

Annexe A

Comptage de boîtes sur l'Ensemble de Cantor

Dans cette annexe, on étudie le comportement de la méthode du comptage de boîtes lorsqu'elle est appliquée à l'Ensemble de Cantor avec des boîtes de dé- a -uplant (c'est-à-dire des boîtes de tailles $(\frac{1}{a^k})_{k \geq 0}$) pour a différent de 3. Plus précisément, on montre que le nombre de boîtes non vides de taille $\frac{1}{a^k}$ avec $a > 3$ (resp. $a < 3$), est supérieur (resp. inférieur) au nombre de boîtes non vides de taille $\frac{1}{3^k}$.

On appelle Ensemble de Cantor C_N l'ensemble construit par itération jusqu'au niveau limite N à partir d'ensembles C_i de segments de longueur maximale $\frac{1}{3^i}$:

$$\begin{aligned} C_0 &= [0, 1] \\ C_1 &= \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] \\ &\dots \\ C_N &= \lim_{i \rightarrow N} C_i \end{aligned}$$

On définit les ensembles D_i homothétiques de rapport $(3/a)^i$ des ensembles C_i et D_N la poussière limite des D_i :

$$\begin{aligned} \forall i \geq 0 \quad D_i &= Hom_{(3/a)^i}(C_i) \\ D_N &= \lim_{i \rightarrow N} D_i \end{aligned}$$

Soit $N(k)$ le nombre de boîtes non vides de C_N de taille $\frac{1}{3^k}$

$N''(k)$ le nombre de boîtes non vides de C_N de taille $\frac{1}{a^k}$

et $N'(k)$ le nombre de boîtes non vides de D_N de taille $\frac{1}{a^k}$

On a : $\forall k \geq 0 \quad N'(k) = N(k)$.

De plus, comme D_N est inclus dans C_N (en effet les D_i sont inclus dans les C_i car les rapports $(3/a)^i$ des homothéties sont toujours inférieurs à 1), $\forall k \geq 0 \quad N'(k) \leq N''(k)$.
Ainsi :

$$\forall k \geq 0 \quad N(k) \leq N''(k)$$

On montre de la même manière que le nombre de boîtes non vides de taille $\frac{1}{2^k}$ est inférieur au nombre de boîtes non vides de taille $\frac{1}{3^k}$.

Annexe B

Caractérisation des lois de type algébrique

Dans cette annexe, on démontre la caractérisation des lois de type algébrique par le comportement limite de leur *FMDR* (proposition énoncée en section 3.3.2).

Rappelons qu'une loi est de type algébrique de coefficient $q > 1$ si sa fonction de survie au delà de 1 s'écrit :

$$\forall s \geq 1 \quad G(s) = r(s) \cdot s^{-q}$$

avec r vérifiant :

$$\begin{cases} (C1) & r(s) \xrightarrow{s \rightarrow \infty} K \text{ et } r \text{ croissante sur } [1, +\infty[\\ (C2) & \forall s \geq 1 \quad r'(s) \leq As^{-\theta} \text{ avec } A > 0 \text{ et } \theta > 1 \end{cases}$$

Montrons que les conditions (C1) et (C2) sont équivalentes aux conditions :

$$\begin{cases} (C3) & \forall s \geq 1 \quad 0 \leq q - 1 - \frac{1}{h(s)} \leq Cs^{-\nu} \text{ avec } C > 0 \text{ et } \nu > 0 \\ (C4) & \forall s \geq 1 \quad 0 \leq -h'(s) \leq Ds^{-\theta'} \text{ avec } D > 0 \text{ et } \theta' > 1 \end{cases}$$

Supposons que les conditions (C1) et (C2) sont vérifiées. Alors :

$$\int_x^\infty t^{-q} r(t) dt \geq r(1) \cdot \int_x^\infty t^{-q} dt = \frac{r(1)}{q-1} x^{1-q}$$

et

$$\int_x^\infty t^{1-q} r'(t) dt \leq A \cdot \int_x^\infty t^{-q} dt = \frac{A}{q+\theta-2} x^{-q-\theta+2}$$

En intégrant par parties,

$$\int_x^\infty t^{-q} r(t) dt = \frac{1}{q-1} \left\{ r(x) \cdot x^{1-q} + \int_x^\infty t^{1-q} r'(t) dt \right\}$$

et il s'ensuit que :

$$q - 1 - \frac{1}{h(x)} = \frac{\int_x^\infty t^{1-q} r'(t) dt}{\int_x^\infty t^{-q} r(t) dt} \leq C.x^{1-\theta}$$

avec $C = \frac{A.(q-1)}{r(1).(\theta+q-2)}$ ce qui montre que (C3) est vérifiée.

Pour montrer que (C4) est vérifiée, on calcule :

$$-h'(x) = \frac{1}{q-1} \left\{ \frac{r'(x)}{r(x)} + \int_x^\infty t^{1-q} r'(t) dt. \left[\frac{r'(x)}{r^2(x).x^{1-q}} - \frac{q-1}{r(x).x^{2-q}} \right] \right\}$$

et il s'ensuit que $-h'(x) \leq \frac{1}{q-1} \left\{ \frac{A}{r(1)} + \frac{A.x^{1-2\theta}}{(q-1)r(1).x^{1-q}} \right\}$ ce qui montre que (C4) est vérifiée en remarquant que le premier terme est dominant et que r est croissante.

Supposons que les conditions (C1) et (C2) sont vérifiées. Posons H la fonction définie par :

$$H(x) = \int_x^\infty t^{-q} r(t) dt$$

On a :

$$h(x) = \frac{-H(x)}{x.H'(x)}$$

et :

$$\frac{1}{x} \left(q - 1 - \frac{1}{h(x)} \right) = \frac{H'(x)}{H(x)} + \frac{(q-1).x^{q-2}}{x^{q-1}}$$

En intégrant cette relation, on obtient :

$$H(x) = H(1).x^{1-q}. \exp \left\{ \int_1^x \left(q - 1 - \frac{1}{h(t)} \right) \frac{dt}{t} \right\}$$

et ainsi :

$$r(x) = \frac{H(1)}{h(x)} \exp \left\{ \int_1^x \left(q - 1 - \frac{1}{h(t)} \right) \frac{dt}{t} \right\}$$

Donc r est continue, croissante car $\frac{1}{h}$ est croissante et tend vers $q-1$. La condition (C3) implique que l'intégrale converge quand x tend vers l'infini, donc r tend vers une limite finie K et la condition (C1) est satisfaite.

Pour montrer que la condition (C2) est satisfaite, constatons que :

$$\frac{r'(x)}{r(x)} = \frac{1}{x} \left(q - 1 - \frac{1}{h(x)} \right) - \frac{h'(x)}{h(x)}$$

La condition (C3) implique que le premier terme est majoré par $C.x^{-\theta}$. De plus, elle implique que $\frac{1}{h}$ est majorée par $q - 1$ et la condition (C4) entraîne que le second terme est majoré par $D.(q - 1).x^{-q}$, ce qui démontre que (C2) est vérifiée.

Annexe C

Calculs du biais et de la variance de \widehat{h}

Dans cette annexe, on calcule formellement, et dans le cas général, le biais et la variance de l'estimateur \widehat{h} présenté en section 3.4.3.

C.1 Biais

Soit s un réel supérieur à 1. On a :

$$\widehat{h}(s) = \frac{1}{s} \sum_{i=1}^N (X_i - s) \cdot 1_{X_i > s} \times \frac{1}{1 + \sum_{j \neq i} 1_{X_j > s}}$$

et comme les $(X_i)_{i=1 \dots N}$ sont indépendantes :

$$E \left[\widehat{h}(s) \right] = H(s) \cdot E \left[\frac{1}{1 + \sum_{j \neq 1} 1_{X_j > s}} \right]$$

Or $\sum_{j \neq 1} 1_{X_j > s} \sim \text{Bin}(N-1, G(s))$, donc :

$$E \left[\widehat{h}(s) \right] = H(s) \cdot \sum_{k=1}^N \frac{1}{1+k} C_{N-1}^k G^k(s) F^{N-1-k}(s) = \frac{H(s)}{sG(s)} \left[(F(s) + G(s))^N - F^N(s) \right]$$

C.2 Variance

Soit s un réel supérieur à 1. On a :

$$\widehat{h}^2(s) = \frac{1}{s^2} \sum_{i=1}^N \frac{(X_i - s) \cdot 1_{X_i > s}}{\sum_j 1_{X_j > s}} = \frac{1}{s^2} (S_1 + S_2)$$

avec

$$S_1 = \sum_{i=1}^N \frac{(X_i - s)^2 \cdot 1_{X_i > s}}{\left(1 + \sum_{j \neq i} 1_{X_j > s}\right)^2} \text{ et } S_2 = \sum_{i=1}^N \sum_{j \neq i} \frac{(X_i - s) \cdot (X_j - s) \cdot 1_{X_i > s} \cdot 1_{X_j > s}}{\left(2 + \sum_{k \neq i, j} 1_{X_k > s}\right)^2}$$

et comme les $(X_i)_{i=1 \dots N}$ sont indépendantes et comme $\sum_{k \neq 1, 2} 1_{X_k > s} \sim \text{Bin}(N-2, G(s))$, on

a :

$$E[S_2] = N(N-1)H^2(s) \cdot \sum_{k=0}^{N-2} \frac{1}{(2+k)^2} C_{N-2}^k G^k(s) F^{N-2-k}(s)$$

Or :

$$\sum_{k=0}^{N-2} \frac{1}{(2+k)^2} C_{N-2}^k G^k(s) F^{N-2-k}(s) = \frac{1}{N(N-1)G^2(s)} [1 - G^N(s) - S_N(s)]$$

avec $S_N(s) = \sum_{k=1}^N \frac{1}{k} C_N^k G^k(s) F^{N-k}(s)$. Ainsi :

$$E[S_2] = \frac{H^2(s)}{G^2(s)} \cdot [1 - F^N(s) - S_N(s)]$$

En intégrant deux fois par parties :

$$\forall i \geq 1 \quad E[(X_i - s)^2 \cdot 1_{X_i > s}] = 2 \int_s^\infty H$$

et $E[S_1] = 2N \cdot \left(\int_s^\infty H\right) \cdot \sum_{k=1}^N \frac{1}{(1+k)^2} C_{N-1}^k G^k(s) F^{N-1-k}(s)$. Comme :

$$\sum_{k=1}^N \frac{1}{(1+k)^2} C_{N-1}^k G^k(s) F^{N-1-k}(s) = \frac{S_N(s)}{NG(s)}$$

le résultat découle directement puisque :

$$E[\widehat{h}^2(s)] = \frac{2S_N(s)}{s^2 G(s)} \int_s^\infty H + \frac{H^2(s)}{s^2 G^2(s)} \cdot [1 - F^N(s) - S_N(s)]$$

Annexe D

Auto-similarité et Multifractalité

On rencontre un grand nombre de définitions non-équivalentes de l'auto-similarité dans la littérature. Parmi elles, la plus courante est celle selon laquelle un processus continu $\{Y(t), t \in T\}$ est auto-similaire (de paramètre H) si et seulement si :

$$Y(t) \stackrel{L}{=} a^{-H} Y(at) \quad \forall t \in T, \quad \forall a > 0, \quad 0 \leq H < 1 \quad (1)$$

Remarque 37 *Ce processus ne peut être stationnaire. On suppose néanmoins que ses incréments le sont.*

Une deuxième définition de l'auto-similarité, plus appropriée dans ce contexte de séries chronologiques classiques, concerne un processus stationnaire $\{X(i), i \geq 1\}$. On définit le processus agrégé à pas de temps m par :

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X(i) \quad \forall k \geq 1 \quad (2)$$

Si X est le processus d'incrément du processus Y défini précédemment (ie $X(i) = Y(i+1) - Y(i)$), alors pour tout $m \geq 1$:

$$X(t) \stackrel{L}{=} m^{1-H} X^{(m)} \quad (3)$$

Définition 38 *Une série stationnaire $\{X(i), i \geq 1\}$ est dite **exactement auto-similaire** si et seulement si elle vérifie (3).*

*Elle est dite **asymptotiquement auto-similaire** si et seulement si (3) est vrai quand $m \rightarrow \infty$.*

*Elle est dite **auto-similaire du second ordre** si et seulement si X et $m^{1-H} X^{(m)}$ ont même variance et auto-corrélation et **asymptotiquement auto-similaire du second ordre** si et seulement si la variance et l'auto-corrélation de $m^{1-H} X^{(m)}$ tend vers celle de X quand $m \rightarrow \infty$.*

Remarque 39 *Si le processus (X) est gaussien, les définitions sont équivalentes.*

Définition 40 *Pour une série stationnaire $\{X(i), i \geq 1\}$, on définit les moments absolus d'ordre q par :*

$$\mu^{(m)}(q) = E \left| X^{(m)} \right|^q = E \left| \frac{1}{m} \sum_{i=1}^m X(i) \right|^q$$

Définition 41 *Troisième définition de l'auto-similarité (impliquée mais pas équivalente à la définition (3)) :*

$$\log \mu^{(m)}(q) = \beta(q) \log m + C(q) \text{ et } \beta(q) = q(H - 1)$$

Définition 42 *Processus multifractal : Un processus positif est dit multifractal si et seulement si la relation log log entre ses moments d'ordre q et son niveau d'agrégation m est linéaire :*

$$\log \mu^{(m)}(q) = \beta(q) \log m + C(q)$$

Définition 43 *Processus multifractal signé : Un processus est dit multifractal signé si et seulement si la relation log log entre ses moments d'ordre q et son niveau d'agrégation m est linéaire :*

$$\log \mu^{(m)}(q) = \beta(q) \log m + C(q)$$

Remarque 44 *Un processus multifractal signé est une généralisation d'un processus auto-similaire.*

Remarque 45 *Un processus stationnaire ne peut être auto-similaire ou asymptotiquement auto-similaire s'il n'est pas centré. En effet, par (3), $EX = m^{1-H} EX^{(m)} = m^{1-H} EX$ (ou $m^{1-H} EX \xrightarrow{m \rightarrow \infty} EX$). Cependant, il peut être multifractal.*

Annexe E

Liste des stations

Dans cette annexe, on fournit le détail des séries de plus de 100 ans utilisée dans la *section 3.5.4*.

Figure E.1: Liste des stations des séries annuelles.