Estimation of the threshold using a proper scoring rule

May 22, 2014

1   Introduction

When using the generalized Pareto distribution, the choice of threshold is very important. Typically, the selection of threshold is done through the use of diagnostic plots. The goal of this paper is to explore the use of proper scoring rules in order to select an ideal threshold for GPD analysis.

2   Decision framework

In extremes analysis, the role of the threshold is very important. If the selected threshold is too low, it can result in high bias when estimating extreme return levels. If the selected threshold is too high, it can result in high variance for the estimated return levels. Furthermore if a predictive distribution is reasonable for the data, the τth quantile of the predictive distribution, $q(\tau)$, should be reasonably close to τth quantile in the observed data, $q_n(\tau)$. To select an optimal threshold, we propose two scoring functions. The first is a modification to continuous ranked probability score (CRPS) (Gneiting and Raftery, 2007). The second is an integrated quantile score. In the traditional framework scoring functions should be maximized, but we use the negation which treats these as loss functions.

2.1   Continuous ranked probability score

The CRPS (Gneiting and Raftery, 2007) is given by

$$CRPS = -\int_{-\infty}^{\infty} [F(q) - I(x \leq q(\tau))]^2 \, \mathrm{d}q. \tag{1}$$

We propose

$$CRPS_m = \int_{-T^*}^{\infty} [F(q) - I(x \leq q(\tau))]^2 \, \mathrm{d}q. \tag{2}$$

2.2   Integrated quantile score

The Brier score does not account for the closeness of a predicted quantile when assigning weight. The quantile score (Gneiting and Raftery, 2007) is a modification of the Brier score that replaces one of the quadratic terms by the distance between the predicted quantile and the observed value. The quantile score is given by

$$S(q, x) = [I(x \leq q(\tau)) - \tau](x - q(\tau)). \tag{3}$$

We propose

$$iQS = \int_{\tau^*}^{1} (q(\tau) - x)[I(x \leq q(\tau)) - \tau]\mathrm{d}\tau \tag{4}$$

When considering the quantile function of the GPD, iQS is only finite when $\xi \leq 1$.

3 Computation

4 Simulation study

5 Conclusions

Acknowledgments

References

Gneiting, T. and Raftery, A. E. (2007) Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, **102**, 359–378.