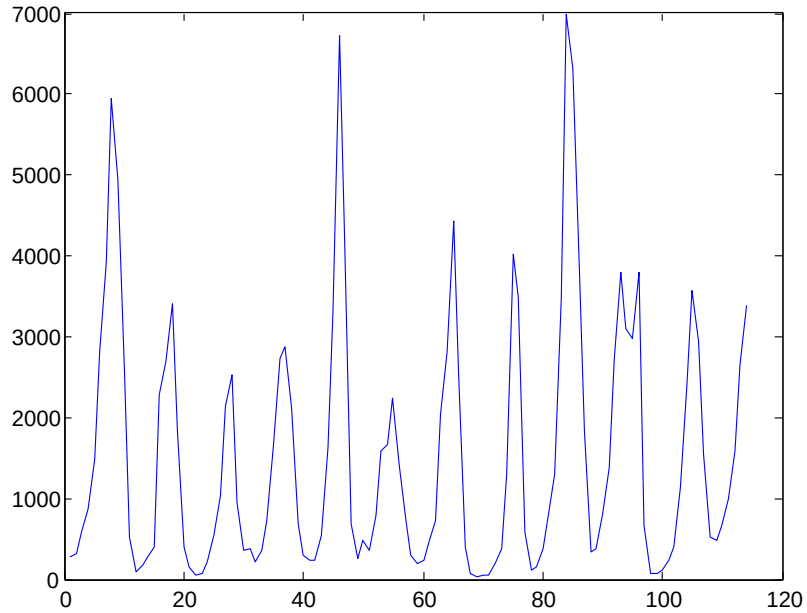
Computing the maximum of random processes and series

Jean-Marc AZAÏS Joint work with : Alan Genz ,Washington State University ; Agnes Lagnoux Toulouse

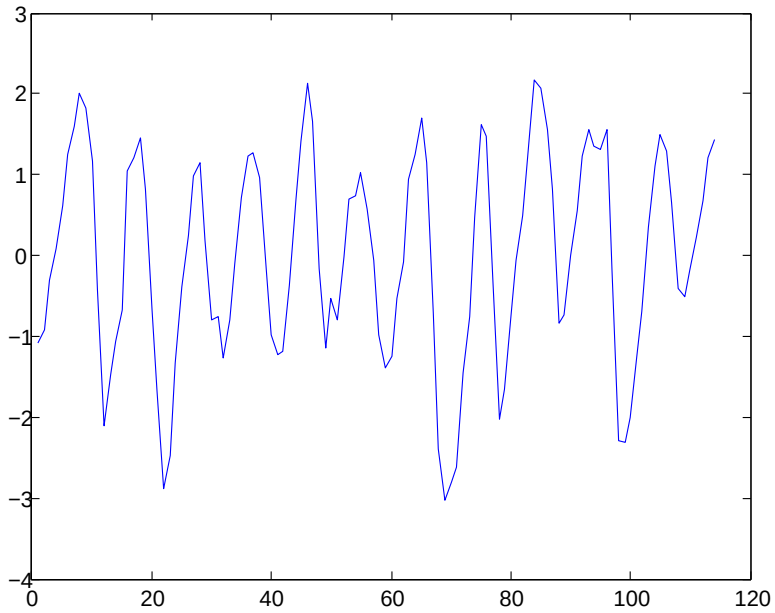
Université de Toulouse
EVA ,Fort collins, june 2009

A toy example : The lynx data



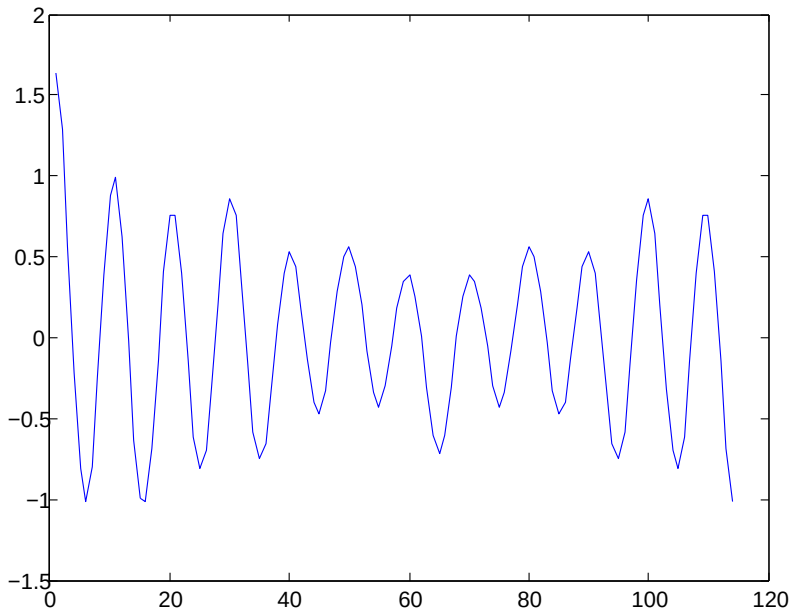
Annual record of the number of the Canadian lynx ”trapped” in the Mackenzie River district of the North-West Canada for the period 1821 - 1934, (Elton and Nicholson, 1942)

After passage to the log and centering



Testing

The maximum of absolute value of the series is 3.0224. An estimation of the covariance with WAFO gives



Can we judge the significativity of 3.0224 ?

Example 2 : Simultaneous bands in curve prediction

A curve is $S(t)$ is expressed onto some basis

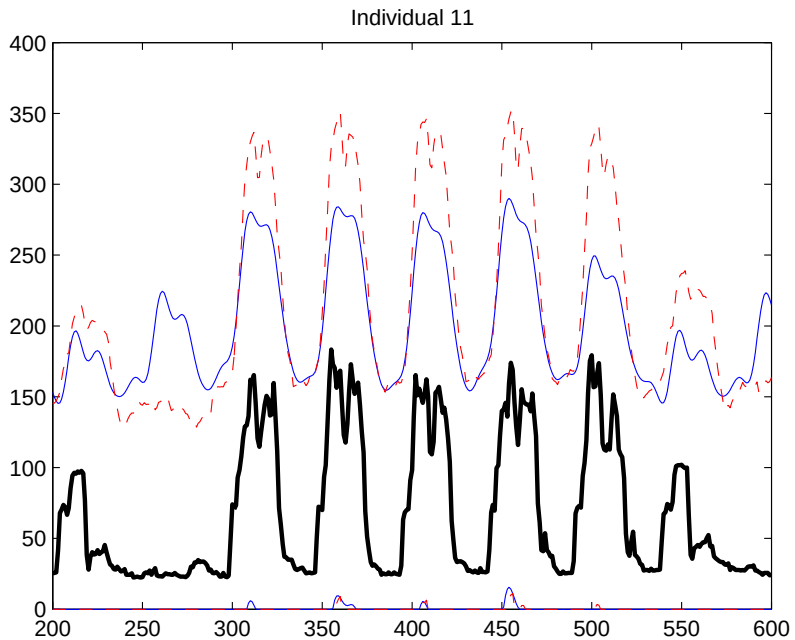
$$S(t) = \sum_{k=1}^{K} a_k \phi_k(t)$$

Each coefficient is estimated by regression on a large learning sample.
The error of regression (variance) can be computed and a **Student-type statistics** can be computed by

$$T(t) := \frac{\widehat{S}(t) - S(t)}{\sqrt{\mathrm{Var}(\widehat{S}(t))}}$$

Confidence bands demand an evaluation of the distribution of

$$M^* := \max_t |T(t)|$$



Typical week load curve for Firm 11

- The Bonferroni method

$$\mathrm{P}\{M^* > u\} < n\mathrm{P}\{|N(0,1)| > u\}$$

- The Sidak (1967) bound

$$\mathrm{P}\{M^* > u\} < \left(\mathrm{P}\{|N(0,1)| > u\}\right)^n$$

- The W of Efron (1997)

$$\mathrm{P}\{M^* > u\} < \mathrm{P}\{|N(0,1)| > u\} + \phi(u) \sum_{i=2}^{n} \frac{\Phi(uL_i/2) - 1/2}{u/2}$$

  Where $L_i := \arccos\left(\mathrm{Cor}(T(i-1), T(i))\right)$
- Numerical exact evaluation by MCQMC method Genz (1992)
- Extreme results

$$\mathrm{P}\left(a_n(M_n^* - b_n) \leq x\right) \to \exp(-\exp(-x)) \quad \text{as} \quad n \to +\infty$$

- The equivalent used in thresholding

$$M_n^* \simeq \sqrt{2\log n}$$

- The Bonferroni method

$$\mathrm{P}\{M^* > u\} < n\mathrm{P}\{|N(0,1)| > u\}$$

- The Sidak (1967) bound

$$\mathrm{P}\{M^* > u\} < \left(\mathrm{P}\{|N(0,1)| > u\}\right)^n$$

- The W of Efron (1997)

$$\mathrm{P}\{M^* > u\} < \mathrm{P}\{|N(0,1)| > u\} + \phi(u)\sum_{i=2}^{n}\frac{\Phi(uL_i/2) - 1/2}{u/2}$$

  Where $L_i := \arccos\left(\mathrm{Cor}(T(i-1), T(i))\right)$
- Numerical exact evaluation by MCQMC method Genz (1992)
- Extreme results

$$\mathrm{P}\left(a_n(M_n^* - b_n) \le x\right) \to \exp(-\exp(-x)) \quad \text{as} \quad n \to +\infty$$

- The equivalent used in thresholding

$$M_n^* \simeq \sqrt{2\log n}$$

- The Bonferroni method

$$\mathrm{P}\{M^* > u\} < n\mathrm{P}\{|N(0,1)| > u\}$$

- The Sidak (1967) bound

$$\mathrm{P}\{M^* > u\} < \left(\mathrm{P}\{|N(0,1)| > u\}\right)^n$$

- The W of Efron (1997)

$$\mathrm{P}\{M^* > u\} < \mathrm{P}\{|N(0,1)| > u\} + \phi(u) \sum_{i=2}^{n} \frac{\Phi(uL_i/2) - 1/2}{u/2}$$

Where $L_i := \arccos\left(\mathrm{Cor}(T(i-1), T(i))\right)$

- Numerical exact evaluation by MCQMC method Genz (1992)
- Extreme results

$$\mathrm{P}\left(a_n(M_n^* - b_n) \leq x\right) \to \exp(-\exp(-x)) \quad \text{as} \quad n \to +\infty$$

- The equivalent used in thresholding

$$M_n^* \simeq \sqrt{2\log n}$$

- The Bonferroni method

$$\mathrm{P}\{M^* > u\} < n\mathrm{P}\{|N(0,1)| > u\}$$

- The Sidak (1967) bound

$$\mathrm{P}\{M^* > u\} < \left(\mathrm{P}\{|N(0,1)| > u\}\right)^n$$

- The W of Efron (1997)

$$\mathrm{P}\{M^* > u\} < \mathrm{P}\{|N(0,1)| > u\} + \phi(u) \sum_{i=2}^{n} \frac{\Phi(uL_i/2) - 1/2}{u/2}$$

Where $L_i := \arccos\left(\mathrm{Cor}(T(i-1), T(i))\right)$

- Numerical exact evaluation by MCQMC method Genz (1992)
- Extreme results

$$\mathrm{P}\left(a_n(M_n^* - b_n) \leq x\right) \to \exp(-\exp(-x)) \quad \text{as} \quad n \to +\infty$$

- The equivalent used in thresholding

$$M_n^* \simeq \sqrt{2\log n}$$

- The Bonferroni method

$$\mathrm{P}\{M^* > u\} < n\mathrm{P}\{|N(0,1)| > u\}$$

- The Sidak (1967) bound

$$\mathrm{P}\{M^* > u\} < \left(\mathrm{P}\{|N(0,1)| > u\}\right)^n$$

- The W of Efron (1997)

$$\mathrm{P}\{M^* > u\} < \mathrm{P}\{|N(0,1)| > u\} + \phi(u) \sum_{i=2}^{n} \frac{\Phi(uL_i/2) - 1/2}{u/2}$$

Where $L_i := \arccos\left(\mathrm{Cor}(T(i-1), T(i))\right)$

- Numerical exact evaluation by MCQMC method Genz (1992)
- Extreme results

$$\mathrm{P}\left(a_n(M_n^* - b_n) \leq x\right) \to \exp(-\exp(-x)) \quad \text{as} \quad n \to +\infty$$

- The equivalent used in thresholding

$$M_n^* \simeq \sqrt{2\log n}$$

- The Bonferroni method

$$\mathrm{P}\{M^* > u\} < n\mathrm{P}\{|N(0,1)| > u\}$$

- The Sidak (1967) bound

$$\mathrm{P}\{M^* > u\} < \left(\mathrm{P}\{|N(0,1)| > u\}\right)^n$$

- The W of Efron (1997)

$$\mathrm{P}\{M^* > u\} < \mathrm{P}\{|N(0,1)| > u\} + \phi(u) \sum_{i=2}^{n} \frac{\Phi(uL_i/2) - 1/2}{u/2}$$

Where $L_i := \arccos\left(\mathrm{Cor}(T(i-1), T(i))\right)$

- Numerical exact evaluation by MCQMC method Genz (1992)
- Extreme results

$$\mathrm{P}\left(a_n(M_n^* - b_n) \leq x\right) \to \exp(-\exp(-x)) \quad \text{as} \quad n \to +\infty$$

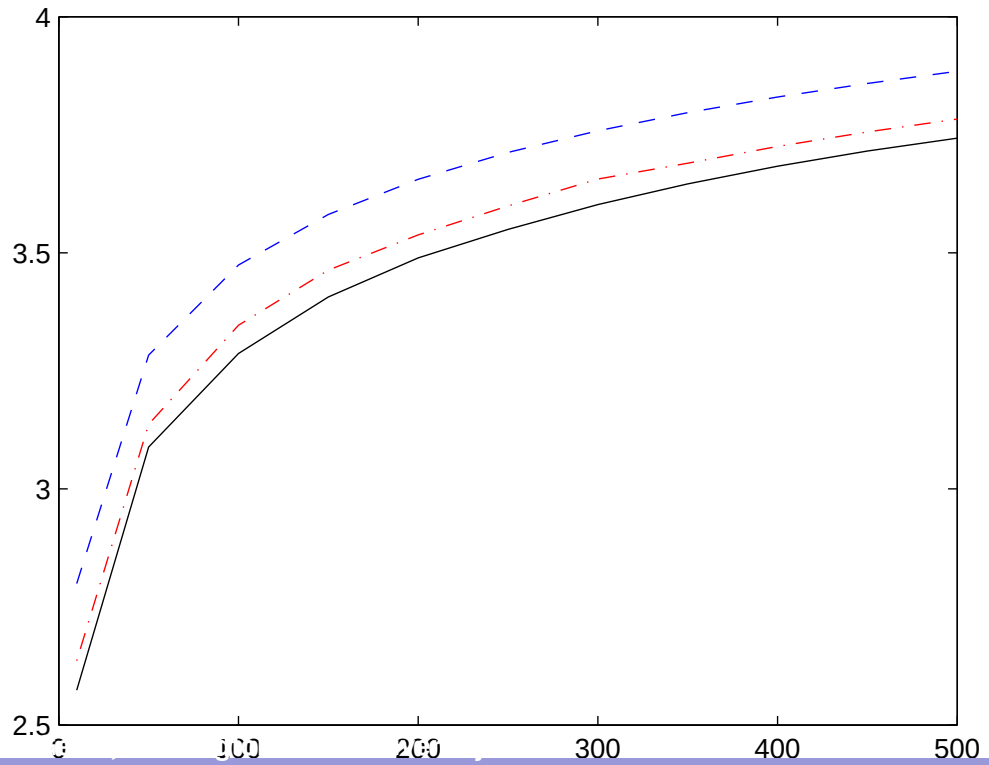
- The equivalent used in thresholding

$$M_n^* \simeq \sqrt{2\log n}$$

Note that the **Sheffé method** cannot be applied in the case of confidence bands

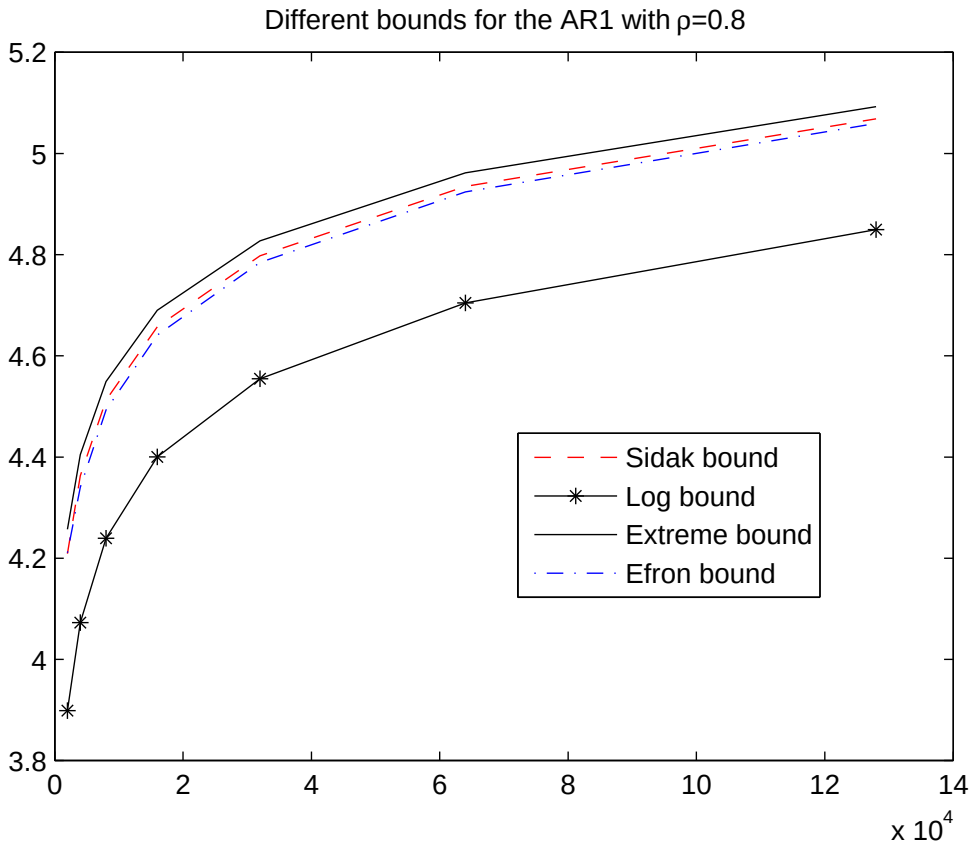
Numerical results, small n

Fractile 0.95 for the AR1 process with $\rho = 0.9$
Sidak W exact



Numerical results, large n

Fractile 0.95 for the AR1 process with $\rho = 0.8$



Conclusion

- $n < 1000$ **MCQMC** is as simple as W and better
- $N > 1000$ **Sidak** or W

Let us consider our problem in a general setting. Σ is a $n \times n$ covariance matrix

$$I := \int_{l_1}^{u_1} \cdots \int_{l_n}^{u_n} \varphi_\Sigma(\mathbf{x}) d\mathbf{x} \tag{1}$$

By conditioning or By Choleski decomposition we can write

$$\begin{aligned} x_1 &= T_{11} z_1 \\ x_2 &= T_{12} z_1 + T_{22} z_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

Where the Z_i's are independent standard. Integral I becomes

$$I := \int_{l_1/T_{11}}^{u_1/T_{11}} \varphi(z_1) dz_1 \int_{\frac{l_2 - T_{12} z_1}{T_{22}}}^{\frac{u_2 - T_{12} z_1}{T_{22}}} \varphi(z_2) dz_2 \cdots\cdots \tag{2}$$

Now making the change of variables $t_i = \Phi(z_i)$

$$I := \int_{\Phi^{-1}(l_1/T_{11})}^{\Phi^{-1}(u_1/T_{11})} dt_1 \int_{\Phi^{-1}\left(\frac{l_2 - T_{12}\Phi^{-1}(t_1)}{T_{22}}\right)}^{\Phi^{-1}\left(\frac{u_2 - T_{12}\Phi^{-1}(t_1)}{T_{22}}\right)} dt_2 \cdots \cdots \quad (3)$$

And by a final scaling this integral can be written as an integral on the hypercube $[0,1]^n$.

$$I := \int_{[0,1]^n} h(\mathbf{t}) d\mathbf{t}. \quad (4)$$

At this stage, if form (4) is evaluated by MC it corresponds to an important reduction of variance $(10^{-2}, 10^{-3})$ with respect to the form (1). The transformation up to there is elementary but efficient.

QMC

In the form (4) the MC evaluation is based on

$$\widehat{I} = 1/M \sum_{i=1}^{M} h(\mathbf{t}_i)$$

it is well known that its convergence is slow : $\mathcal{O}(M^{-1/2})$.
The Quasi Monte Carlo Method is based on the of searching sequences that are “more random than random”. A popular method is based on lattice rules. Let $\mathbf{Z}$ be a “nice integer sequence” in $\mathbb{N}^n$, the rule consist of choosing

$$\mathbf{t}_i = \left\{ \frac{i.\mathbf{z}}{M} \right\},$$

where the notation $\{\}$ means that we have taken the fractional part componentwise. M is chosen prime.

Theorem

(Nuyens and cools, 2006) Assume that h is the tensorial product of periodic functions that belong to a Koborov space (RKHS). Then the minimax sequence and the worst error can be calculated by a polynomial algorithm. Numerical results show that the convergence is roughly $\mathcal{O}(M^{-1})$.

A meta theorem

If h does not satisfies the conditions of the preceding theorem we can still hope QMC to be faster than MC

MCQMC

Let (t_i, i) be the lattice sequence, the way of estimating the integral can be turn to be random but exactly unbiased by setting

$$\widehat{I} = 1/M \sum_{i=1}^{M} h\big(\{\mathbf{t}_i + U\}\big)$$

where U is uniform on $[0,1]^n$.
By the meta theorem $\widehat{I}$ has small variance.

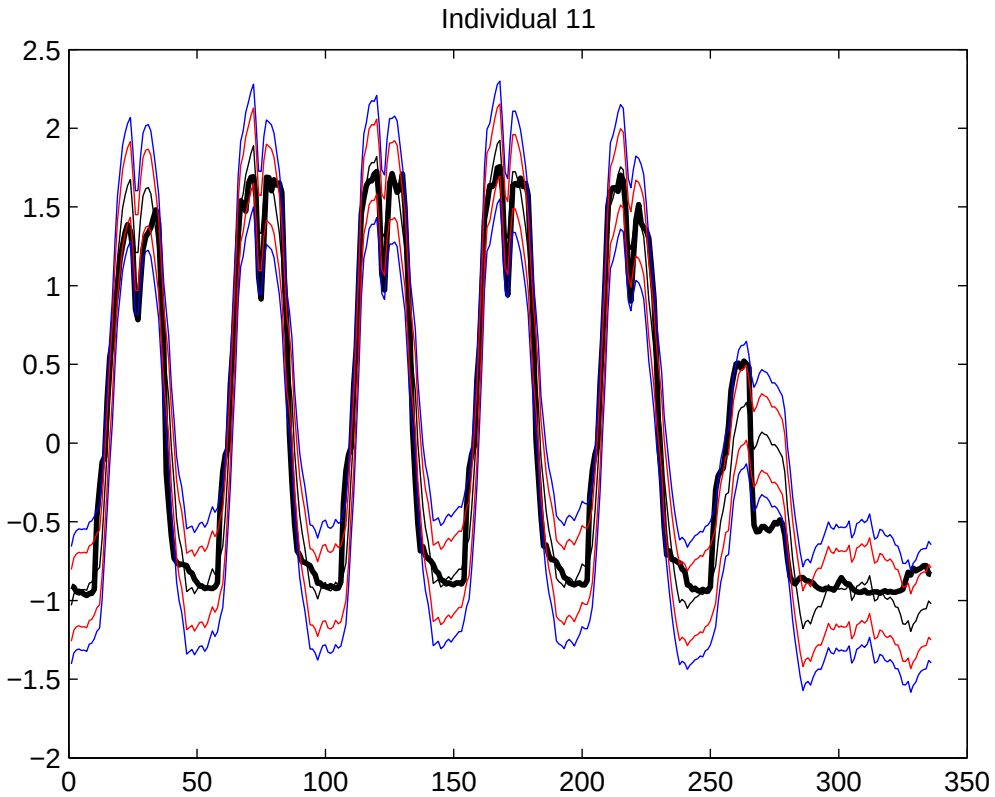
So we can make N independent replications of this calculation and construct Student-type confidence intervals. It is correct whatever the properties of the function h are.
N must be chosen small : in practical 12.

Conclusion : At the cost of a small loss in speed ($\sqrt{12}$) we have a reliable estimation of error.

Conclusion

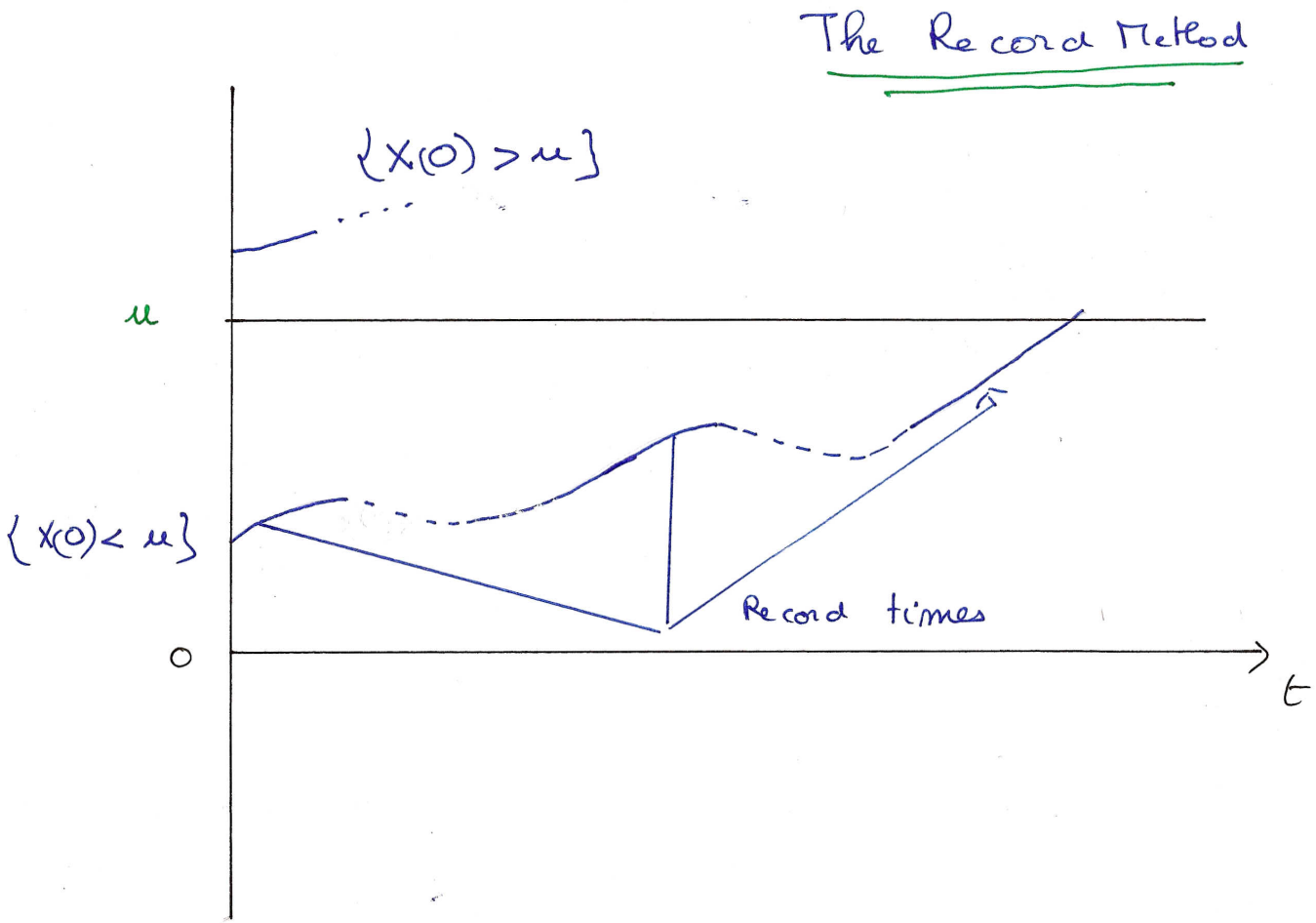
The MCQMC method works with sizes up to 500-1000. This covers most of the cases and it is the safer choice.

Load curve prediction



The record method

$$\mathrm{P}\{M > u\} = \mathrm{P}\{X(0) > u\} + \int_0^T \mathbb{E}\left(X'(t)^+ \mathbb{1}_{X(s)\leq u, \forall s<t} \middle| X(t) = u\right) p_{X(t)}(u) dt \tag{5}$$


The Record Method
{X(0) > u}
u
{X(0) < u}
Record times
0
t

after discretization of $[0,T]$, $D_n = \{0, T/n, 2T/n, \ldots, T\}$ Then

$$\mathrm{P}\{\sup_{t\in D_n} X(t) > u\} \leq \mathrm{P}\{M > u\} \leq \mathrm{P}\{X(0) > u\}$$
$$+ \int_0^T \mathbb{E}\left(X'(t)^+ \mathbb{1}_{X(s)\leq u, \forall s<t, s\in D_n} \middle| X(t) = u\right) p_{X(t)}(u)dt \quad (6)$$

Now the integral is replaced by a trapezoidal rule using the same discretization. Error of the trapezoidal rule is easy to evaluate .

Moreover that the different terms involved can be computed in a recursive way.

An example

Using MGP written by Genz , let us consider the centered stationary Gaussian process with covariance $\exp(-t^2/2)$

[pl, pu, el, eu, en, eq] = MGP(100000, 0.5, 50, @(t)exp(-t.^ 2/2),0,4)

pu upper bound with

eu = estimate for total error,

en = estimate for discretization error, and

eq = estimate for MCQMC error ;

pl lower bound

el = error estimate (MCQMC)

Extensions

Treat all the cases : maximum of the absolute value, non centered, non-stationary. In each case some tricks have to be used.

A great challenge is to use such formulas for fields .

THANK-YOU

References

Azaïs Genz (2009),Computation of the distribution of the maximum of stationary Gaussian sequences and processes. W. paper

Azais Bercu Lagoux Lé. (2009) Simultaneous confidence bands in curve prediction applied to load curves. W. paper

Efron B. (1997).The length heuristic for simultaneous hypothesis tests. Biometrika

Allan Genz web site
http ://www.math.wsu.edu/faculty/genz/homepage

Mercadier, C. (2005). MAGP tooolbox,
http ://math.univ-lyon1.fr/ mercadier/

Mercadier, C. (2006), Numerical Bounds for the Distribution of the Maximum of Some One- and Two-Parameter Gaussian Processes, *Adv. in Appl. Probab.* **38**, pp. 149–170.

Nuyens, D., and Cools, R. (2006), Fast algorithms for component-by-component construction of rank-1 lattice rules