Tolerance regions for record values

ALEXANDRE M. BERRED
Université du Havre, France

Having observed a sequence of record values in a sample of size n or only a number of record values, what should be the magnitude of future record values ? This problem has been dealt with, under different settings, in Ahsanullah (1993, 1998), Balakrishnan and Chan (1994), Dunsmore (1983), Hill (1992) and Smith and Miller (1986) among others

Since we want to make inference on future record values, we have to make some minimal assumptions on the tail of the distribution of the random variable generating these record values. Let $\{X_n : n \geq 1\}$ be an i.i.d. sequence of random variables having a continuous distribution function $F(x) = \mathrm{P}(X_1 \leq x)$ with regularly varying upper tail. Namely, for $a \geq 0$,

$$1 - F(x) = x^{-1/\alpha} L(x), \text{ for } x \geq a,$$

where $\alpha > 0$ and L is a slowly varying function at infinity. Denote by $X_{1:n} \leq \cdots \leq X_{n:n}$ the order statistics associated to the sample $X_1, \ldots, X_n$. Define the sequences of the *record times* and *values* by

$$\begin{aligned}
&\tau(1) = 1, \quad \tau(i+1) = \min\left\{ j > \tau(i),\ X_j > X_{\tau(i)} \right\}, \\
&X(i) = X_{\tau(i)}, \quad i \geq 1.
\end{aligned}$$

Denote by $N(n)$ the number of record values within the sample $X_1, \ldots, X_n$, for $n \geq 1$.

We base our estimation on

$$\gamma_{m,n} = \frac{\log X\,(N(n) + m) - \log X\,(N(n))}{\hat{\alpha}},$$

where $m \geq 1$, $r \geq 1$ and $\hat{\alpha}$ is an estimator of α. The estimator $\hat{\alpha}$ will be chosen according to the following. If we deal with series of observations containing only record values, we use the estimator (Berred (1992))

$$\alpha_{k,r}^{(1)} = \frac{1}{p} \sum_{i=1}^{k} \log X(r - i + 1),$$

where $1 \leq k < r$ and $p = kr - k(k-1)/2$. The estimator $\alpha_{k,r}^{(1)}$ is investigated in Berred (1992), Feuerverger and Hall (1994). Otherwise we use the estimator (Hill (1975))

$$\alpha_{k,n}^{(2)} = \frac{1}{k} \sum_{i=1}^{k} \log X_{n-i+1:n} - \log X_{n-k:n},$$

for $1 \leq k < n$.

The asymptotic properties of $\gamma_{m,n}$ will be discussed. These results will be used to construct tolerance regions for future record values.

Acknowledgements

This work was supported in part by CREST, INSEE, Malakoff and URA 1321, Paris, France.

References

AHSANULLAH, M. (1993). Linear prediction of record values for the two parameter exponential distribution. *Ann. Inst. Stat. Math.* **32**, 363–368.

AHSANULLAH, M. (1988). *Introduction to Record Values*. Ginn Press, Needham Heights, Massachusetts.

BALAKRISHNAN, N. AND CHAN, P. S. (1994). Record values from Rayleigh and Weibull distributions and associated inference. *Proceedings of the Conference on Extreme Value Theory and its Applications* (NIST, Gaithersburg, Maryland, USA, May 1993, J. Galambos, J. Lechner, E. Simiu Eds) **3**, 41–51.

BERRED, A. M. (1992). On record values and the exponent of a distribution with a regularly varying upper tail. *J. Appl. Probab.* **29**, 575–586.

BERRED, A. M. (1998). Prediction of record values. To appear in *Communications in Statistics: Theory and Methods* **27**, no. 9.

FEUERVERGER, A. AND HALL, P. (1994). On statistical inference based on record values. To appear in *Extreme*.

DUNSMORE, I. R. (1983). The future occurrence of records. *Ann. Inst. Statist. Math.* **35**, 267–277.

HILL, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163–1174.

HILL, B. M. (1994). Bayesian Forecasting of record values in an exchangeable sequence. *J. Res. Natl. Inst. Stand. Technol.* **99**, 521–538.

SMITH, R. L. and Miller, J. E (1986). A non-gaussian state space model and application to prediction of records. *J. R. Statist. Soc. B* **48**, 79–88.