Introduction to the **extreme events** functionality

Ajay Shah, Vimal Balasubramaniam and Vikram Bahure

May 6, 2013

Abstract

The *eventstudies* package includes an extreme events functionality. This package has *ees* function which does extreme event analysis by fusing the consecutive extreme events in a single event. The methods and functions are elucidated by employing data-set of S&P 500 and Nifty.

1 Introduction

Using this function, one can understand the distribution and run length of the clustered events, quantile values for the extreme events and yearly distribution of the extreme events. In the sections below we replicate the analysis for S&P 500 from the Patnaik, Shah and Singh (2013) and we generate the extreme event study plot for event on S&P 500 and response of NIFTY. A detail methodology is also discussed in the paper.

2 Extreme event analysis

This function needs input in returns format on which extreme event analysis is to be done. Further, we define tail events for given probability value. For instance, if *prob.value* is 5 then both side 5% tail events are considered as extreme, lower tail and upper tail (5% to 95%).

```
> library(eventstudies)
> data(eesData)
> input <- eesData$sp500
> # Suppress messages
>   deprintize<-function(f){
+     return(function(...) {capture.output(w<-f(...));return(w);});
+   }
> output <- deprintize(ees)(input, prob.value=5)
```

The output is a list and consists of summary statistics for complete data-set, extreme event analysis for lower tail and extreme event analysis for upper tail. Further, these lower tail and upper tail list objects consists of 5 more list objects with following output:

1. Extreme events dataset

2. Distribution of clustered and unclustered

3. Run length distribution

4. Quantile values of extreme events

5. Yearly distribution of extreme events

2.1 Summary statistics

Here we have data summary for the complete data-set which shows minimum, 5%, 25%, median, mean, 75%, 95%, maximum, standard deviation (sd), inter-quartile range (IQR) and number of observations. The output shown below matches with the fourth column in Table 1 of the paper.

```
> output$data.summary

        summary
Min      -9.47
5%       -2.44
25%      -0.68
Median    0.09
Mean      0.00
75%       0.72
95%       2.25
Max      12.40
sd        1.57
IQR       1.40
Obs.   2035.00
```

2.2 Extreme events dataset

The output for upper tail and lower tail are in the same format as mentioned above. The data-set is a time series object which has 2 columns. The first column is *event.series* column which has returns for extreme events and the second column is *cluster.pattern* which signifies the number of consecutive days in the cluster. Here we show results for the lower tail of S&P 500. Below is the extreme event data set on which analysis is done.

```
> str(output$lower.tail$data)

List of 3
 $ All       :âĂŸzooâĂŹ series from 2000-02-23 to 2011-07-12
  Data: num [1:59, 1:2] -2.63 -2.6 -3.05 -2.57 -2.76 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : chr [1:59] "2000-02-23" "2000-03-08" "2000-05-09" "2000-05-23" ...
  .. ..$ : chr [1:2] "event.series" "cluster.pattern"
  Index:  Date[1:59], format: "2000-02-23" "2000-03-08" ...
 $ Clustered :âĂŸzooâĂŹ series from 2001-03-12 to 2009-01-09
  Data: num [1:3, 1:2] -4.41 -3.46 -4.44 2 2 ...
  ..- attr(*, "dimnames")=List of 2
```

```
 .. ..$ : chr [1:3] "2001-03-12" "2002-07-10" "2009-01-09"
 .. ..$ : chr [1:2] "event.series" "cluster.pattern"
 Index:  Date[1:3], format: "2001-03-12" "2002-07-10" ...
$ Unclustered:âĂŸzooâĂŹ series from 2000-02-23 to 2011-07-12
 Data: num [1:56, 1:2] -2.63 -2.6 -3.05 -2.57 -2.76 ...
 ..- attr(*, "dimnames")=List of 2
 .. ..$ : chr [1:56] "2000-02-23" "2000-03-08" "2000-05-09" "2000-05-23" ...
 .. ..$ : chr [1:2] "event.series" "cluster.pattern"
 Index:  Date[1:56], format: "2000-02-23" "2000-03-08" ...
```

2.3   Distribution of clustered and unclustered events

In the analysis we have clustered, unclustered and mixed clusters. We remove the mixed clusters and study the rest of the clusters by fusing them. Here we show, number of clustered and unclustered data used in the analysis. The *removed.clstr* refers to mixed cluster which are removed and not used in the analysis.*Tot.used* represents total number of extreme events used for the analysis which is sum of *unclstr* (unclustered events) and *used.clstr* (Used clustered events). *Tot* are the total number of extreme events in the data set. The results shown below match with second row in Table 2 of the paper.

```
> output$lower.tail$extreme.event.distribution

      unclstr used.clstr removed.clstr tot.clstr tot tot.used
lower      56          6            40        46 102       62
```

2.4   Run length distribution of clusters

Clusters used in the analysis are defined as consecutive extreme events. Run length shows total number of clusters with n consecutive days. In the example below we have 3 clusters with *two* consecutive events and 0 clusters with *three* consecutive events. The results shown below match with second row in Table 3 of the paper.

```
> output$lower.tail$runlength

                two three
clustered.events   3     0
```

2.5   Extreme event quantile values

Quantile values show 0%, 25%, median, 75%,100% and mean values for the extreme events data. The results shown below match with second row of Table 4 in the paper.

```
> output$lower.tail$quantile.values

                 0%   25% Median   75%  100%  Mean
extreme.events -7.16 -3.94  -3.05 -2.71 -2.44 -3.55
```

2.6 Yearly distribution of extreme events

This table shows the yearly distribution and the median value for extreme events data. The results shown below match with third and forth column for S&P 500 in the Table 5 of the paper.

```
> output$lower.tail$yearly.extreme.event
```

	number.lowertail	median.lowertail
2000	7	-2.634046
2001	5	-4.414078
2002	12	-3.106038
2003	2	-3.795204
2004	1	-2.441617
2007	6	-2.953726
2008	11	-2.980506
2009	9	-3.403248
2010	5	-3.163582
2011	1	-2.524164

The yearly distribution for extreme events include unclustered event and clustered events which are fused. While in extreme event distribution of clustered and unclustered event, the clustered events are defined as total events in a cluster. For example, if there is a clustered event with three consecutive extreme events then yearly distribution will treat it as one single event. Here below the relationship between the Tables is explained through equations:

Sum of yearly distribution for lower tail = 59
Unclustered events for lower tail = 56

Clustered events for lower tail = 3 + 0
Total events in clusters (Adding number of events in each cluster) = 3*2 + 0*3 = 6
Total used events = Unclustered events for lower tail + Total events in clusters
= 56 + 6 = 62

Sum of yearly distribution for lower tail = Unclustered events for lower tail + Total events in clusters
= 56 + 3 =59

```
> sum(output$lower.tail$yearly.extreme.event[,"number.lowertail"])

[1] 59

> output$lower.tail$extreme.event.distribution[,"unclstr"]

[1] 56

> output$lower.tail$runlength
```

	two	three
clustered.events	3	0

3 Extreme event study plot

Here, we replicate the Figure 7, from the paper Patnaik, Shah and Singh (2013). First, we need to have a merged time series object with event series and response series with no missing values for unerring results. After getting the time series object we just need to use the following function and fill the relevant arguments to generate the extreme event study plot.

The function generates extreme values for the event series with the given probability value. Once the values are generated, clustered extreme events are fused together for the response series and extreme evenstudy plot is generated for very bad and very good events. The detail methodology is mentioned in the paper.

```
> eesPlot(z=eesData, response.series.name="nifty", event.series.name="sp500",
+         titlestring="S&P500", ylab="(Cum.) change in NIFTY", prob.value=5,
+         width=5)

Pattern of: 6 ; Discarded event: 1
Pattern of: 5 ; Discarded event: 1
Pattern of: 4 ; Discarded event: 5
Pattern of: 3 ; Discarded event: 7
Pattern of: 2 ; Discarded event: 10
Maximum length after removing mixed clusters is 3
[1] "Clustering events."
```

4 Computational details

The package code is purely written in R. It has dependencies to zoo (Zeileis 2012) and boot (Ripley 2013). R itself as well as these packages can be obtained from CRAN.

Figure 1: Extreme event on S&P500 and response of NIFTY