MANAGEMENT SCIENCE AND ENGINEERING
Vol. **5**, No. **1**, **2011**, pp. 19-25
www.cscanada.org

ISSN 1913-0341 [Print]
ISSN 1913-035X [Online]
www.cscanada.net

Extreme Value Analysis for Record Loss Prediction during Volatile Market

Utkarsh Shrivastava[1]
Gaurav Dawar[2]
Savita Dhingra[3]
Malika Rani[4]

Abstract: Year after year stock markets of the world kept on breaking records. They reached new heights and plunged to new depths. During financial crisis of 2008 many markets shed as many points as they never did in their history. It is extremely difficult to predict future index value due to their high randomness but is it possible to know if markets are going to achieve a record fall in near future or not. Daily changes in stock market index are not normally distributed, analysis showed they exhibit fatter tails that normal distribution, while extreme fall and rise generally follow generalized extreme value distribution explained by Extreme Value theory. The study models worst losses suffered in a day by National Stock Exchange index CNX-Nifty by fitting GEV distribution on yearly and quarterly maximum losses. GEV distribution function hence obtained is used for predicting probability of obtaining a record maximum loss next year / quarter of 2008. As Indian markets shed maximum point in a day during financial crisis of 2008, study verifies if model gives indication about such extreme event.
Key words: GEV distribution; Extreme Value Theory; Record Loss; Frechet Density Function; Block Maxima

1. INTRODUCTION

Financial risk managers in the world keep on predicting the extremities which their returns and assets will be subjected to. They analyze probabilities and magnitudes of potential losses such as trading crisis, market scandals currency crisis and large bond defaults. These extreme events are high quintile and high probability events in the probability distribution. Today a number of parametric and non parametric techniques exists to analyze a stationary time series and they give acceptable results within specified constraints but all series are not perfectly stable as there are outliers and should be dealt separately. Extreme value analysis basically deals with all these outliers in the time series and develops probability distributions to predict them and study their nature. Without extreme value analysis one will have to study

[1] Assistant Professor, Lovely School of Business, Lovely Professional University, Punjab (INDIA) -144402
[2] Lecturer , NC college of Engineering, Panipat, Haryana (INDIA) – 131001
[3] Lecturer, School of Business, Lovely Professional University, Punjab (INDIA) -144402
[4] Lecturer, School of Business, Lovely Professional University, Punjab (INDIA) -144402
*Received 11 February 2011; accepted 10 March 2011.

quartiles and tail probabilities which are unobservable if we analyze normal data set. Hence all the extreme values are taken away from the data set and analyzed separately.

Extreme Value Theory (EVT) has gained much attention lately (Reiss and Thomas, 1997; Leadbetter et al, 1993; Embrechts et al, 1997) and there have been a number of applications in the field of finance (Longin, 1996; Longin, 2000; Longin and Solnik, 1998; Danielson and De Vries, 1997; Danielson and De Vries, 1998; Danielson et al, 1998; Diebold et al, 1999; Emmer et al, 1998; McNeil, 1998; McNeil and Frey, 1998; Ferreira, Mendes and Duarte Jr, 2000; Ho, Burridge, Cadle and Theobald, 2000; among other studies). Understanding the influence of extreme market events, such as the East Asian financial crises, is of great importance for risk managers (Ewing,1995; Longin, 2000). Since all risk measurement methodologies used to estimate the Value-at-Risk (VaR) of a portfolio assume that the market behavior is stable, extreme market events demand a special approach from risk managers. A more recent methodology to estimating VaR focuses on modeling the tail of the distribution based on extreme value theory Longin (2000), Danielson and De Vries (1997).

Indian stock markets and those of other developing countries are centers of foreign investment post recession but they also had to face extreme events during downturn of 2008 as CNX-Nifty registered its biggest fall in terms of number of index points. Index had registered its biggest fall in 21-jan-2008 of 496.5 points, second biggest fall ever was on 24-Oct-2008 of 359.15 points. So two biggest fall of NSE are in 2008. Did extreme value analysis predict record fall in the year 2008 through analysis of historical market data till 2007. This study cross verifies the statement by predicting probability of a record breaking fall in 2008. Dataset consists of daily observations of CNX-Nifty index of NSE (National Stock Exchange) i.e from 3-Jul-1990 till 31-dec-2007 and predictions for 2008 are made using this.

2. EXTREME VALUE THEORY

Extreme Value Theory is about behavior and patterns of extreme values (maxima or minima) of random variable. Its role is as important is central limit theorem which is used to model the sum of random variables. To understand its concept let V_1 , V_2 , V_3 …. be independent identical distributed random variables representing risk and losses and following an unknown cumulative distribution function F(v) = $\Pr(V_i \le v)$.As in this study V_i would represent index points shed by CNX-Nifty index in a day due to some extraordinary event. Since risks or losses are being modeled, hence these would be taken as positive and are generally found on the right tail of loss distribution F. If W_n = Max (V_1 ,.. V_n) or is the worst loss in the sample of n losses. Extreme value theory basically deals with the distribution of W_n. On using iid assumption CDF of W_n will be given by

$$\Pr\{W_n \le v\} = \Pr\{V_1 \le v, .. V_n \le v\} = \prod_{i=1}^{n} F(v) = F^n(v) \qquad (1)$$

$F^n(v)$ is unknown and is poorly estimated by empirical distribution function . Fisher and Tippett in 1928 devised a theorem to make inference about $F^n(v)$ and get asymptotic inferences about $F^n(v)$. As central limit theorem applies to normalized sum of random variables, the fischer tippet theorem applies to Asymptotic approximation is based on standardize maximum value,

$$M_n = (W_n - \mu_n) / \sigma_n \qquad (2)$$

where μ_n is location measure and σ_n is scale measure such that both are sequence of real numbers and σ_n > 0. Using this theorem in current situation if standardize maximum of worst falls in the index point converges to some non degenerate function then it will be Generalized extreme value distribution. Its equation is given by

$$H_\xi(x) = e^{-(1+\xi x)^{-1/\xi}} \quad \text{for } \xi \ne 0 \text{ and } 1 + \xi x > 0 \qquad (3)$$

$$H_\xi(x) = e^{(-x)^{-x}} \quad \text{for } \xi = 0 \text{ and } -\infty \le x \le \infty \qquad (4)$$

The parameter ξ here is called shape parameter and determines tail behavior of GEV distribution. GEV distribution characterizes the (3) and (4) characterizes the limiting distribution of standardize maximum (2). For large values of n the Fischer Tippet Theorem may be interpreted as follows

$$\Pr\{M_n < m\} = \Pr\{(W_n - \mu_n)/\sigma_n \le m\} \quad (5)$$

If $x = \sigma_n m + \mu_n$ then

$$\Pr\{W_n \le x\} \approx H_{\xi,\mu,\sigma}\{(x - \mu_n)/\sigma_n\} \approx H_{\xi,\mu,\sigma}(x) \quad (6)$$

The Cumulative Distribution Function F of the data being analyzed decides the characteristics of H_ξ. If the tail of F declines then H_ξ is of the Gumbel type and equal to zero . Generally thin tailed distributions like normal, log normal, exponential and gamma make Gumble type GEV. If tail of F declines by a power function then H_ξ is of Frechet type and $\xi > 0$. If CDF of data is fat tailed distribution like Pareto, Cauchy, student's T then H_ξ will be Frechet type. If F is finite then H_ξ is of Weibull type and $\xi < 0$.

3. METHODOLOGY

As a GEV distributions is defined by σ_n , μ_n and Maximum likelihood estimation is used to obtain them. If V_1 , V_2 , V_3 be identically distributed losses of CNX-Nifty on daily basis with unknown CDF F and T be total number of days under consideration. Let W_T be sample maximum or in this case maximum points shed by CNX-Nifty. Since whole sample has only one maximum hence, it's not possible to estimate required parameters using MLE. If we divide data into sub parts let on yearly basis, then yearly maxima can be obtained and log likelihood function for parameters σ_n , μ_n and ξ can be generated for M_T. In this study daily losses are further divided into blocks on yearly basis. CNX-Nifty historical data is available from Jul-1990. Hence, data to be analyzed is from Jul-1990 till Dec-2007. If divided on yearly basis total 19 blocks are obtained. Suppose, these blocks are represented as,

$$[\, V_1, \ldots\ldots V_n \mid V_{n+1} \ldots\ldots V_{2n} \mid \ldots\ldots \mid V_{(m-1)(n+1)} \ldots\ldots V_{nm} \,] \quad (7)$$

If W_n^j is maximum value of V_i in block j = 1,2…m . The likelihood ratio function for the parameters σ_n , μ_n and ξ of GEV distribution is obtained from the sample of block maxima [W_n^1 , W_n^2 ,… W_n^m . For such sample of block maxima, if we assume that these are iid observations from a GEV distribution with $\xi \neq 0$ then log likelihood function is given by

$$L(\sigma_n, \mu_n, \xi) = -m\ln(\sigma) - (1 + 1/\xi)\sum_{i=1}^{m} \ln[1 + \xi\{(W_n^j - \mu)/\sigma\}] - \sum_{i=1}^{m}[1 + \xi\{(W_n^j - \mu)/\sigma\}^{-1/\xi}] \quad (8)$$

such that $[1 + \xi\{(W_n^j - \mu)/\sigma\}] > 0$.

Embrechts et. al. (1997) and Coles (2001) discussed details of maximum likelihood estimation. McNeil(1998) showed how finite sample properties of MLE will depend upon the number of blocks m and block size n. The bias of MLE is reduced by increasing block size n and the variance of the MLE is reduced by increasing number of blocks m.

4. EMPIRICAL ANALYSIS

S&P CNX Nifty is a well diversified 50 stock index of National Stock Exchange (NSE) accounting for 23 sectors of the Indian economy. It is computed based on free float economy. Nifty stocks represent about 56% of the Free Float Market Capitalization. Difference in current day closing price and previous closing prices is recorded from 4-jul-1990 till 31-dec-2007. Table 1 shows that biggest fall in a day of index was 270.7 points till 2007 but in 2008 all records were broken and index shed 496.5 points. Daily data is then subdivided into 19 blocks on yearly basis to obtain block maxima or worst loss in a year so as to estimate parameters of GEV distribution. Next probability to attain a new record in terms of maximum loss in the next year 2008 is obtained. Normal QQ plot of data (Figure 2) shows that for values with lesser probabilities or tail values are quite different from normal distribution i.e change in index closing value w.r.t previous day close has fatter tail and thus indicates that it follows Frechet family of GEV distribution with $\xi > 0$ for yearly block maxima of negative changes. Table 1 shows statistics of closing price difference. Largest fall in the market is of 270.7 points on 17 Dec-2007.

Table 1: Data Summary from 4-jul-1990 till 31-Dec-2007

Year	Min	1st Qu.	Mean	Median	3ed Qu.	Max	Total	Std. Dev.
1990-07	-270.70	-9.635	1.407229	1.3000	12.30000	289.700	4168	29.4323
2008	-496.5	-78.52	-12.57146	-9.1500	59.200	349.900	247.0	113.3489

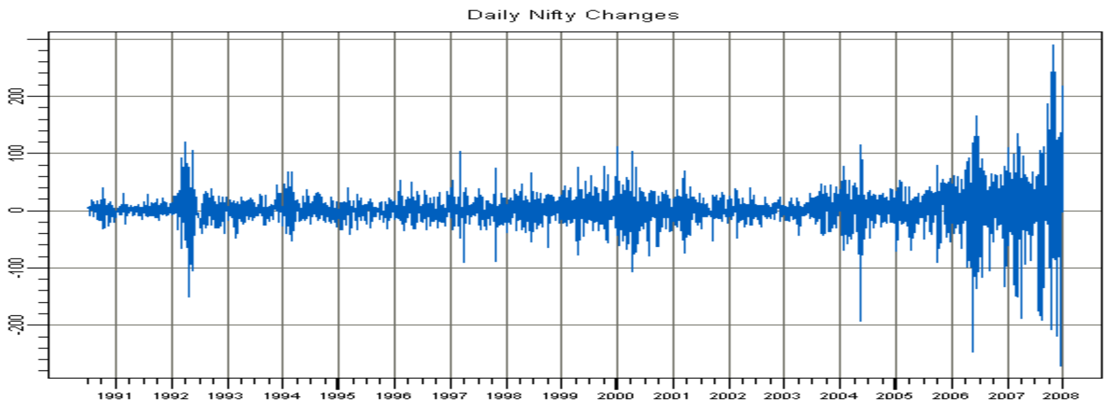

Figure 1: Plot of Daily Changes in Nifty Prices from 1991 till 2008

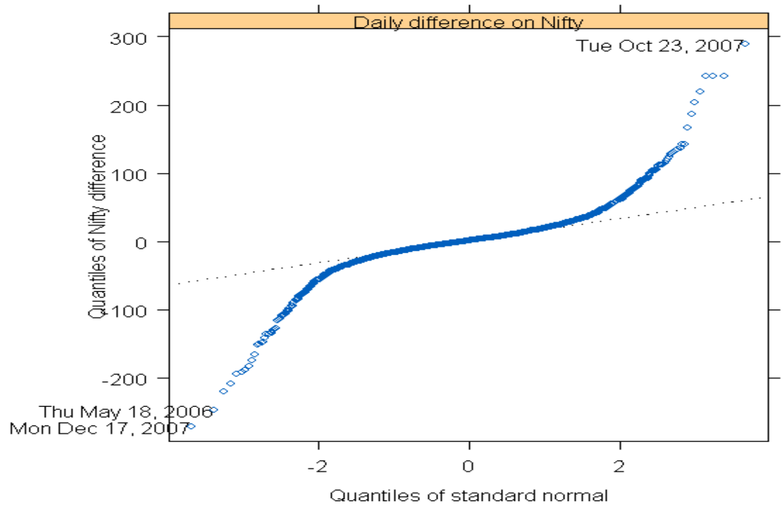

Figure 2: Normal QQ Plot for Daily Changes in Index From Jul-1990 till Dec-2007

Analysis of Block Maxima

As study focuses on maximum points which index has shed in a day and to find the probability that the next year may see biggest drop in a day. Daily losses from 1990 till 2007 are subdivided into 18 blocks on yearly basis and maximum losses are separated to fit GEV distributions and calculate the values of parameters σ_n, μ_n, ξ. Figure 3, 4 and 5 depicts the exploratory data analysis of annual block maxima. Largest negative loss in an annual block is in 2007 of 270.7 points. Histogram plot in Figure 4 resembles a Frechet density. The plot of record development

Figure 5 records and illustrates the developments of records (new maxima) for the daily index loss along with the expected number of records for iid data. This figure depicts a unique pattern which three quick records after a long gap. It is inconsistent with iid assumption. Table 2 and 3 gives parameter values for GEV fit for yearly and quarterly maximum losses and probability for a new record in next year. Log likelihood value Table 3 of GEV fit on annual maximum loss is greater than the same for quarterly maximum loss is a day. As more the value of Log likelihood function greater the probability of occurrence of given annual data for estimated values of model parameter. Table also gives probability of obtaining a record loss next year which comes out to be 3.2% and for next quarter is .9 %

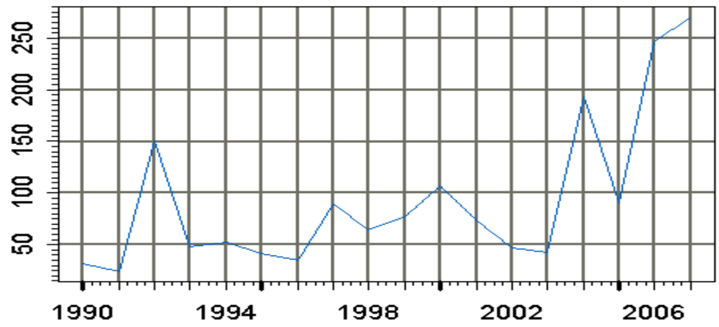

Figure 3: Plot of Block Maxima (i.e maximum one day loss) Yearly Basis

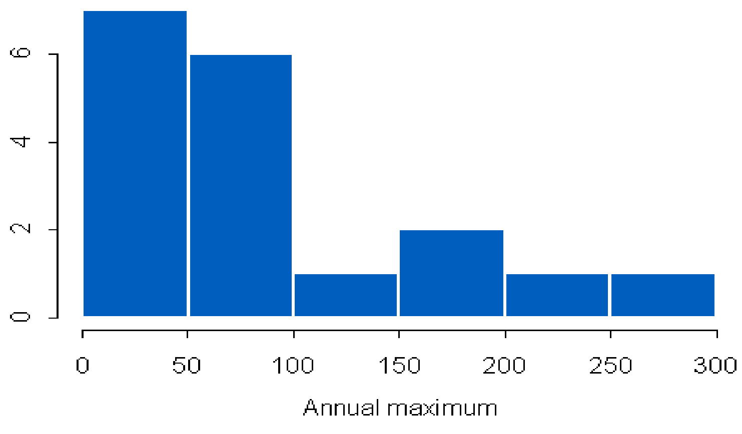

Figure 4: Histogram Showing Yearly Maximum and Its Frequency

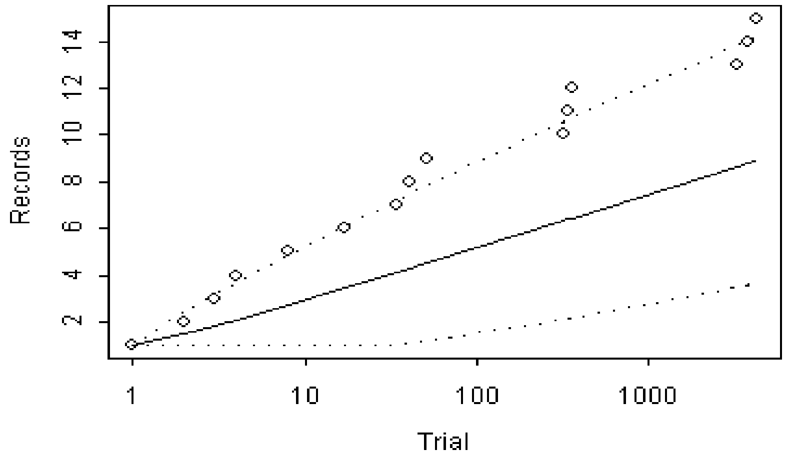

Figure 5: Plot of Records vs i.i.d

Table 2: Parameters of GEV Distribution When Fit on Yearly and Quarterly Maximum Loss

Parameter	Estimates	Standard Errors	t-ratios	Block size
ζ	0.2052	0.1780	1.1532	year
ξ	0.2637	0.0770	3.4263	quarter
σ	44.1366	9.0058	4.9009	year
σ	25.7355	2.7310	9.4233	quarter
μ	52.1128	11.3391	4.5959	year
μ	35.6930	3.3271	10.7279	quarter

Table 3: Prediction Based on GEV Distributions

Block Size	Log Likelihood Value for GEV fit	Probability that next year / Quarter would lead to new record minima?
Year	-104.3	0.032272
Quater	-354	0.00951

CONCLUSION

Extreme value analysis CNX-Nifty index firstly concludes that daily losses are not normally distributed as generally considered. They exhibit fat tail behavior and resembles greatly with Frechet type GEV distribution. Analysis of maximum and minimum loss on yearly and quarterly basis shows, annual maximum losses fits better to GEV distribution than quarterly losses. Plot of records shows a unique trend according to which initially records will be tumbled steadily but later records may tumble in a bunch of three i.e. three big losses in quick succession and then steady for a longer period which is inconsistent with the assumption of i.i.d. Log likelihood value suggests that GEV distribution fitted better with yearly maximum values. Figures which come out after the analysis are not so supportive to accurately predict extreme events of 2008, however on yearly basis it gives little indication though 3.2% probability that coming year will be unstable and prediction for next quarter to have new record minimum is .9%, which can be attributed to poor fit in case of quarterly data.

REFERENCES

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. London: Springer-Verlag.

Danielsson, J., and C. De Vries. (1997). *Value at Risk and Extreme Returns, Discussion Paper*, 273, LSE Financial Markets Group. London School of Economics.

Danielsson, J., and C. De Vries. (1998). *Beyond the Sample: Extreme Quantile and Probability Estimation*. Discussion Paper, 298, LSE Financial Markets Group. London School of Economics.

Danielsson, J., Hartmann, P., and C. De Vries. (1998). The Cost of Conservatism, *Risk, 11*, 103-107.

Embrechts,P., Kloppelberg C., and T. Mikosch (1997). *Modelling Extrenal Events*, Berlin: Springer-Verlag,.

Ewing, R., (1995). *Identifying, Measuring and Managing Market Risk and Stress Testing Portfolios*. U.K., London: Risk Conferences & Publications.

Ferreira, R., Mendes, B., and A. Duarte Jr. (2000). Extreme Market Events in Latin American Stock Markets. *Journal of Emerging Markets, 5* (1), 21-48.

Leadbetter, M., Lindberg, G., and H. Rootzen. (1993). *Extremes and Related Properties of Random Sequences and Processes*. Berlin: Springer-Verlag.

Longin, F. (1996). The Asymptotic Distribution of Extreme Stock Market Returns. *Journal of Business, 69* (3), 383-408.

Longin, F. (2000). From Value at Risk to Stress Testing: the Extreme Value Approach. *Journal of Banking and Finance, 24* (7), 1097-1130.

Longin, F., and B. Solnik. (1998). Correlation Structure of International Equity Markets during Extreme Volatile Periods. *Working Paper*, 97-039 ESSEC, Cergy-Pontoise, France.

McNeil A.J. and T. Saladin. (2000). *Developing Scenarios for Future Extreme Losses Using the POT Method*. in P. Embrechts (ed.) Extremes and Integrated Risk Management. London: RISK books.

McNeil, A. (1998). *Calculating Quantile Risk Measures for Financial Return Series Using Extreme Value Theory*. Department Mathematik, ETH, Zentrum, Zurich.

McNeil, A., and R. FREY. (1998). *Estimation of Tail-Related Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach*.Department Mathematik, ETH, Zentrum, Zurich.

Mendes, B. (2000). Computing Robust Risk Measures in Emerging Markets Using Extreme Value Theory. *Emerging Markets Quarterly, 4* (2), 25-42.

Reiss, R., and M. Thomas. (1997). *Statistical Analysis of Extreme Values*. Birkhauser-Verlag.