EPFL, under the supervision of Pr. Stephan Morgenthaler

Master Thesis Project : Extreme Value Theory

– Introductory Task –

Killian Martin–Horgassan

February 18, 2015

1 Summary - tasks assigned

Let us consider a sequence of i.i.d. random variables $X_1, X_2, ... \sim \mathcal{N}(0,1)$[1]. Let us define the maximum at stage n as $M_n = \max(X_1, ..., X_n)$.
Let us consider the following example :

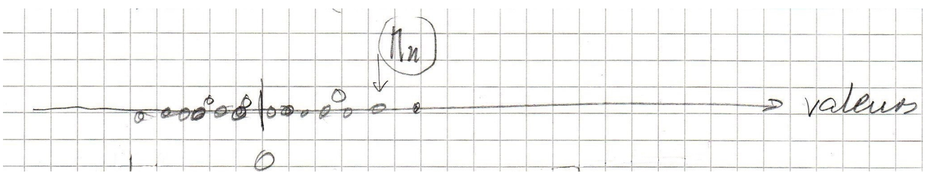
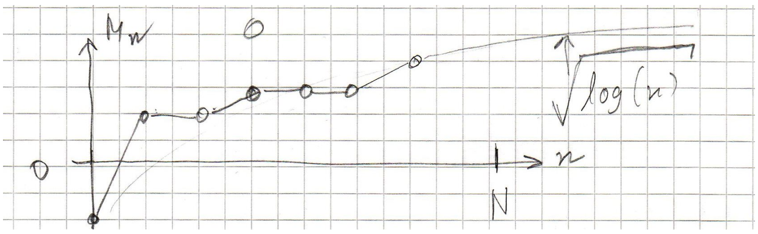


Figure 1.1: Distribution of the $x_1,...,x_n$ around 0.

Let us draw the plotting of M_n against n :



(M_n) is a non-decreasing sequence with growth is of order $\sqrt{\log(n)}$.

[1] could be extended to $\mathcal{N}(\mu, \sigma^2)$

1.1 Sub-task 1

Write a R program that will generate the sequence of the maxima $M_1,...,M_n$ and make the plot of M_n against n. The signature is the following : NormMax(N,m,s,DIST), where :

- N is the number of stages
- m is the mean
- s is the standard deviation
- DIST is the common distribution of the X_i[2]

1.1.1 Sub-task 2

Let us make the following observation, if F_n denotes the cumulated distribution function of M_n then : $F_n(t) = \Pr(\{M_n \leq t\}) = \Pr(\{\max(X_1, ..., X_n) \leq t\}) = \Pr(\{X_1 \leq t\} \cap ... \cap \{X_n \leq t\})$.
By independence of the X_i, it boils down to $F_n(t) = \Phi(t)^n$.

- What is the limit, if any, when n goes to infinity ?
- Is it possible to find two sequences (a_n) and (b_n) such that $\Pr(a_n + b_n M_n \leq t)$ has a limit distribution ?

2 Results

2.1 R programming

Soon to be added...

2.2 Theoretical part

What is the limit, if any, when n goes to infinity ?

$\forall t \in \mathbb{R}$, $0 < \Phi(t) < 1$. Hence, $\forall t \in \mathbb{R}$, $\lim_{n\to\infty} \Phi(t) = 0$. Obviously, there is no convergence in distribution[3] !
Here, the issue is particularly acute as the upper-bound limit of Ω which I hereafter denote by t_+[4] is infinity. If t_+ is finite there is a limiting cumulative distribution function but it is degenerate. That raises an issue as a degenerate distribution is not something it is convenient to work with. It is possible to circumvent that issue by considering a linear renormalization of M_n namely $M_n^* = \frac{M_n - b_n}{a_n}$[5]. That brings us to the next issue : is it possible to find two sequences (a_n) and (b_n) such that (M_n^*) admits a limiting distribution ?

---

[2] gaussian, exponential or Cauchy for instance.

[3] Of course there is no convergence in distribution : the limiting function is not a cdf. Indeed, a cdf must take value 0 at $-\infty$ and value 1 at ∞ and here the limiting function is constant.

[4] that is the smallest t such that $F(t) = 1$

[5] I use here the same notation as Coles in his book.

IS IT POSSIBLE TO FIND TWO SEQUENCES (a_n) AND (b_n) SUCH THAT $\Pr(a_n + b_n M_n \leq t)$ HAS A LIMIT DISTRIBUTION ?

The extremal types theorem[6] provides a related result. It states that **if** there exists $((a_n),(b_n)) \in {\mathbb{R}_+^*}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}$ such that $\lim_{n\to\infty} \Pr(\{\frac{M_n - b_n}{a_n} \leq t\}) = F(t)$ **then** F is non-degenerate and either a Grumbel, a Fréchet or a Weibull distribution. These distributions can be regrouped into the Generalized Extreme Value Distribution (**GEV**).

3 WORK DONE

Please report to file 'JournalDeBord.pdf'[7], section *Week I*.

[6] also referred to as Fisher-Tippett-Gnedenko theorem in the literature.
[7] available on `https://github.com/CillianMH/pdmExtremeValueTheory`