$ (see, [5]). We remark that the MLE of θ is the value of θ with highest plausibility.

3.2   Prediction Using Belief Functions

Let X be a random variable with parametric density function $f_\theta(x)$ for $\theta \in \Theta$ and assume that we have observed $X = x$. Given the belief function Bel_x^{Θ} about θ, we can predict the future value of a random variable Y whose pdf $g_\theta(y)$ also depends on θ. In the approach introduced in [5], Y is written as a function of the parameter θ and an unobserved auxiliary variable $u \in \mathbb{U}$ with known probability measure μ not depending on θ:

$$Y = \varphi(\theta, u). \tag{10}$$

Using Equations (9) and (10), we can compose the multi-valued mapping Γ_x from $[0, 1] \to 2^{\Theta}$ with φ to get a new multi-valued mapping Γ_x' from $[0, 1] \times \mathbb{U}$ to $2^{\mathbb{Y}}$ defined as

$$\Gamma_x' : [0, 1] \times \mathbb{U} \to 2^{\mathbb{Y}}$$
$$(\omega, u) \to \varphi(\Gamma_x(\omega), u).$$

We can then define the predictive belief $(Bel_x^{\mathbb{Y}})$ and plausibility $(Pl_x^{\mathbb{Y}})$ functions on $\mathbb{Y}$ as

$$Bel_x^{\mathbb{Y}}(A) = (\lambda \otimes \mu)(\{(\omega, u) \in [0, 1] | \varphi(\Gamma_x(\omega, u) \subseteq A\}). \tag{11a}$$
$$Pl_x^{\mathbb{Y}}(A) = (\lambda \otimes \mu)(\{(\omega, u) \in [0, 1] | \varphi(\Gamma_x(\omega, u) \cap A \neq \emptyset\}), \tag{11b}$$

for all $A \subseteq \mathbb{Y}$.

4   Application to Stock Market Prediction

4.1   Model

The CAPM measures the sensitivity of the expected excess return on security to expected market risk premium. The equation of CAPM is a linear function of the security market line:

$$E(R_A) - R_F = \beta_0 + \beta_1 E(R_M - R_F), \tag{12}$$

where $E(R_A)$ is the expected return of the asset, R_M is the expected market portfolio return, R_F is the risk free rate, β_0 is the intercept and β_1 is the equity beta, representing market risk. Suppose we have observed the historical returns of stock $R_A = (r_{a1}, \cdots, r_{an})$ and returns from market $R_M = (r_{m1}, \cdots, r_{mn})$. The errors will be assumed to be iid with density function (5). The likelihood function is given by (6).

4.2   Experimental Results

The data contain the weekly log returns of the integrated oil and gas company, Chesapeake Energy (CHK), during 2010-2013. The ML estimates of the parameters are shown in Table 1 for different values of α.

Table 1. Parameter estimation results. Standard errors are given in parentheses.

Stock Name	Parameters	$\alpha = 0.40$	$\alpha = 0.50$	$\alpha = 0.60$
CHK	β_0	-0.011 (0.002)	-0.004 (0.003)	0.005 (0.000)
	β_1	1.379 (0.163)	1.442 (0.005)	1.304 (0.017)
	σ	0.002 (0.001)	0.016 (0.001)	0.016 (0.010)

We used a nonlinear optimization algorithm to maximize the likelihood (6) with respect to $\theta = (\beta_{(0,\alpha)}, \beta_{(1,\alpha)}, \sigma_\alpha)$. The plausibility function on θ is then defined by the relative likelihood (8) and the marginal contour function on a specific parameter is obtained by take the supremum with respect to the others parameters, e.g.,

$$pl_{R_A}(\beta_{(0,\alpha)}) = \sup_{\beta_{(1,\alpha)},\sigma} pl_{R_A}(\beta_{(0,\alpha)}). \tag{13}$$

Figure 1 displays two-dimensional marginal contour functions and Figure 2 shows the marginal contour functions for parameters β_0, β_1 and σ. The 0.15 threshold corresponds to an approximate 95% confidence interval and gives us an interval of plausible values of each of the three parameters.
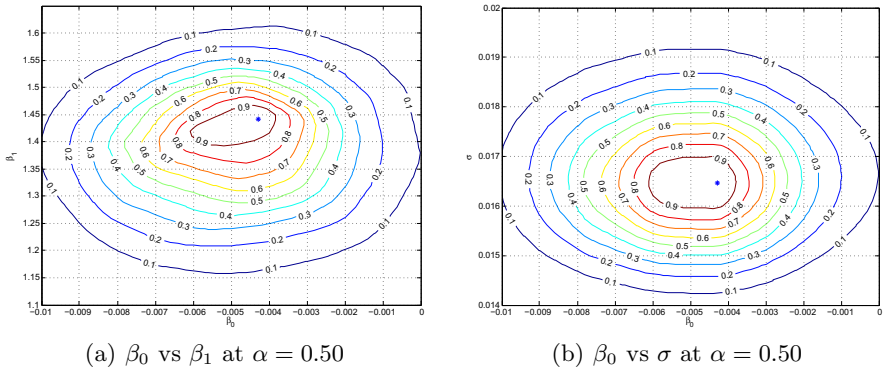

(a) β_0 vs β_1 at $\alpha = 0.50$ (b) β_0 vs σ at $\alpha = 0.50$

Fig. 1. Marginal contour functions $pl_x(\theta_1, \theta_2)$ in two-dimensional parameter subspaces

(a) β_0 at $\alpha = 0.5$

(b) β_1 at $\alpha = 0.5$
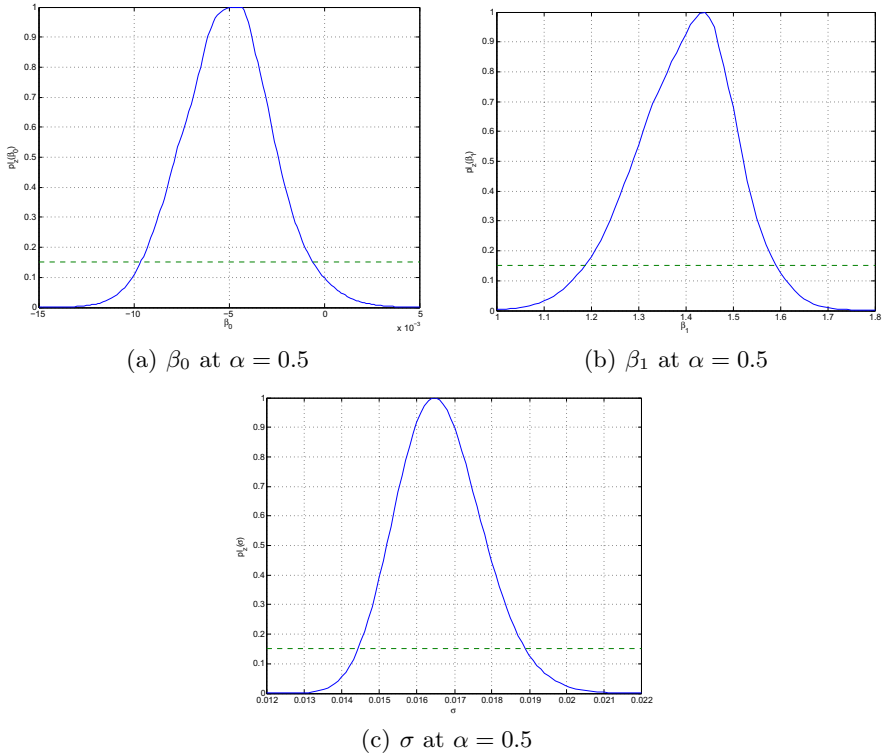
(c) σ at $\alpha = 0.5$

Fig. 2. Marginal plausibility of $\beta_0^{0.5}$ (a), $\beta_1^{0.5}$ (b) and $\sigma^{0.5}$ (c)

To predict the expected return of the asset $r_{a,n+1}$ for a new market portfolio return $r_{m,n+1}$, we compute the minimum and maximum of $r_{a,n+1}$ given $r_{m,n+1}$ at fixed α by

$$r_{a,n+1} = \beta_{0,\alpha} + \beta_{1,\alpha} r_{m,n+1} + \sigma_\alpha F_{\varepsilon_\alpha}^{-1}(u), \tag{14}$$

under the constraint $pl_{R_A}(\beta_{i,\alpha}, \sigma_\alpha) > \omega$, where $F_{\varepsilon_\alpha}^{-1}$ is the inverse cumulative distribution function (cdf) of the asymmetric Laplace distribution $ALD(\alpha, 0, 1)$ and u, ω are independent random variables with the same uniform distribution $U([0,1])$. Given (14), we randomize independently N pairs of the random number $(\omega_i, u_i), i = 1, 2, \cdots, N$ resulting in N intervals $[r_a^L(\omega_i, u_i), r_a^U(\omega_i, u_i)]$. For any $A \subseteq \mathbb{R}$, the stock returns $Bel_{r_{ai}}(A)$ and $Pl_{r_{ai}}(A)$ can be estimated by equation (11). The estimated lower and upper expectations of $r_{a,n+1}$ are then:

$$\overline{R}_A^L = \sum_{i=1}^{N} \frac{r_a^L(\omega_i, u_i)}{N} \tag{15a}$$

$$\overline{R}_A^U = \sum_{i=1}^{N} \frac{r_a^U(\omega_i, u_i)}{N}. \tag{15b}$$

Figures 3 displays the lower and upper cdfs $Bel_{R_A}((-\infty, R_A])$ and $Pl_{R_A}((-\infty, R_A])$ at given $r_m = 0.05$. This function give us the summary of the predictive belief function Bel_{R_A}. Figure 4 shows the upper and lower predictive quantiles of the stock returns (see [5]), defined by the inequalities $pl(R_A \geqslant q_{\alpha'}^L) = \alpha'$ and $pl(R_A \leqslant q_{\alpha'}^U) = \alpha'$. As shown in [5], the following inequalities hold:

$$Bel(q_{\alpha'}^L \leqslant R_A \leqslant q_{1-\alpha'}^U) \geqslant 1 - 2\alpha'. \qquad (16)$$
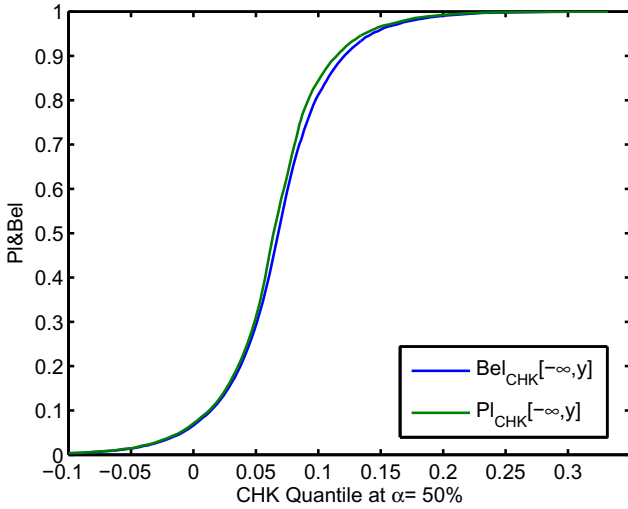

Fig. 3. Upper and lower cumulative distribution function given $r_m = 0.05$
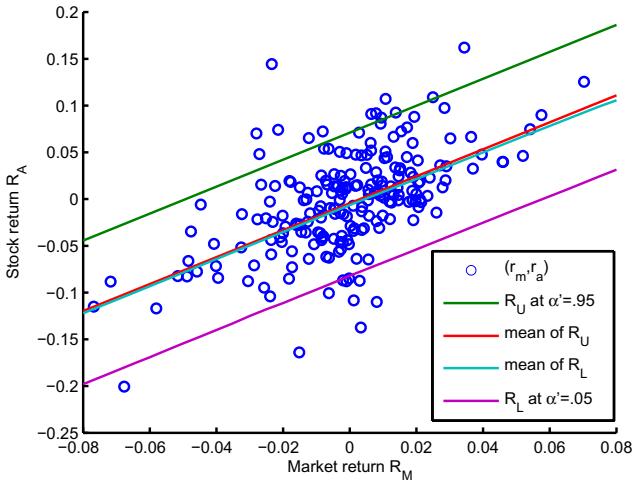

Fig. 4. Lower and upper quantiles of stock returns at $\alpha = 0.5$

Lower-upper expectations and quantiles are other representations of prediction uncertainty, taking into account both parameter estimation uncertainty and randomness. For these data, the very narrow gap between the lower and upper cdfs shows that estimation uncertainty is small as compared to random uncertainty. In practice, these results can be used to increase the performance of the investment portfolio.

5   Conclusions

In this paper, we studied the method of quantile CAPM with ALD for stocks in S&P500 in the belief function framework. We used the Dempster-Shafer theory of belief functions to model the uncertainty referring to the statistical prediction based on historical data and a financial model. This method consists of two steps. First, a consonant belief function representing the uncertainty on the parameter vector θ is defined from the normalized likelihood function given the past data. The return of individual stock R_A is then expressed as $\varphi(\theta, u)$, where u is a stochastic variable with known distribution and the beliefs on θ and u are transferred through φ, resulting in a belief function on R_A. This approach has been adapted to the prediction of stock returns. A possible extension of this work is to consider uncertainty on the independent variable r_m, which can also be expressed as a belief function and combined with other uncertainties to compute a belief function on R_A. This issue will be addressed in future work.

References

1. 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)
2. Ben Abdallah, N., Mouhous Voyneau, N., Denoeux, T.: Combining statistical and expert evidence using belief functions: Application to centennial sea level estimation taking into account climate change. International Journal of Approximate Reasoning 55, 341–345 (2014)
3. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics 38, 325–339 (1967)
4. Denoeux, T.: Likelihood-based belief function: justification and some extensions to low-quality data. International Journal of Approximate Reasoning (2013) (accepted for publication), http://dx.doi.org/10.1016.j.ijar.2013.06.007
5. Kanjanatarakul, O., Sriboonchitta, S., Denoeux, T.: Forecasting using belief functions: an application to marketing econometrics. International Journal of Approximate Reasoning 55(5), 1113–1128 (2014)
6. Linden, M.: A Model for Stock Return Distribution. International Journal of Finance and Economics 6, 159–169 (2001)
7. Shafer, G.: A mathematical theory of evidence. Princeton University Press, Princeton (1976)
8. Sánchez, B.L., Lachos, H.V., Labra, V.F.: Likelihood Based Inference for Quantile Regression Using the Asymmetric Laplace Distribution. Journal of Statistical Computation and Simulation 81, 1565–1578 (2013)