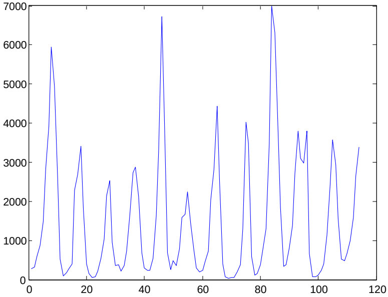
Computing the maximum of random processes and series

Jean-Marc Azaïs Joint work with : Alan Genz ,Washington State University ; Cécile Mercadier Lyon, France and Mario Wschebor

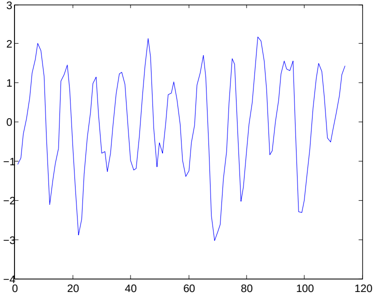
Université de Toulouse

The lynx data



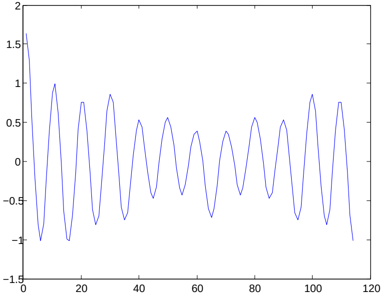
Annual record of the number of the Canadian lynx "trapped" in the Mackenzie River district of the North-West Canada for the period 1821 - 1934, (Elton and Nicholson, 1942)

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 this quantity ?

We assume the series is Gaussian.
Let φ_Σ the Gaussian density in $\mathbb{R}^{114}$. We have to compute

$$\int\int_{-3.0224}^{3.0224} \varphi_\Sigma(x_1,\dots,x_{114})dx_1,\dots,dx_{114}$$

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

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

By conditioning or By Choleski decomposition we can write

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

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

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

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 $\mathbf{Z}_1$ 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 $\mathcal{O}(M^{-1})$.

This result concerns the "worst case" so it is not so relevant

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

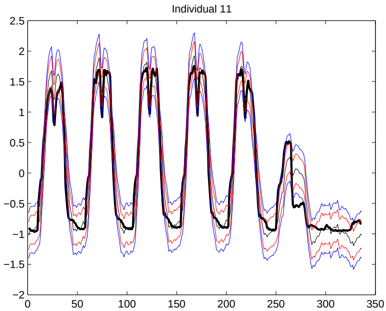
$$\widehat{I} = 1/M \sum_{i=1}^{M} h\big(\{\mathbf{t}_i + U\}\big)$$

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.

This method has been used to construct confidence bands for electrical load curves prediction. Azaïs, Bercu, Fort, Lagnoux Lé (2009)



In this figure the weekly cycle of a firm is decomposed on a Fourier basis. Each coefficient is learned on a **large learning sample** The error is supposed Gaussian and its variance structure can be computed by linear models formulas.

Do processes exist ?

In this part $X(t)$ is a Gaussian process defined on a compact interval $[0, T]$.
Since such a process is always observed in a finite set of times and since the previous method work with say $n = 1000$, is it relevant to consider continuous case ?
Answer yes : random process occur as limit statistics. Consider for example the simple mixture model

$$\begin{cases} H_0 & : & Y \sim N(0,1) \\ H_1 & : & Y \sim pN(0,1) + (1-p)N(\mu,1) \quad p \in [0,1], \quad \mu \in \mathcal{M} \subset \mathbb{R} \end{cases} \tag{5}$$

Theorem (Asymptotic distribution of the LRT)

Under some conditions the LRT of H_0 against H_1 has , under H_0, the distribution of the random variable

$$\frac{1}{2}\sup_{t\in\mathcal{M}}\{Z^2(t)\}, \tag{6}$$

where $Z(.)$ is a centered Gaussian process covariance function

$$r(s,t) = \frac{e^{st}-1}{\sqrt{e^{s^2}-1}\sqrt{e^{t^2}-1}}.$$

In this case there is no discretization.

The record method

$$P\{M > u\} = P\{X(0) > u\} + \int_0^T \mathbb{E}\big(X'(t)^+ \mathbb{I}_{X(s)\le u, \forall s<t} \big| X(t) = u\big) p_{X(t)}(u)dt \quad (7)$$

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

$$P\{\sup_{t\in D_n} X(t) > u\} \le P\{M > u\} \le P\{X(0) > u\} + \int_0^T \mathbb{E}\big(X'(t)^+ \mathbb{I}_{X(s)\le u, \forall s<t, s\in D_n} \big| X(t) = u\big) p_{X(t)}(u)dt \quad (8)$$

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.`$^2/2), 0, 4)$;

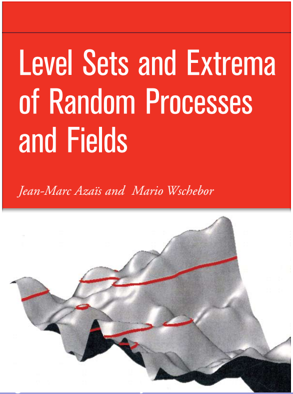
```
pu upper bound with
eu = estimate for total error,
en = estimate for discretization error, and
eq = estimate for MCQMC error;
pl lower bound
el = error estimate (MCQMC)
```

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 .

References


Level Sets and Extrema of Random Processes and Fields
Jean-Marc Azaïs and Mario Wschebor

Azaïs Genz, A. (2009),Computation of the distribution of the maximum of stationary Gaussian sequences and processes. In preparation

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 in shift-invariant reproducing kernel Hilbert spaces, *Math. Comp* **75**, pp. 903–920.

THANK-YOU

MERCI

GRACIAS

THANK-YOU

MERCI

GRACIAS