

# **LOI DE ZIPF.**

## Sommaire

.....	1
.....	2
<b>Introduction</b> .....	4
<b>Chapitre 1 :Lois Puissance</b> .....	6
1. Définition.....	6
2 Loi de Pareto.....	6
3. Loi de Zipf.....	9
4. Loi log-normale.....	10
5. Statistique d'ordre.....	11
<b>Chapitre 2: APPLICATION DE LOI DE ZIPF</b> .....	13
1- Distribution de la Population des villes.....	13
2-Modèle de Base ( loi de Gibrat).....	15
3- Modèle canonique de Gabaix ( économique):.....	18
4-Modèles de Croissance aléatoire et effets de taille.....	20
5-Modèle Dynamique urbaine (nouvelles villes ).....	26
6-Modèle Rang-Taille :.....	30
7-Le modèle de Simon et Steindl.....	33
8-Système urbaine former par des régions hétérogènes.....	35
9-Dérivation alternative du mécanisme de croissance avec les processus Kesten.....	36
<b>Chapitre 3 :ESTIMATION</b> .....	39
1-Modèle log rang taille.....	39
2-modèle de Lotka.....	39
3-L'estimateur de Hill.....	40
4-Méthode des moindres carrés généralisée.....	41
<b>Conclusion</b> .....	45



## Introduction

Dans les années 1930, c'est-à-dire bien avant l'apparition des ordinateurs, un principe mathématique appelé la loi de Zipf a prédit la taille des méga-villes dans le monde entier. En 1949, le linguiste George Zipf a remarqué quelque chose d'étrange quant à la fréquence à laquelle les gens utilisent les mots dans une langue donnée. Il a constaté qu'un petit nombre de mots sont utilisés tout le temps, alors que la grande majorité est utilisée très rarement. Il a remarqué en classant les mots dans l'ordre de popularité que Le mot classé numéro un a toujours été utilisé deux fois plus souvent que le mot de deuxième rang et trois fois plus souvent que le troisième rang. Il a appelé cela une règle par rapport à la fréquence et a constaté qu'il pourrait également être utilisé pour décrire les distributions de revenus dans un pays donné, la personne la plus riche gagnant deux fois plus d'argent que la plus proche, et ainsi de suite.

Plus tard, la loi de Zipf, la règle de fréquence par rapport au rang fonctionne également si vous l'appliquez à la taille des villes. La ville avec la plus grande population dans n'importe quel pays est généralement deux fois plus grande que la plus grande, et ainsi de suite. La loi de Zipf pour les villes est restée vraie pour tous les pays du monde depuis un siècle. Cette loi permet de résumer en une seule valeur, celle du coefficient de hiérarchisation, les informations qui régissent la formation et l'évolution des hiérarchies urbaines d'un pays ou d'une région.

De nombreux chercheurs, économistes, statisticiens ou géographes, se sont focalisés, ces dernières années, sur le sujet des hiérarchies urbaines et de leur évolution dans le but d'étudier la Distribution des villes selon leur taille.

La ville peut, aujourd'hui, être définie par des critères quantitatifs, tels que la taille de sa population, la surface de son étendue, sa densité ou même le volume de sa production marchande. Elle peut également faire l'objet de définitions fonctionnelles comme son statut administratif, sa spécialisation économique ou son rôle dans la structuration des échanges et des communications.

Notre travail consiste a



# Chapitre 1 :Lois Puissance

## 1. Définition

On peut définir une loi puissance comme étant une fonction associant deux quantités. Elle peut être de la forme suivante  $y = ax^{-b}$  avec a et b des constantes. Une représentation graphique dans un repère (x,y) est la suivante :

En général, les lois puissance sont représentées dans un repère bi-logarithme. En choisissant une tel repère, on obtient que la représentation graphique d'une distribution en loi puissance est une droite. On a donc  $\ln y = \ln a - b \ln x$

Les modèles de loi puissance les plus connus sont la loi de Zipf et la loi de Pareto. On les retrouve dans de nombreux domaines tel que l'économie, la démographie, la musique, l'internet ou encore la génétique.

## 2 Loi de Pareto

La loi de Pareto ou encore la « loi des 80/20 » est l'une des premières loi puissance (1897) mise en place par l'économiste Wilfried Pareto. En étudiant la répartition des revenus personnels des individus dans les principaux pays industrialisés, Pareto a constaté que cette répartition suivait une loi puissance.

Il base ses démonstrations à partir d'une définition simplifiée des variations de l'inégalité dans la distribution. En effet il dit que « la diminution de cette inégalité sera définie par le faite que le nombre des pauvres va en diminuant par rapport au nombre de riches ou ce qui est la même chose par rapport au nombre total des membres de la société ».

Soit X une variable de paramètres k et  $x_{min}$ . On dit que X suit une loi de Pareto notée

$L_p(k, x_{min})$  si la variable statistique X présente la présente la fonction de répartition ayant les caractéristiques suivantes :

$$F(x) = \begin{cases} 0 & \text{si } x < x_{min} \\ 1 - \left(\frac{x_{min}}{x}\right)^k & \text{si } x \geq x_{min} \end{cases}$$

Où F(x) représente la proportion de personnes ayant un revenu inférieur à x.

Si l'on s'intéresse à la probabilité qu'un individu pris au hasard dans la population ait un revenu X supérieur

$$P(X > x) = \left(\frac{x}{x_{min}}\right)^{-k}$$

Avec  $x_{min}$  le revenu minimum et k une constante positive.

L'espérance mathématique du revenu est donnée par la formule suivante :

$$E(X) = x_{min} \frac{k}{k-1}$$

Sa variance est :

$$\text{Var}(x) = \left( \frac{x_{\min}}{k-1} \right)^2 - \frac{k}{k-2}$$

La valeur de  $k$  mesure la pente de la distribution et mesure l'inégalité des revenus.

On observe cette loi dans les domaines tels que les revenus de particuliers mais aussi les chiffres d'affaires des entreprises.

A travers ces études Pareto a souvent souligné le fait que la réduction de la pauvreté qui se traduit généralement par un relèvement du revenu moyen n'implique pas forcément une réduction des inégalités de la distribution. En effet à travers sa loi, il a montré qu'une baisse de la pauvreté se traduit par une augmentation de  $x_{\min}$  et par conséquent un accroissement de  $k$  provoquant ainsi une augmentation de l'inégalité.

Pour Pareto le principal souci n'est pas l'existence d'inégalités sociales mais leur importance. Donc une société peut relativement égalitaire et la répartition des revenus suivre quand même une loi de Pareto. Selon Pareto les 20% des individus les plus riches possèdent 80% de la richesse d'un pays.

Cependant la loi de Pareto ne modélise que les revenus les plus élevés correctement.

En effet durant le vingtième siècle certains facteurs tels que le développement des classes moyennes et l'instauration de minima sociaux ont engendrées des différences entre ces travaux et les études récentes.

#### Simulation d'une loi de Pareto

La densité de la loi de Pareto de paramètre  $k$  sur l'intervalle  $1 \leq x < \infty$  est :

$$f(x) = \begin{cases} k x^{-k-1} & \text{si } x \geq 1 \\ 0 & \text{sinon} \end{cases}$$

On calcule de fonction de répartition  $F(x)$  de cette loi

$$F(x) = P(X \leq x) = \int_1^x k t^{-k-1} dt = [-t^{-k}]_1^x = 1 - x^{-k}$$

On calcule ensuite l'inverse de  $F$  :

On cherche à écrire  $x$  en fonction de  $u$  avec  $u = 1 - x^{-k}$ .

Et donc  $F(u)^{-1} = (1-u)^{-1/k}$

On peut maintenant simuler la loi de Pareto à l'aide de la fonction de répartition inverse. Cela consiste à construire un échantillon uniformément distribué sur l'intervalle  $[0, 1]$

Et lui applique l'inverse de la fonction de répartition.

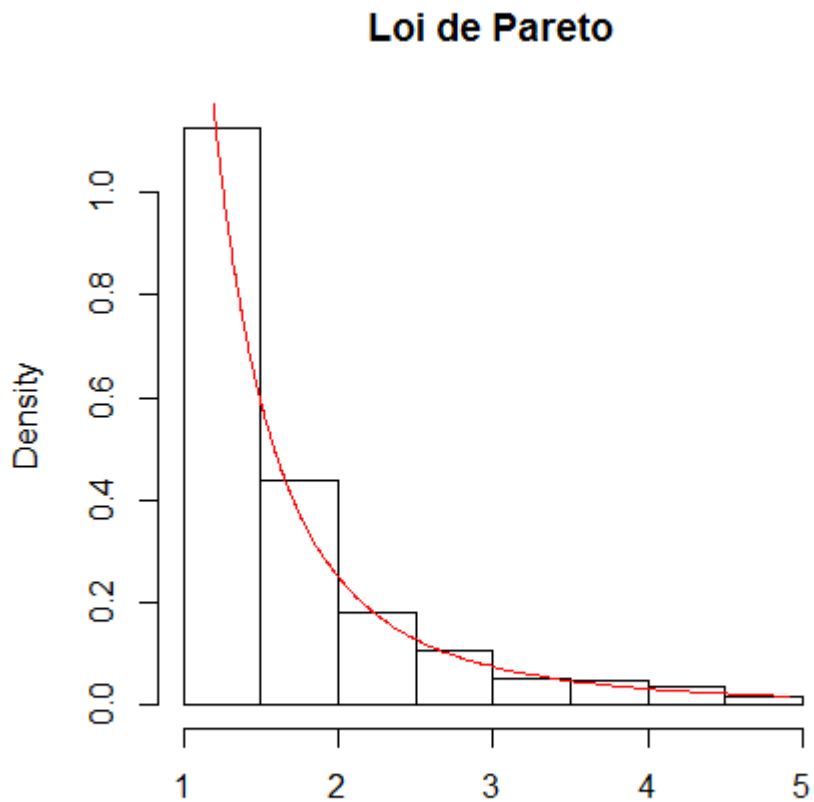
On choisit ici :

- 1)  $n=1000$  (nombre d'observations)
- 2)  $k=2$  (le paramètre)

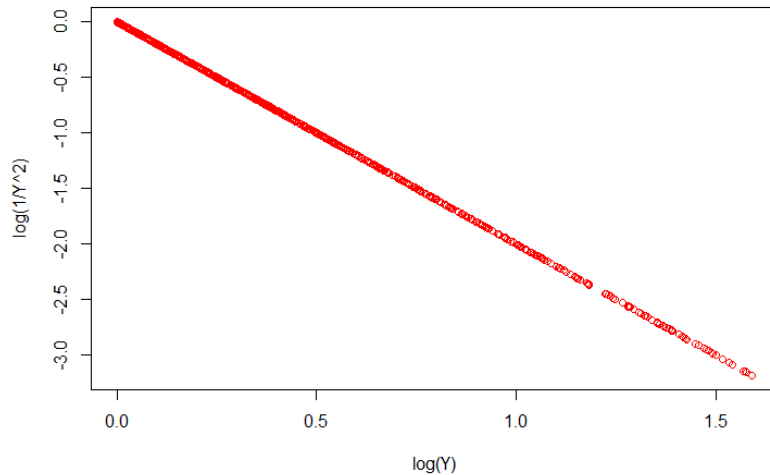
Avec ces lignes de code sur R.studio

```
set.seed(1)
n <- 1000
X <- 1/sqrt(runif(n))
hist(X[X<5],prob=TRUE,main="Loi de Pareto",xlab="")
Y <- sort(X[X<5])
lines(Y,2/Y^3, col="red")
plot(log(Y),log(1/Y^2), col="red")
```

On peut avoir les graphiques suivants :







### 3. Loi de Zipf

#### 1-Loi de Zipf et les mots

La loi de Zips est avant tout une loi empirique, dans ces études il a constaté que la fréquence d'utilisation d'un mot dans un texte était inversement proportionnelle à son rang.

La loi de Zips s'intéresse à la répartition statistique des fréquences d'apparition des différents mots dans un langage. Et à l'inverse de la loi de Pareto qui s'intéresse à des probabilités, on utilise ici directement les résultats du comptage des mots, fréquences absolues.

Si  $x$  est le  $x^e$  mot le plus fréquent dans un texte, il doit avoir alors la fréquence

$$f(x) = \frac{C}{x} \text{ où } C \text{ est une constante.}$$

Selon la loi de Zipf, on peut assimiler les fréquences d'utilisation d'un mot rangé par ordre décroissant à une loi puissance. En d'autres termes, si l'on désigne par  $f$  la fréquence d'apparition du mot de rang  $x$  dans la suite, la relation est de la forme suivante :

$$f = C x^{-a}$$

Avec  $C$  et  $a$  des constantes.

#### 2-Loi de Zipf et les villes

Zips a étendu son travail et cette loi a différents domaines tels que les systèmes de villes.

Selon Zipf, si  $X$  est la variable qui associe à chaque ville sa population,  $X$  suit une loi de Pareto, dont la fonction densité est de la forme

$$f(x) = C x^{-a}$$

Avec  $a$  est un coefficient d hiérarchisation et  $C$  est une constante qui dépend de la taille de la plus grande ville du pays. Si  $P(X > x)$  est la probabilité qu'une ville ait une population supérieure à  $x$ , la fonction de répartition complémentaire associé à  $X$  est de la forme suivante :

$$F(x) = P(X > x)$$

$$\int_x^{\infty} f(t) dt = \int_x^{\infty} C t^{-a} dt$$

$$F(x) = \frac{C x^{1-a}}{a-1}$$

Avec  $a > 1$ .

## 4. Loi log-normale

La loi log-normale également appelée loi de Galton est une loi de probabilité qui sert souvent à modéliser des distributions.

En théorie une variable aléatoire  $X$  suit une loi log-normale de paramètres  $\mu$  et  $\sigma^2$  si la variable  $Y = \ln(X)$  suit une loi normale d'espérance  $\mu$  et de variance  $\sigma^2$ .

Une distribution est modélisée par la loi log-normale quand les effets de nombreux facteurs indépendants se multiplient.

Cette loi est généralement notée Log-N ( $\mu, \sigma^2$ ).

On a donc que si  $X$  suit une loi log-puissance, sa fonction de densité est de la forme suivante :

$$f(x, \mu, \sigma) = \frac{1}{x \sigma \sqrt{2\pi}} \exp\left[-\frac{(\ln(x-\mu))^2}{2\sigma^2}\right]$$

Avec  $\mu$  l'espérance et  $\sigma$  l'écart type.

La fonction de répartition de la variable  $X$ , ou encore fonction de réparation de la loi log-normale est liée à la fonction d'erreur notée ERF.

Par intégration de la fonction de densité on obtient :

$$F(x) = P(X \leq x)$$

$$\begin{aligned} \int_{-\infty}^x f(t, \mu, \sigma^2) dt &= \int_{-\infty}^x \frac{1}{t \sigma \sqrt{2\pi}} \exp\left[-\frac{(\ln(t-\mu))^2}{2\sigma^2}\right] dt \\ &= \frac{1}{2} + \frac{1}{2} \text{ERF}\left[\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right] \end{aligned}$$

L'espérance est la suivant :

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

La variance se présente de la façon suivante :

$$V(X) = (\exp(2\mu + \sigma^2))(\exp\sigma^2 - 1)$$

### *Loi Gibrat*

Historiquement nommée loi de l'effet proportionnel mais également loi log-normale à 3 paramètres, la loi de Gibrat est une loi de probabilité énoncée par l'ingénieur français Robert Gibrat.

Cette loi est une généralisation de la loi log-normale obtenue par l'ajout d'une simple translation en posant  $Y = \ln(X - X_0)$

Elle est notée  $\log N(X_0, \mu, \sigma^2)$  et ne concerne que des valeurs  $X > X_0$ .

En utilisant cette loi comme modèle, le taux de croissance d'une ville serait indépendant de sa taille.

## 5. Statistique d'ordre

Soit  $X_1 < X_2 < \dots < X_n$

Ici

$$X_1 = \min(X_1, \dots, X_n)$$

$$X_n = \max(X_1, \dots, X_n)$$

Et

$$X_k = k \text{ où } k \text{ est la } k\text{-ième plus petite valeur de } X_1, \dots, X_n$$

**Théorème :** Supposons que  $X_1, \dots, X_n$  sont des variables aléatoires I.I.D de densité  $f$ . La densité jointe de la statistique d'ordre  $n$  peut être définie de la façon suivante :

$$f_{X_1, X_2, \dots, X_n}(y_1, \dots, y_n) = \begin{cases} n! \prod_{k=1}^n f(y_k) & \text{si: } y_1 < y_2 < \dots < y_n \\ 0 & \text{sinon} \end{cases}$$

### 1. Variable de commandes exponentielles

Soit  $X_1, \dots, X_n$  des variables aléatoires I.I.D de distribution  $e^\theta$ . On peut donc définir la fonction de densité de chaque  $X_i$  par :

$$f(x, \theta) = \begin{cases} \frac{1}{\theta} & \text{si: } x \geq 0 \\ 0 & \text{si: } x < 0 \end{cases} \quad (\text{#####})$$

Nous nous sommes intéressés par les différences entre les variables d'ordre

$$X_1, X_i - X_{i-1} \quad i=2, \dots, n$$

**Théorème :** Supposons que  $X_1, \dots, X_n$  sont des variables aléatoires I.I.D qui suivent la loi  $\exp(1)$ .

Alors on a :

1.

$$X_1 \sim \exp\left(\frac{1}{n}\right), \quad X_i - X_{i-1} \sim \exp\left(\frac{1}{n+1-i}\right)$$

2.

$X_1, X_i - X_{i-1}$  pour  $i=2, \dots, n$  sont  $n$  variables aléatoires indépendantes.

**Démonstration :**

Nous définissons  $Y_i$  pour  $i=1, \dots, n$  par :

$$Y_1 = X_1, \quad Y_i = X_i - X_{i-1}$$

Alors nous introduisons

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

Si :

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{pmatrix} \quad X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_n \end{pmatrix}$$

On a :  $Y = AX$

Il est évident que la matrice inverse  $A^{-1}$  existe, il est possible de trouver  $X$  à partir de  $Y$  avec la décomposition suivante :

$$Y_1 = X_1, \quad X_i = Y_i + Y_{i-1} + \dots + Y_1$$

On peut donc écrire ces égalités précédentes sous forme de matrice.

On obtient :

$$X = A^{-1}Y$$

En appliquant un changement de variable on obtient :

$$f_Y(y) = f_X(A^{-1}y) \frac{1}{|\det(A)|}$$

Ensuite à l'aide de l'égalité du théorème 1, on peut écrire :

$$f_X(A^{-1}y) = n! f(y_1) f(y_1 + y_2) \dots f(y_1 + y_2 + \dots + y_n)$$

Sachant que  $y_1 < y_1 + y_2 < \dots < y_1 + y_n + \dots + y_n$

Et que  $f(x) = e^{-x}$ , on a :

$$f(y_1) f(y_1 + y_2) \dots f(y_1 + y_2 + \dots + y_n) = e^{-y_1} e^{-(y_1 + y_2)} \dots e^{-(y_1 + y_2 + \dots + y_n)}$$

Posons  $y_1 = x_1$  et  $y_i = x_i - x_{i-1}$

Alors

$$\begin{aligned}
f(y_1)f(y_1+y_2)\dots f(y_1+y_2+\dots+y_n) &= e^{-ny_1}e^{-(n-1)y_2}\dots e^{-2y_{n-1}}e^{-y_n} \\
&= e^{-nx_1}e^{-(n-1)(x_2-x_1)}\dots e^{-(x_n-x_{n-1})} \frac{1}{|\det A|}
\end{aligned}$$

En insérant le résultat précédent dans  $f_Y(y) = f_X(A^{-1}y) \frac{1}{|\det(A)|}$  et en distribuant les facteurs  $n! = n(n-1)\dots 1$  dans le produit des exponentielles on obtient :

$$f_Y(y) = n e^{-nx_1} (n-1) e^{-(n-1)(x_2-x_1)} \dots e^{-(x_n-x_{n-1})} \frac{1}{|\det A|}$$

Maintenant on a que  $A$  est une matrice triangulaire, son déterminant est égal au produit de ses termes diagonaux.

Par conséquent  $\det A = 1$ .

En d'autres termes, on a obtenu :

$$\begin{aligned}
& f_{X_1, X_2 - X_1, \dots, X_n - X_{n-1}}(X_1, X_2 - X_1, \dots, X_n - X_{n-1}) \quad \% \% \\
& = n e^{-nx_1} (n-1) e^{-(n-1)(x_2-x_1)} \dots 2 e^{-2(x_{n-1}-x_{n-2})} e^{-(x_n-x_{n-1})}
\end{aligned}$$

En se rapportant à l'égalité (####) on constate que  $n e^{-nx_1}$  est la densité de probabilité de

$\exp\left(\frac{1}{n}\right)$ ,  $(n-1) e^{-(n-1)(x_2-x_1)}$  n'est que la densité de probabilité de  $\exp\left(\frac{1}{n-1}\right)$ , mais comme le facteur générique dans le produit de l'égalité %% est  $(n+1-i) e^{-^{(n+1-i)}(x_i-x_{i-1})}$  qu'est la densité de probabilité de  $\exp\left(\frac{1}{n+1-i}\right)$

En conclusion nous avons que le produit %% est le produit respectivement de densité de probabilité pour les variables  $X_1 \sim \exp\left(\frac{1}{n}\right)$  et  $X_i - X_{i-1} \sim \exp\left(\frac{1}{n+1-i}\right)$ .

## Chapitre 2: APPLICATION DE LOI DE ZIPF

### 1- Distribution de la Population des villes

La plupart des chercheurs qui travaillent sur les distributions des villes admettent une relation linéaire entre les logarithmes du rang et de la taille démographique d'une ville. Cet « a priori » fut récemment mis en cause par un ensemble d'approches qui s'interrogent sur les déviations éventuelles des distributions réelles vis-à-vis de la loi de Pareto, voire sur une meilleure adéquation de celles-ci avec d'autres lois statistiques, telles que la loi log-normale.

Selon Zipf, si  $S$  est la variable qui associe à chaque ville sa population, la fonction densité de  $S$  suit une loi de Pareto, en épousant la forme :

$$f(s) = bs^\beta \quad **$$

qu'on peut écrire aussi:

$$f(s) = C \zeta s^{(-\zeta-1)} \quad **$$

où:

- ◆  $C$  est une constante qui dépend de la taille de la plus grande ville du pays et du nombre de villes incluses dans l'échantillon  $e$
- ◆  $\zeta$  est un coefficient de hiérarchisation ou bien l'exposant de la loi de Zipf proche de 1.

Si on note par  $P(S > s)$  la probabilité qu'une ville ait une population supérieure à  $s$ , la fonction de répartition complémentaire associée à  $S$  est :

$$\bar{F}(s) = P(S > s) = \int_s^\infty f(t) dt = \int_s^\infty C \zeta s^{(-\zeta-1)} = Cs^{-\zeta} \quad **$$

**Remarque :**

- Gell-Mann [1994, p. 95] propose la modification de la formule :

$$P(S > s) = \frac{c}{(s+m)^\zeta} \quad ** \quad \text{tel que } c \text{ une constante .}$$

- On va s'intéresser au cas où  $m=0$  pour la raison : le nombre et La qualité des recherche qui traite ce cas.

Si l'on arrange les villes, selon leur ordre décroissant de population , de la façon suivante :

$s_1 > s_2 > s_3 > \dots > s_i > \dots > s_n$  , avec  $r(s)$  le rang de la ville de taille  $s$ , on obtient :

$$P(S > s) = \frac{r(s)}{n} \quad **$$

$$r(s) = a \cdot s^{-\zeta} \quad \text{tel que } a = nC \quad **$$

où  $a$  est une constante et  $\zeta$  proche de 1.

La relation linéaire entre les logarithmes de rang et de taille correspond à la version la plus connue de la loi rang-taille qui stipule que le rang d'une ville donnée est inversement proportionnel à sa taille :

$$\ln r(s) = \ln(a) - \zeta \ln(s) \quad **$$

Lorsque  $\zeta = 1$  est égal à 1, on obtient la loi de Zipf.

Zipf avance une vague justification de la régularité de cette loi dans la nature humaine qui conditionne les choix de vie et de localisation des individus.

Graphiquement dans un repère bi-logarithmique. Si la répartition des villes selon la population suit une loi puissance, on obtient une droite avec une pente proche de -1.

**Remarque :**

Au début, on va établir quelque notation. Si  $\tilde{S}$  est la taille d'une ville, la loi de zipf peut s'exprimer comme :

$$P(\tilde{S} > s) = c \cdot s^{-\zeta(s)} \quad ** \text{ pour un certain } c.$$

On retrouve une telle répartition en loi puissance dans la plupart des pays. Plusieurs interprétations ont été proposées pour expliquer cette loi. La plupart d'entre elles se basent sur la modélisation de la croissance des villes par un processus aléatoire, selon le même principe que celui employé par Simon pour les textes.

## **2-Modèle de Base ( loi de Gibrat)**

On suppose que notre échantillon  $(P_1, P_2, \dots, P_N)$  fixe contient la taille des villes tel que les différentes villes croissent aléatoirement avec le même taux de croissance au cours du temps et la même variance.

On normalise notre échantillon (en divisant chaque taille de villes par la populations urbaine total)  $(S_1, S_2, \dots, S_N)$  .

Ainsi on a  $\sum_{i=1}^N S_i = 1$  toujours vérifier à n'importe quelle moment, la probabilité que  $\tilde{S}$  à une taille supérieure à  $S$  est  $P(\tilde{S} > S) = P(S) = c S^{-\zeta}$  caractériser par la densité

$$f(s) = \zeta c s^{-\zeta-1} \text{ ainsi } \zeta + 1 = -s \frac{f'(s)}{f(s)}, \text{ d'une autre manière } \zeta = -s \frac{P'(s)}{P(s)}$$

Selon Gabaix, la croissance des villes peut être modélisée par la loi de Gibrat, alors notre distribution doit tendre vers la loi de Zip-f, c'est à dire

une loi puissance avec un exposant égal à  $-1$ .

En effet :

soit  $\gamma_{t+1}^i$  le taux de croissance de la ville  $i$  à la date  $t$ . La taille de la ville  $i$  à la date  $t+1$  est :

$$S_{t+1}^i = \gamma_{t+1}^i S_t^i \text{ et } \sum_1^N S_t^i = 1 \quad \forall t$$

donc :

$$E(\gamma) = \int_0^{\infty} \gamma f(\gamma) d\gamma = 1 \quad **$$

on note par  $G_t(S) = P(S_t > s)$  la proba que la taille de la ville soit plus grande que  $s$  à l'instant  $t$ .

$$\begin{aligned} G_{t+1}(S) &= P(S_{t+1} > s) \\ &= P(\gamma_{t+1} S_t > s) \\ &= E(1_{S_t > s/\gamma_{t+1}}) \\ &= E(E(1_{S_t > s/\gamma_{t+1}} | \gamma_{t+1})) \\ &= E(G_t(\frac{S}{\gamma})) \\ &= \int_0^{\infty} G_t(\frac{S}{\gamma}) f(\gamma) d\gamma \quad * \end{aligned}$$

supposons qu'on est en régime permanent c'est à dire que la fonction  $G$  ne dépend pas du temps donc:

$$G_t = G \quad \forall t$$

donc :

$$G = \int_0^{\infty} G(\frac{S}{\gamma}) f(\gamma) d\gamma \quad **$$



Une solution de cette équation est donnée par la loi de Zipf :

$$G(s) = \frac{C}{s} \quad ** \text{ avec } C \text{ une constante}$$

$$\text{donc : } \frac{C}{s} = \int_0^{\infty} \frac{C}{s} \gamma f(\gamma) d\gamma$$

$$1 = \int_0^{\infty} \gamma f(\gamma) d\gamma$$

Il est possible de rendre le résultat plus intuitive en deux parties :

- l'existence d'une loi de puissance qui est due à un simple principe de physique ; car le taux moyen de la croissance est supposé le même pour toutes les villes au cours du temps
- l'existence d'une loi de puissance de -1, un exemple pour clarifier: On construit un échantillon tel que la croissance d'une ville soit le double, ou la moitié sous condition que la taille moyenne soit toujours constante au cours du temps; c'est à dire que le nombre de villes avec une taille (2S) doit être la moitié de villes ayant une taille S, le nombre de villes avec une taille S/2 doit être le double des villes avec une taille S.

on pose la proba  $P(s=S/2)=a$  et  $P(s=2S)=a/2$ .

Et on a:  $a+a/2=1 \Rightarrow a=2/3$

Ainsi la proba qu'une ville aura une taille S/2, 2S est égale à 2/3 respectivement 1/3.

D'abord, si la variance est nulle on aura jamais une convergence. Krugman [1996a, pp 96-97] montre que ce processus prend du temps, Pour que notre échantillon de villes converge vers la loi de Zipf avec un exposant  $\zeta=1$ . Parce que il faut une variation suffisante de taux de croissance de la ville.

En utilisant la méthode de Monte Carlo, on constate que le coefficient de hiérarchisation soit proche de 1 dans environ une période de sept décennies c'est à dire dès que la variance total en partant de la distribution initial jusqu'à l'instant T atteint l'ordre  $\sigma \cdot T=0,7$ , la distribution est très proche d'une loi de puissance avec un exposant  $\zeta=1,05$ , dans un double temps le coefficient atteint la valeur  $\zeta=1,001$ .<sup>19</sup>

On conclut que selon ce modèle même si on prends un système urbaine petit et dynamique, il

va satisfaire la loi de zipf dans un certain temps.

**Remarque :Exemple de Variance pour différentes pays :**

-Dobkins et Loannides [1998a]donnent la variance pour les les villes de l'U.S dans le vingtième siècles égale à 20 % par décade.

-Pour U.K en 1800-1850 la variance a était calculée par Bairoch, Batou et Chèvre (1998) ,la variance était au environ de 5 % par décade .

-Eaton et Eckstein ont emmené une étude sur les villes des France et China dans le vingtième siècles telle que la variance était 1,26 % et 1,32 % par décade respectivement.

**3- Modèle canonique de Gabaix ( économique):**

En 1999 Gabaix développe un modèle de croissance urbaine qui conduit, à une distribution rang-taille des villes qui converge vers la loi de Zipf. Ce modèle suppose, de façon générale, que le changement des tailles des villes suit un processus stochastique, conforme à la loi de Gibrat. Ceci implique que toutes les villes ont la même espérance et variance de taux de croissance .

Gabaix s'appuie sur un ensemble d'hypothèses .L'échantillon est caractérisé par une population totale croissante .La mobilité du travail est réduite aux jeunes ménages à condition qu'ils peuvent migrer une seule fois, au début de leur entrée dans la vie professionnelle.Une fois leur choix de localisation effectué, ils ne changent plus de lieu d'habitation . Enfin, la production ne dépend que du facteur travail.

Sous ces hypothèses, la croissance urbaine apparaît comme une marche aléatoire qui dépend uniquement des mouvements migratoires des jeunes agents économique . Les choix de migration, dans chaque période, sont liés aux impacts de chocs exogènes(événement naturelles ou historique, impact des politique économique ,variétés des services municipaux ...) distribués de façon aléatoire.

Soit une ville  $i$  à l'instant  $t$  on note  $a_{it}$  représente le niveau d'aménités de celle ci, et  $c$  la consommation d'un agent économique habitant cette ville.Alors ce dernier a une fonction d'utilité sous forme :

$$u(c)=a_{it}c^{**}$$

Les niveaux d'aménités  $a_{it}$  dans les différentes villes sont i.i.d. Si  $w_{it}$  est le salaire unitaire dans la ville  $i$ , alors, à l'équilibre, les salaires ajustés par le niveau d'utilité seront

identiques dans toutes les villes :

$$a_{it} w_{it} = u_{it} \forall (i, t) \quad **$$

Posons  $S_i^m$  est le volume des migrants vers une ville  $i$  à l'instant  $t$  et  $S_i^0$  la population déjà en place, la fonction de production localisée au sein de cette ville, sous un régime de rendements constants, est :

$$F(S_i^0, S_i^m) = S_i^0 f(S_i^m / S_i^0) \quad **$$

avec  $f$  la fonction du volume de la population. Donc la production dépend que du salaire ainsi, les salaires dans chaque ville sont fonction du volume de la population :

$$w_i = \frac{\delta F}{\delta S_i^m} = f'(S_i^m / S_i^0) \quad **$$

Dans ce cas :

$$S_i^m = S_i^0 f'^{-1}(u_{it} / a_{it}) \quad **$$

En tenant compte du taux de mortalité des agents économiques qu'on note  $\theta$  donc la croissance de la population d'une ville  $i$  à l'instant  $t$  est :

$$\Delta S_{it} = S_{it}^m - \theta S_{it}^0$$

Ainsi le taux de croissance d'une ville  $i$  à l'instant  $t$  :

$$\gamma_{it} = \Delta S_{it} / S_{it} = f'^{-1}(u_{it} / a_{it}) - \theta = S_{it}^t / S_{it}^{t-1} \quad **$$

Comme la distribution des niveaux d'aménités  $a_{it}$  est indépendante de la taille des villes initial de la ville  $i$ , ainsi, la même espérance de croissance démographique .

Puisque notre échantillon est normalisé la somme totale des tailles des villes reste constante est égale à 1.

Donc on trouve le même résultat que le modèle de base :

$$G(s) = \int_0^{\infty} G\left(\frac{s}{y}\right) f(y) dy \quad \text{et} \quad E(y) = \int_0^{\infty} y f(y) dy = 1$$

avec  $G(S)$  représente la fonction de répartition complémentaire des tailles des villes les plus grandes, au moment  $t$ , alors, que celle-ci suit un processus stationnaire qui vérifie l'équation :

$$G(s) = \frac{C}{s} \quad **$$

Ceci signifie qu'au-delà d'une certaine taille urbaine, la croissance des villes de la distribution

rang-taille qu'obéit à la loi de Gibrat suit la loi de Zipf.

Dans ce modèle, la variance des taux de croissance ne décroît pas avec la taille des villes, contrairement à ce que supposent la plupart des études empiriques, selon lesquelles la diversification industrielle des grandes villes les rend moins exposées aux chocs sectoriels que les petites villes, davantage spécialisées. En suivant Kalecki (1945), Gabaix (1999) suggère que le taux de croissance d'une ville  $i$ ,  $\gamma_{it}$ , peut être décomposé en une fraction commune pour toutes les villes considérées et une fraction spécifique propre à chaque ville, région ou industrie :

$$\gamma_{it} = \bar{\gamma} + \gamma_{it}^{pol} + \gamma_{it}^{reg} + \gamma_{it}^{ind} \quad **$$

où  $\bar{\gamma}$  représente le taux de croissance démographique,  $\gamma_{it}^{pol}$  les impacts démographiques liés aux politiques, il est clair que sa variance ne dépend pas de la taille de la ville donc

$\sigma_{pol}^2(S) = \sigma_{pol}^2$ ,  $\gamma_{it}^{reg}$  les impacts liés aux performances macroéconomiques de la région urbaine tel que sa variance ne dépend pas de la taille car ces impacts sont distribués en égalités sur toutes les villes selon Gabaix  $\sigma_{reg}^2(S) = \sigma_{reg}^2$ .  $\gamma_{it}^{ind}$  qui représente les impacts démographiques des chocs économiques, est décroissante vis-à-vis de la taille urbaine, comme le suggèrent la plupart des études empiriques la variance est égale  $\sigma_{ind}^2(S) = \sigma_{ind}^2/S$ .

La variance du taux de croissance de la taille  $S$  de la ville  $i$  peut être décomposée de la même manière :

$$\sigma^2(S) = \sigma_{pol}^2 + \sigma_{reg}^2 + \sigma_{ind}^2/S \quad **$$

L'hypothèse que la variance est indépendante de la taille de la ville serait vraie dans la partie supérieure de la queue de distribution car on aura :

$$\sigma_{pol}^2 + \sigma_{reg}^2 \gg \sigma_{ind}^2/S \quad **$$

On conclut la variance totale est globalement indépendante vis-à-vis de la taille des villes.

$$\sigma^2(S) \simeq \sigma_{pol}^2 + \sigma_{reg}^2 \quad **$$

## **4-Modèles de Croissance aléatoire et effets de taille**

### **Définition : Mouvement brownien géométrique**

Soit  $\sigma > 0$  et  $r \in \mathbb{R}$  on appelle mouvement brownien géométrique la solution de l'équation

différentielle stochastique :

$$dX_t = \sigma X_t dB_t + rX_t dt \quad **$$

**Proposition :**

la solution de l'équation différentielle du mouvement géométrique brownien est :

$$X_t = X_0 \exp\left(\sigma B_t + \left(r - \frac{\sigma^2}{2}\right)t\right) \quad **$$

Démonstration :

- ◆ Traitons le cas particulière (cas  $r=0$ ) .

$$\frac{dX_t}{X_t} = \sigma dB_t$$

( Cas générale :  $\frac{dX_t}{X_t} = \sigma dB_t + rdt$  avec  $r - \frac{\sigma^2}{2} < 0$  )

On pose  $L_t = \log(X_t) = F(X_t)$  ,d'après la Formule d'Itô :

$$L_t = L_0 + \int_0^t F'(S_u) dS_u + \frac{1}{2} \int_0^t F''(S_u) d\langle S_u \rangle$$

Avec  $d\langle S_u \rangle = \sigma^2 S_u^2 du$  ( crochet d'Itô)

$$F'(x) = \frac{1}{x}$$

$$F''(x) = -\frac{1}{x^2}$$

Ainsi

$$\begin{aligned} L_t &= L_0 + \int_0^t \frac{dX_u}{X_u} - \frac{\sigma^2}{2} \int_0^t du \\ &= L_0 - \frac{\sigma^2}{2} t + \frac{\sigma^2 t}{2} \quad \left(\text{mouvement brownien avec drift } -\frac{\sigma^2}{2}\right) \\ &\Rightarrow X_t = X_0 e^{\sigma B_t - \frac{\sigma^2 t}{2}} \end{aligned}$$

**Remarque**

en remarque que si la valeur initiale  $X_0$  est strictement positive, la solution le reste en tout

temps  $t \geq 0$  . le mouvement brownien géométrique est utilisé dans le célèbre modèle de Black et Scholes en mathématiques financières . La raison de l'apparition de ce processus tient à la hypothèse économique d'indépendance des rendements successifs :

$$\frac{X_{t_2} - X_{t_1}}{X_{t_1}}, \frac{X_{t_3} - X_{t_2}}{X_{t_2}}, \dots, \frac{X_{t_n} - X_{t_{n-1}}}{X_{t_{n-1}}}$$

## Hypothèse du modèle

Cette section exposera le mécanisme qui génère la loi de Zipf avec plus de détail. Le point de fond à faire ici nécessite un mécanisme qui empêche les petites villes de devenir trop petites un tel mécanisme est donnée par une marche aléatoire avec une barrière inférieure.

Le modèle fonctionne en temps continu, On définit le processus de la taille d'une ville  $i$  à l'instant  $t$  par la formule :

$$dP_t = \mu P_t dt + \sigma P_t dB_t \quad **$$

On normalise notre distribution :

$S_{it} = P_{it} / P_t^T = P_{it} / P_0^T \exp(gt)$  tel que  $P_t^T$  est la population total à l'instant  $t$  et  $g$  la croissance attendu et  $\gamma_{it} = \Delta S_{it} / S_{it} = f^{-1}(u_{it} / a_{it}) - \theta = S_{it}^t / S_{it}^{t-1}$  avec :  $\Delta S_{it} = S_{it} - S_{it-1}$

$$\frac{dP_t}{P_t} = g dt + \sigma dB_t$$

tel que  $B_{it}$  est un mouvement brownien,  $\sigma$  et  $\gamma$  qui dépend des paramètres discrets du modèle  $(u_{it}, a_{it}, \theta)$  On obtient la formule suivante:

$$dS_{it} / S_{it} = \mu dt + \sigma dB_{it} \quad **$$

tel que  $\mu = \gamma(S) - \bar{\gamma}$  La croissance attendue en taille normalisée c'est la différence entre le taux de croissance des villes de taille  $S$  et le taux moyen de croissance .

### **Ajout d'une barrière infinitésimale $S_{\min}$**

D'après la proposition puisque la croissance de  $S_t$  est caractérisé par un mouvement brownien on a:

$$S_t = S_0 \exp(\sigma B_t + (\mu - \frac{\sigma^2}{2})t) \quad \text{avec} \quad \mu - \frac{\sigma^2}{2} < 0$$

le processus n'admet pas de distribution en régime permanent (S ne dépend pas de t), donc La distribution de la taille-ville serait juste une **log-normal**, où la plupart des villes auraient une taille infinitésimale, ainsi la nécessité d'un mécanisme qui empêche les petites villes de devenir trop petites.

On introduit une barrière inférieure  $S_{min}$  tel que :

On pose  $Y_t = \frac{S_t}{S_{min}}$  et Soit  $M_t = \log(Y_t) \quad t > 0$

Alors  $(M_t)_{t>0}$  devient un mouvement brownien avec drift réfléchi en zéro

En effet , par itô :

$$dM_t = \sigma dB_t - \frac{\sigma^2 dt}{2} \quad Y_t > 1 \Leftrightarrow M_t > 0 \quad \text{tel que} \quad \frac{-\sigma^2}{2} \quad \text{est le drift.}$$

**Remarque :**

- $Y_t$  est toujours supérieur ou égale à 1  $\Leftrightarrow M_t \geq 0$
- $S_t \rightarrow 0$  p.s quand t tend vers l'infini, il n'y a pas de loi invariante non dégénérée.
- Sans drift on a deux cas envisageable :
  - $S_t = S_0 e^{\sigma B_t + \mu t}$  donc soit  $S_t$  tends vers l'infini
  - cas  $\mu = 0$  ainsi  $S_t = S_0 e^{\sigma B_t}$  ne converge pas

On conclut que le drift négative qui consiste à réfléchir  $M_t$  en 0 pour la forcer à rester positif et crée la mesure invariante .

**proposition 1:**

Soit  $S_{min}$  la taille normalisée minimale. Supposons que la distribution de la taille

normalisée des villes suit le processus :

$$\begin{cases} dS_t = S_t \max(\mu dt + \sigma dB_t, 0) & \text{pour : } S_t < S_{min} \\ dS_t / S_t = \mu dt + \sigma dB_t & \text{pour : } S_t > S_{min} \end{cases}$$

Alors la distribution converge vers la loi de Zipf avec un coefficient de hiérarchisation

$$\zeta = 1 / (1 - S_{min} / \bar{S}) \quad \text{tel que} \quad E(S) = \bar{S} \quad \text{est la moyenne des tailles-ville.}$$

### Remarque:

- Suppose qu'il existe une taille minimale est suffisant pour que la distribution tend la loi Zipf .En effet ,si  $S_{min}$  tend vers 0 l'exposant converge vers 1 .Ainsi , il suffit de prendre une barrière infinitésimale pour assurer la convergence avec un exposant proche de 1
- la taille  $S_{min}$  doit avoir une taille relative fixe.Selon Harrison [1985, p. 15] la barrière infétesimale doit avoir un taux de croissance  $\gamma > \bar{\gamma} - \sigma^2$

### Preuve:

Considérons le Log-taille  $s = \ln(S)$  qui suit le processus :

$$ds = \sigma dB_t + \left(\mu - \frac{\sigma^2}{2}\right) dt \quad **$$

et  $s > s_{min} = \ln(S_{min})$  , Harrison [1985 , p. 15] montre que la distribution de  $s$  suit une loi exponentiel tel que :  $P(s > s') = \exp(-\zeta(s' - s_{min}))$  pour  $s' > s_{min}$  et un certain  $\zeta$

Donc  $P(S > S') = (S'/S_{min})^{-\zeta}$  pour  $S' > S_{min}$

$$\begin{aligned} E(S) &= \zeta S_{min}^{\zeta} \int_{S_{min}}^{\infty} S'^{-\zeta} dS' \\ &= \zeta S_{min}^{\zeta} [S'^{1-\zeta} / (1-\zeta)]_{S_{min}}^{\infty} \\ &= \zeta S_{min} / (\zeta - 1) = \bar{S} \end{aligned}$$

Ainsi le coefficient de hiérarchisation égale :

$$\zeta = \frac{\bar{S}}{\bar{S} - S_{min}} = 1 / (1 - S_{min} / \bar{S})$$

### 3 Modèles général

cette partie présente un ensemble de travaux qui explorent, de façon empirique, la nature du changement urbain, en admettant un rôle moins déterministe des effets de taille. Gabelle et Ioannides [2004] admettent que la loi de Gibrat (modele de base) n'est pas toujours validée du point de vue empirique. Alors, ils supposent une hypothèse moins restrictive, où le taux de croissance démographique, ainsi que sa variance, ne sont pas nécessairement indépendants de la taille.



Dans ce modèle on normalise notre distribution , la taille de la ville normalisé i varie, durant la période t , de la façon suivante :

$$dS_t/S_t = \mu(S_t) dt + \sigma(S_t) dB_t$$

avec  $\mu(S_t) = \gamma(S_t) - \bar{\gamma}$  la moyenne espérée du taux de croissance d' une ville de taille  $S_t$  au temps t ,  $\sigma(S_t)$  son écart type et  $B_t$  le mouvement brownien géométrique qui illustre la trajectoire irrégulière du taux de croissance.

**Remarque : (Dans le cas de base)**

- Sous les conditions de loi Gibrat, La moyenne des taille normalisées doit rester constante donc  $E(S) = \frac{1}{N}$  ce qui implique  $\mu = 0$  car le taux de croissance espérée est indépendant de la taille donc égale au taux moyen de croissance.

On note  $p(S, t)$  la distribution de la taille s à l'instant t est une fonction de classes  $C^2$

Selon la formule d'itô l'équation du mouvement s'écrit sous la forme suivante:

$$\frac{\partial}{\partial t} p(S, t) = \frac{\partial}{\partial S} (\mu(S) S p(S, t)) + \frac{1}{2} \frac{\partial^2}{\partial S^2} (\sigma^2(S) S^2 p(S, t)) \quad **$$

Remarque :

- la moyenne de taille à l'instant t est égale :

$$E(S_t) = E(S_0) \exp(\mu t)$$

- Considérons le cas l'état stationnaire sous la conditionne  $\mu = 0$  i.e :

$$\begin{cases} \frac{\partial}{\partial t} p(S, t) = 0 \\ p(S, t) = p(S) \\ \frac{1}{2} \frac{\partial^2}{\partial S^2} (\sigma^2(S) S^2 p(S)) = 0 \end{cases}$$

Une solution du Système est  $p(S) = cS^{-1}$  avec c une constante qui dépend toujours de la queue de la distribution .

La distribution rang-taille des villes tend, alors, vers la loi de Zipf.

- A l'inverse, lorsque la croissance des villes dépend de leur taille ( $\mu \neq 0$ ).

$$-\mu(S)Sp(S) + \frac{1}{2} \frac{\partial^2}{\partial S^2} (\sigma^2(S)S^2 p(S)) = 0 \quad **$$

Donc la distribution rang-taille des villes converge vers une distribution théorique dont le coefficient de hiérarchisation est égal à :

$$\zeta(S) = 1 - 2\mu \frac{(S)}{\sigma^2(S)} + \frac{S}{\sigma^2(S)} \frac{\partial \sigma^2(S)}{\partial S} \quad (*****)$$

Selon Gabaix et Ioannides (2004), les déviations vis-à-vis de la loi de Zipf sont liées à deux causes. Si certaines villes ont un taux de croissance élevé  $\mu(S) = \gamma(S) - \bar{\gamma} > 0$ , alors la distribution pourraient avoir tendance à décroître moins rapidement (cas de la loi Zipf pure car  $\zeta(S) < 1$ ).

Si, à l'inverse, la variance du taux de croissance est élevée, la courbe de la distribution rang-taille a tendance à s'aplatir.

### Conclusion

la formule (\*\*\*\*\*) aide à évaluer dans quelle mesure le modèle économique (pour les petites et moyennes villes; par exemple de rang > 100) peut dévier de la loi de Gibrat.

Étudions la convergence de  $\mu(S)$  et  $\sigma(S)$  :

Barro et Sala-Martin [1995] déduits que empiriquement  $0.8 < \zeta(S) < 1.2$  et

$$\left| \frac{S}{\sigma^2(S)} \frac{\partial}{\partial S} \sigma^2(S) \right| < 0.2 \quad **$$

On revenant aux résultats du modèle Basic (Gibrat) où la variance est estimée par :

$\sigma^2 = 0.1 / \text{décennie}$  la formule (\*\*\*\*\*) indique que le taux de croissance vérifie l'inégalité

$|\gamma(S) - \bar{\gamma}| < 0.4$  et  $\sigma^2/2 = 0.2$  par décennie.

Par conséquent, l'écart par rapport à la loi de Gibrat pour le taux de croissance moyenne est assez faible.

## 5-Modèle Dynamique urbaine (nouvelles villes)

Dans cette partie on suppose la possibilité d'apparition des nouvelles villes dans la région

urbaine étudiée. On cherche sous quelles conditions cette hypothèse n'affecte pas la tendance de la distribution vers une loi de Zipf

**Proposition :**

notons  $\nu$  le taux de croissance de nombre de ville,  $\gamma$  le taux de croissance taille-ville des villes existantes.

Si  $\nu \leq \gamma$  :

Alors, les queues de distributions satisfont la loi de Zipf avec un coefficient de hiérarchisation égal à 1

Si  $\nu \geq \gamma$  : Dans le cas du temps continue, l'état stationnaire de la distribution suit une loi de puissance avec un exposant qui est la racine positive de l'équation de deuxième ordre au dessous :

$$\zeta^2 - (1 - 2\gamma/\sigma^2)\zeta - 2\nu/\sigma^2 = 0 \quad **$$

**Remarque :**

- En particulier  $\zeta > 1$  .

**Preuve :**

- ◆ cas  $\nu < \gamma$  :

Considérons  $t$  la limite du temps pour laquelle on mesure la taille des villes temps qui tend vers l'infini .

On fixe  $T = t/2$  , on différencie entre les anciennes villes qui existent avant  $T$  et les nouvelles qui apparaissent après  $T$ . On note  $S = \exp(s)$  une taille-ville qui se situe à la queue de distribution (très grande)

la probabilité de trouver une ville avec une taille plus grande que  $S$  est :

$$P(\tilde{S} > S) = P(\tilde{S} > S, b < T) + P(\tilde{S} > S, b \geq T) \quad **$$

tel que  $b$  La variable aléatoire de la date de naissance des villes.

On va montrer que le premier terme domine le deuxième, donc les nouvelles villes n'affectent pas sur la tendance vers une loi de Zipf dans la queue de distribution.

D'abord, on a pour les villes nées à la date  $\tau < T$

$$P(\tilde{S} > S, b = \tau) \sim \exp(-(s - \gamma(t - \tau)) \quad (**)$$

(car  $S = \exp(s)$  et la constante  $c$  caractériser par la queue de distribution inférieure  $\{ S_{min} \}$  égale à l'instant  $\tau = \exp(\gamma(t - \tau))$ )

L'age de ces villes à l'instant  $t$  est supérieure à  $t - T$  qui tend vers l'infini, donc la (proportion 1) sans les nouvelles villes est vérifiée. Le terme  $\exp(-\gamma(t - \tau))$  est La normalisation appropriée pour les cités née à l'instant  $b = \tau$  tel que la moyenne des tailles à l'instant  $t$  est  $\exp(\gamma(t - \tau))$ , rappelons que  $E(dS_t/S_t) = \bar{\gamma}$ . Le processus de naissance des nouvelles ville à l'instant  $t$  est défini par :

$$N_t = \int_0^t v e^{v\tau} d\tau = e^{vt} - 1 \quad **$$

La densité qui définit ce processus à l'instant  $t$  des villes nées à l'instant  $\tau$  est

$$v e^{vt} / N_t = v e^{v(t - n_t)} \quad \text{telle que } n_t = \ln(N_t)$$

On intègre (\*\*), on obtient :

$$\begin{aligned} P(\tilde{S} > S, b < T) &\sim \int_0^T \exp(-s + \gamma(t - \tau)) v \exp(v\tau - n_t) \\ &= e^{(-s + \gamma t - n_t)} v \int_0^T e^{-(\gamma - v)\tau} d\tau \\ &= e^{(-s + \gamma t - n_t)} v \left( \frac{e^{-(\gamma - v)T} - 1}{\gamma - v} \right) \end{aligned}$$

puisque  $v \leq \gamma$  et  $T$  tend vers l'infini alors :  $P(\tilde{S} > S, b < \tau) \sim e^{(-s + \gamma t - n_t)} v / (\gamma - v)$

Ainsi les villes qui sont nées à l'instant  $\tau > T$  n'ont pas le temps pour converger vers une loi de Zipf

Dans le cas  $b = \tau$  :  $P(\tilde{S} > S, b = \tau) \leq E[\tilde{S}/S | b = \tau] = e^{-s + \gamma(t - \tau)}$

Ainsi  $P(\tilde{S} > S, b \geq T) \leq \int_T^t e^{-s + \gamma(t - \tau)} v e^{v\tau - n_t} d\tau \sim \frac{v e^{-s + \gamma t - (\gamma - v)T - n_t}}{\gamma - v}$

$$\begin{aligned} P(\tilde{S} > S, b \geq T) &= O(P(S_t > S, b < T) e^{-(\gamma - v)T}) \\ &= o(P(S_t > S, b < T)) \end{aligned}$$

Pour T très grande

$$\begin{aligned}
 P(S_t > S) &= P(S_t > S, b < T) + P(S_t > S, b \geq T) \\
 &\sim P(S_t > S, b < T) \\
 &\sim e^{(-s + \gamma t - n_t)} \nu / (\gamma - \nu)
 \end{aligned}$$

Par suite :

$$P(S_t > S) = a_t e^{-s} = a_t / S$$

- cas  $\nu \geq \gamma$

Pour étudier ce cas on va se mettre dans le cas du temps continu. On note  $p(S, t)$  la distribution (densité) de  $S$  à l'instant  $t$ . L'évolution de  $p(S, t)$  est donnée par l'équation de Kolmogorov [Forward Equation] modifiée pour tenir compte de l'apparition des nouvelles villes.

$$\frac{\partial}{\partial t} p(S, t) = -\frac{\partial}{\partial s} (\gamma S p(S, t)) + \frac{1}{2} \frac{\partial^2}{\partial s^2} (\sigma^2 S^2 p(S, t)) - \nu p(S, t)$$

telle que le terme  $\nu p(S, t)$  explique le fait qu'il y a des nouvelles villes.

Dans le cas d'état stationnaire puisque la distribution ne dépend pas du temps c'est à dire la densité dépend que de la taille de ville :

$$-\gamma (S p(S, t))' + \frac{1}{2} \sigma^2 (S^2 p(S, t))'' - \nu p(S, t) = 0 \quad **$$

On a d'une autre part  $\zeta + 1 = -s \frac{p'(s)}{p(s)}$  et  $(\zeta + 1)(\zeta + 2) = -s \frac{p''(s)}{p(s)}$

ainsi on divise par  $p(s) \neq 0$  dans la relation \*\* on trouve :

$$-\gamma(1 - \zeta - 1) + \frac{1}{2} \sigma^2 (2 - 4(\zeta + 1) + (\zeta + 1)(\zeta + 2)) - \nu = 0$$

d'où :

$$h(\zeta) = \zeta^2 - \zeta \left(1 - \frac{2\gamma}{\sigma^2}\right) - \frac{2\nu}{\sigma^2} = 0 \quad (\text{parabole})$$

$h(0) < 0$  donc il existe qu'une seule racine positive, de plus  $h(1) < 0$  quand  $\nu < \gamma$  implique que  $\zeta > 1$ .

## 6-Modèle Rang-Taille :

On peut explorer le problème « règle de distribution rang taille »

On suppose que la probabilité qu'une ville ait une taille supérieure à  $S$  décroît de  $1/S$ .

Nous devons s'attendre que la taille  $S(i)$  d'une ville de rang  $i$  suit une loi de puissance.

La taille de une ville de rang  $i$  varie de  $1/i$  et le ratio de la deuxième plus grande ville par rapport à la plus grande ville est de  $1/2$  et le ratio de la 3eme à la 2eme est  $2/3$  et ainsi de suite d'où la règle de distribution rang taille.

Ces rapports de tailles sont utilisés souvent pour comparer les modèles urbains actuelle aux modèles « idéal » de Zipf.

En réalité même si la loi Zipf est vérifiée exactement, la règle de rang taille sera vérifié approximativement.

Si notre interprétation probabiliste de Zipf est correcte on doit pas s'attendre que les ratios soient trop proches de leur valeurs idéales

### Proposition4 :

On ordonne les villes par taille  $S(1) \geq S(2) \geq \dots$  et on suppose que le queue de distribution

satisfait la loi Zipf avec un exposant de 1.

alors pour  $i < j$  :

la moyenne de  $S(j)/S(i)$  est :  $i/j$ .

Et l'écart-type est :  $\sqrt{\left(\left(1 - \frac{i}{j}\right)\left(\frac{i}{j(j+1)}\right)\right)}$

Le médiane est :  $B^{-1}(1/2, i, j)$  (inverse de la la fonction Bêta).

ces résultats sont valables aussi pour un nombre de ville fini.

### Démonstration :

tout d'abord on commence par énoncer le théorème de Rényi (Rényi(1953))

#### théorème :

soient  $X_1 \geq X_2 \geq \dots \geq X_n$  les statistiques d'ordre d'un échantion exponentiel, alors

$$(X_n, \dots, X_i, \dots, X_1) \sim \left(\frac{Y_n}{n}, \dots, \sum_{k=i}^n \frac{Y_k}{k}, \dots, \sum_{k=1}^n \frac{Y_k}{k}\right)$$

$Y_{1, \dots, n}$  sont des variable aléatoires exponentielles indépendantes

On ordonne les villes par taille  $S(1) \geq S(2) \geq \dots$

et on note  $s(1) \geq s(2) \geq \dots$  le log taille  $s(i) := \log(S(i))$

par normalisation selon la loi de Zipf, les logs tailles suivent une distribution exponentielle

$P(s > t) = e^{-t}$  pour  $t \geq 0$ .

On peut écrire  $s(i) - s(j) \sim \sum_{k=i}^{j-1} \frac{X_k}{k} = X_{i,j}$   $i < j$  avec les  $X_k$  sont des tirages indépendants d'une distribution exponentielle  $P(X_k > t) = e^{-t}$  pour  $X \geq 0$ .

$$E[\exp(-s X_{i,j})] = \prod_{k=i}^{j-1} E(\exp(\frac{-sL}{k})) \quad \text{avec } L \sim \exp(1)$$

$$= \prod_{k=i}^{j-1} \left( \frac{1}{1+s/k} \right)$$

$$= \frac{i \dots (j-1)}{(i+s) \dots (j-1+s)}$$

$$= \frac{(j-1)!}{(i-1)!} \frac{\Gamma(i+s)}{\Gamma(j+s)}$$

ou  $\Gamma(z) = \int_0^{\infty} \exp(-t) t^{z-1} dt$   
en effet  $\Gamma(z+1) = z \Gamma(z)$

donc  $E[\exp(-s X_{i,j})] = \frac{\Gamma(j) \Gamma(i+s)}{\Gamma(i) \Gamma(s+j)}$   $s > 0$

d'autre part, il est facile de voir que si  $B_{i,j}$  est une loi bêta on a :

$$E[B_{i,j}^s] = \frac{\Gamma(j) \Gamma(i+s)}{\Gamma(i) \Gamma(s+j)} = E[\exp(s \log B_{i,j})]$$

par identification des transformées de Laplace, on en déduit :

$$X_{i,j} = -\log B_{i,j}$$

et donc  $\frac{S(j)}{S(i)} = \exp(-X_{i,j}) \sim B_{i,j}$

donc :

$$E\left[\frac{S(j)}{S(i)}\right] = E[\exp(s(j) - s(i))] = E\left[\sum_{k=i}^{j-1} \frac{X_k}{k}\right] = \prod_{k=i}^{j-1} E\left[\exp\left(\frac{-X_k}{k}\right)\right] = \prod_{k=i}^{j-1} \frac{k}{k+1} = \frac{i}{j}$$

Notons que de la même façon on peut avoir  $E\left[\frac{S(i)}{S(j)}\right] = \frac{j-1}{i-1}$

$$E\left[\left(\frac{S(j)}{S(i)}\right)^2\right] = \prod_{k=1}^{j-1} E\left[\exp\left(-2\frac{X_k}{k}\right)\right] = \prod_{k=1}^{j-1} \frac{k}{k+2} = \frac{i(i+1)}{j(j+1)}$$

d'où :

$$\sigma = \sqrt{\left(1 - \frac{i}{j}\right)\left(\frac{i}{j(j+1)}\right)}$$

Finalement pour la médiane :

$$\begin{aligned} P\left(\frac{S(j)}{S(i)} \leq 1/2\right) &= P(\exp(s(j) - s(i)) \leq 1/2) \\ &= \frac{\int_0^{1/2} u^{i-1}(1-u)^{j-1} du}{\int_0^1 u^{i-1}(1-u)^{j-1} du} = B(1/2, i, j) \end{aligned}$$

et comme la fonction de repartition de Bêta est continue croissante sur  $[0, 1]$  elle est inversible d'où la médiane est donnée par :

$$B^{-1}(1/2, i, j)$$

### Exemple :

statistiques de ratio  $S(j)/S(i)$  entre la  $j$ -ème et la  $i$ -ème plus grandes villes ( $i < j$ )

	$S_{(j)}/S_{(i)}$	$S_{(2)}/S_{(1)}$	$S_{(3)}/S_{(1)}$	$S_{(3)}/S_{(2)}$	$S_{(10)}/S_{(5)}$	$S_{(100)}/S_{(10)}$
Mean	$i/j$	.5	.333	.6667	.5	.1
Standard deviation	$\sqrt{\frac{(1-i/j)i/j}{j+1}}$	.2887	.2357	.2357	.1508	.0299
Median <sup>a</sup>	$B^{-1}(1/2, i, j-i)$	.5	.2929	.7071	.5	.0973

On remarque d'après ce tableau que l'écart type de ces ratios est assez haut par exemple, le rapport de taille entre la ville 2 et la ville 1 a une moyenne de 1/2 mais un écart type assez élevé 0,2887 même le rapport entre la ville 100 et la ville 10, qui a une valeur attendue de 1/10, a un écart type de 0,03.



donc le fait que dans un pays donné la deuxième et troisième villes sont de tailles proches ne réfute pas la loi Zipf,

## 7-Le modèle de Simon et Steindl.

### -Le modèle de croissance de Steindl.

Dans le modèle de Steindl, les villes sont nées à un taux  $v$ , et les villes existantes croissent à un taux  $\gamma$ .

A l'instant  $t$  :

la taille d'une ville est :  $S_t = \exp(\gamma \cdot t)$

le nombre de villes est :  $N_t = N_0 \exp(v \cdot t)$

et la proportion de villes d'âge supérieure ou égale à  $a$  est :  $\frac{N_{t-a}}{N_t} = \exp(-v \cdot a)$

donc :

$$P(S_t \geq S) = P\left(t \geq \frac{\ln S}{\gamma}\right) = P(\text{age} \geq \ln \frac{S}{\gamma})$$

les villes de taille supérieure à  $S$  sont les villes d'âge supérieur à  $a = \ln \frac{S}{\gamma}$

le nombre de ces villes est proportionnel à  $\exp(-v \cdot a) = \exp\left(\frac{-v}{\gamma} \ln S\right) = S^{-v/\gamma}$ , ce

qui donne l'exposant  $\zeta = \frac{v}{\gamma}$

Il en résulte que la distribution de nouvelles villes sera sous la forme d'une loi de puissance, avec un exposant  $\zeta = \frac{v}{\gamma}$ .

Cependant, cela pose problème. Il ne fournit pas le

Résultat que nous voulons, à savoir l'exposant de 1, sauf dans le cas où  $v = \gamma$ .

C'est qui est peu convaincant, en particulier pour les systèmes urbains matures comme Paris par exemple.

### Le modèle de Simon

A l'aide de la loi de Gibrat, Simon (1955) a construit un modèle dont la population n'augmente pas de façon régulière,

Dans ce modèle, de nouveaux migrants arrivent à Chaque période, et avec une probabilité  $\pi$ , ils formeront une nouvelle ville avec probabilité  $(1 - \pi)$  ils vont dans une ville existante. La probabilité qu'ils choisissent de se localiser dans une ville donnée est proportionnelle à sa population.

$$N_{\tau} = N_0 \exp(\gamma_0 \cdot t)$$

$$\text{avec : } \tau = \exp(\gamma_0 \cdot t)$$

$$N_0 = \pi \quad \text{car } \frac{N_{\tau}}{\tau} \rightarrow \pi \quad \text{par la loi des grands nombre}$$

$$\gamma_0 \quad \text{taux de croissance du nombre de villes}$$

pour une ville donnée, la croissance par unité de période est :

$$(1 - \pi) * \frac{S_{\tau}}{U_{\tau}} \quad \text{avec } U_{\tau} \text{ nombre de nouveau habitants par periode } \tau$$

$$\text{d'où l'accroissement en moyenne est : } E[S_{\tau-1} - S_{\tau}] = \frac{(1 - \pi)}{U_{\tau}} \cdot S_{\tau}$$

$$\text{si } S_t' = S_{e^{\gamma_0 t}} \quad \text{donc le taux de croissance de chaque ville}$$

$$\text{ona } dS_t' = (1 - \pi) \gamma_0 S_t' dt$$

$$\gamma = (1 - \pi) \gamma_0$$

$$\text{par Steindl, } \gamma_0 = (1 - \pi) \gamma_0 = \frac{1}{1 - \pi}$$

d'où ce modèle génère une loi de puissance avec l'exposant  $\zeta = 1/(1 - \pi)$ .

### Remarque :

En premier lieu, la teneur théorique du modèle est très faible car la croissance de la population et les choix de localisation des différents blocs ne reposent pas sur des mécanismes économiques. Simon exclut ainsi toute relation entre la croissance urbaine et des facteurs traditionnellement considérés comme déterminants des choix de localisation et de migration des agents, tels que les économies externes et les coûts de transport.

En deuxième lieu, le modèle de Simon converge vers la loi de Zipf, dans un processus infiniment lent, afin d'atteindre la situation où  $\pi = 0$ .

Krugman qualifie de situation de dégénérescence (Krugman, 1996a et 1996b, Fujita et al., 1999a).

Gabaix(1999) explique que le modèle de Simon fonctionne si seulement le taux de croissance du nombre de villes (c'est-à-dire le taux d'apparition de villes

nouvelles) est supérieur ou égal au taux de croissance démographique des villes existantes. Enfin, la troisième critique concerne la définition même des villes qui n'est, pour Simon, qu'une simple addition de blocs. Ceci signifie que l'on pourrait remplacer la ville par d'autres niveaux d'agrégation, tels que la région ou le pays, or, la loi de Zipf ne marche pas pour ce type d'espace. Soit

## **8-Système urbaine former par des régions hétérogènes**

Considérons l'hypothèse que les régions du système urbaine en le même taux de croissance. Car si on suppose que les régions se comportent différemment, on aura pas la convergence vers une loi Zipf.

### **Proposition :**

supposons que le système urbaine se compose de  $R$  régions. Telle que pour chaque régions l'hypothèse la variable taille de région normalisée suit un processus de mouvement Brownien géométrique, ainsi  $\zeta = 1/(1 - S_{min}/\bar{S})$ . ainsi pour chaque région la loi zipf est vérifiée.

En particulier le processus de croissance pour chaque région est identique (pour la queue de distribution) avec les autres région.

Alors la distribution asymptotique du système urbaine existe et satisfasse la loi zipfs avec le même exposent des régions

### **Démonstration :**

**Considérons chaque régions corresponds à un système urbaine sous les même hypothèses, donc pour une région  $j$  on a la probabilité que une ville à une taille supérieure à  $s$  est :**

$$P(\tilde{S} > s | \tilde{S} \in \text{region } j) \sim \frac{c_j}{s^{\zeta}}$$

Notons par  $\lambda_j$  la probabilité qu'une ville soit dans la région  $j$  ( $\sum_{j=1}^R \lambda_j = 1$ ).

On passe au niveau du tout le système on a :

$$\begin{aligned} P(\tilde{S} > s) &= \sum_{j=1}^R P(\tilde{S} > s, \tilde{S} \in \text{région } j) \\ &= \sum_{j=1}^R \lambda_j P(\tilde{S} > s | \tilde{S} \in \text{région } j) \\ &\sim \sum_{j=1}^R \lambda_j \frac{a_j}{s^{\zeta}} = \frac{a}{s^{\zeta}} \quad \text{avec} \quad a = \sum_{j=1}^R \lambda_j a_j \end{aligned}$$

Ainsi, l'Europe doit suivre une loi Zipf si tous les pays européens satisfont à la loi (sauf si le marin sera élu comme présidente), de même pour la France suit la loi Zipf si chaque région tend vers une loi Zipf.

## 9-Dérivation alternative du mécanisme de croissance avec les processus Kesten

Considérons que l'évolution de la distribution des tailles normalisées est caractérisée par

l'équation :  $S_{t+1} = \gamma_{t+1} S_t + \epsilon_{t+1}$ , \*\*\*

tel que  $\epsilon_{t+1}$  est trop petit avec un espérance  $E[\epsilon_{t+1}] = \bar{\epsilon} \geq 0$ .

A la limite  $\bar{\epsilon} = 0$ , égale à zéro, on est dans le cas pur de modèle de Gibrat.

Considérons le cas  $\bar{\epsilon} > 0$ . Quand la taille d'une ville est très grande la variable  $\epsilon_{t+1}$  est négligeable et proche de 0, ce qui mène au cas de Gibrat pour les grandes villes

$S_{t+1} \sim \gamma_{t+1} S_t$ , c'est à dire le terme  $\epsilon$  est significatif dans l'équation \*\*\* juste pour les villes de taille petite. D'autre part c'est le mécanisme qui empêche les petites villes de devenir d'une taille infinitésimale analogue avec la barrière inférieure dans le modèle de croissance aléatoire

### Proposition

Supposons que la distribution est normalisée et que les villes suivent le processus de croissance suivant :  $S_{t+1} = \gamma_{t+1} S_t + \epsilon_{t+1}$ , avec  $E[\epsilon_{t+1}] = \bar{\epsilon} > 0$ .

Alors pour n'importe quelle état initial de la distribution. Elle va tendre vers une loi de puissance Zipf avec un exposant  $\zeta$  tel que  $E(\gamma^\zeta)=1$ .

Pour  $\bar{\epsilon}$  petit on a  $\zeta=1+O(\bar{\epsilon})$ , ainsi quand  $\bar{\epsilon}\rightarrow 0$ , la distribution converge vers une loi Zipf avec un coefficient de hiérarchisation 1.

De plus, si  $\gamma_{t+1}$  suit une loi log-normal avec  $\sigma^2=Var(\ln(\gamma_{t+1}))$ . Alors l'exposant exacte vérifie :  $\zeta=1-2\ln(1-\frac{\bar{\epsilon}}{S})$  tel que  $\bar{S}$  est la moyenne de taille-ville.

**Demonstration :**

puisque on est dans le cas d'une distribution normalisé alors  $E(S_t)=\bar{S}$  pour tout  $t>0$ , alors

$$E(\gamma_{t+1})=1-\frac{\bar{\epsilon}}{S} \text{ et on note } \epsilon=\frac{\bar{\epsilon}}{S} .$$

Par inégalité de Jensen on a :  $E(\ln \gamma)\leq \ln(1-\epsilon)<0$

Le processus de Kesten nous permet de conclure que la distribution converge vers la loi Zipf avec un exposant  $\zeta$  qu'est la racine positive de  $E(\gamma^\zeta)=1$ .

Supposons l'état stationnaire ,on note  $G(S)=P(S_t>S)\sim cS^{-\zeta}$  car  $S_{t+1}\sim \gamma_{t+1}S_t$  dans la queue de distribution supérieure ,rappelons que dans l'état stationnaire on a :

$$\int_0^\infty G(\frac{S}{\gamma^\zeta})f(\gamma)d\gamma\sim G(S)$$

ainsi  $E(\gamma^\zeta)=1$ .

Dans le cas  $\gamma_{t+1}$  suit une loi log-normal avec  $\sigma^2=Var(\ln(\gamma_{t+1}))$  c'est à dire

$$\ln(\gamma_{t+1})\sim N(1-\epsilon, \sigma^2)$$

Soit U une variable qui suit la loi normale centrée réduite ,alors

$$\exp(U)\sim \text{log-normal}(0,1)$$

$$\text{donc } \exp(\ln(1-\epsilon)-\frac{\sigma^2}{2}+\sigma U)\sim \text{log-normal}(\frac{-\sigma^2}{2}+\ln(1-\epsilon), \sigma^2)$$

$$\text{ainsi : } \gamma_{t+1}\sim \exp(\ln(1-\epsilon)-\frac{\sigma^2}{2}+\sigma U)$$

$$\text{d'où : } E(\gamma_{t+1}^\zeta)=\exp((\ln(1-\epsilon)-\frac{\sigma^2}{2})\zeta+\zeta^2\frac{\sigma^2}{2})$$

Par suite la racine positive de  $E(\gamma_{t+1}^\zeta) = 1 \Rightarrow (\ln(1-\epsilon) - \frac{\sigma^2}{2})\zeta + \zeta^2 \frac{\sigma^2}{2} = 0$   
 $\Rightarrow \zeta = 1 - 2\ln(1-\zeta)/\sigma^2$

Dans le cas général où  $\gamma_{t+1}$  n'est pas forcément une log-normal, mais avec  $\epsilon$  est proche de

à, on définit  $\gamma_0 = \frac{\gamma}{1-\epsilon}$ .

Alors  $\zeta$  est la racine positive de  $f(\zeta, \epsilon) = 1$  avec  $f(\zeta, \epsilon) = E(\gamma_0(1-\epsilon)^\zeta)$ .

D'après le théorème de fonction implicite dans le voisinage de (0,1).

on a  $f(0, 1) = 1$

$$f(\epsilon, \zeta) = E(\gamma_0^\zeta (1-\epsilon)^\zeta)$$

$$\frac{\partial f}{\partial \epsilon} = \frac{-\zeta}{1-\epsilon} E(\gamma_0^\zeta (1-\epsilon)^\zeta)$$

$$\frac{\partial f}{\partial \zeta} = E(\ln(\gamma_0(1-\epsilon))(\gamma_0^\zeta (1-\epsilon)^\zeta))$$

Soit  $\zeta = \zeta(\epsilon)$  solution de  $f(\epsilon, \zeta) = 1$

On a  $\zeta = \zeta(0) - \left( \frac{\frac{\partial f}{\partial \epsilon}}{\frac{\partial f}{\partial \zeta}}(0, 1) \right) \epsilon + o(\epsilon)$  (théorème des fonction implicite)

puisque :

$$\frac{\partial f}{\partial \epsilon}(0, 1) = \frac{-1}{1-0} E(\gamma_0^1 (1-0)^1) = -1$$

$$\frac{\partial f}{\partial \zeta}(0, 1) = E(\ln(\gamma_0))$$

Alors :

$$\zeta = \zeta(0) + \frac{\epsilon}{E(\ln(\gamma_0))} + o(\epsilon)$$

avec :

$$E(\ln(\gamma_0)) = E(\ln(\gamma)) + o(\epsilon)$$

( $< 0$ , puisque  $E(\gamma) = 1$  et appliqué l'inégalité de Jensen)

soit  $\zeta = 1 + \frac{\epsilon}{E(\ln(\gamma))} + o(\epsilon)$

## Chapitre 3 :ESTIMATION

Estimation du coefficient de Zipf (coefficient de hiérarchisation) Par la méthode des moindres carrés ordinaire

### 1-Modèle log rang taille

On va chercher à estimer le coefficient de hiérarchisation.

On a la relation entre la taille  $T_i$  d'une ville  $i$  et son rang  $R_i$

$$R_i = A/T_i^\beta$$

En appliquant le log on se ramène à :  $R_i = \ln A - \beta \cdot \ln T_i$

Si  $\beta < 1$  : les villes de grande taille ne suit plus exactement une distribution de loi Zipf

Et l'effet agglomération (la possibilité de réaliser des gains de productivité dans la production grâce à la proximité géographique des entreprises les unes avec les autres) s'intensifie.  
La courbe de la distribution rang taille est strictement convexe

Si  $\beta > 1$  : dans ce cas c'est nombre de villes moyennes qui devient plus important qu'une distribution rang taille conforme à la loi de Zipf

La courbe de la distribution rang taille est strictement concave

### 2-modèle de Lotka

Ce modèle met en relation la taille en fonction du rang et se comporte contrairement au modèle log rang taille avec

$$\ln T_i = \ln C - \gamma \ln R_i$$

Avec  $\gamma$  coefficient de hiérarchisation qui augmente lorsque le poids des grandes villes augment et vice-versa.

$$\gamma = 1/\beta$$

Quand la distribution de la taille des villes suit exactement une loi de Pareto

### **Remarque**

- 3) Plusieurs chercheurs dénoncent un biais assez important des estimateurs du coefficient de hiérarchisation surtout pour les échantillons de petites tailles.

Gabaix et Loannides ont construit un échantillon de ville qui suit parfaitement la loi zipf et ont proposé des estimations de taille n différente.

Ils ont trouvé :

$$\hat{\beta}=0,90 \quad \text{pour } n=20$$

$$\hat{\beta}=0,94 \quad \text{pour } n=100$$

$$\hat{\beta}=0,98 \quad \text{pour } n=500$$

D'où on peut constater que la méthode de MCO a tendance à sous-estimer la valeur de coefficient de hiérarchisation.

- 4) Gabaix et Loannides ont montré aussi qu'il y a une différence importante entre l'écart-type calculé et l'écart-type réel,

$$\text{Pour } n=100 \quad \text{l'écart type calculé} = 0.013 \quad \text{l'écart type réel} = 0.13$$

$$\text{Avec : l'IC par MCO} = [0,97 \quad 1,026 ]$$

$$\text{l'IC réel} = [0,68 \quad 1,20 ]$$

D'où on rejette la loi Zipf.

## **3-L'estimateur de Hill**

### ***Définition :***

Soit  $X_1, \dots, X_n$  un échantillon de variables aléatoire indépendante et identiquement distribuées.

On note  $X_{1,n} \leq \dots \leq X_{n,n}$

Estimateur proposé par Hill (1975) est :

$$\hat{\zeta} = \frac{1}{k_n} \sum_{i=1}^{k_n} i \ln \frac{X_{n-i+1,n}}{X_{n-i,n}} \quad \text{Ou } 1 < k_n < n$$

Certain chercheur préfère utiliser de méthodes semi et non paramétrique pour améliorer le calcul de la pente de distribution avec une queue épaisse



Hill a proposé une méthode semi paramétrique très utilisé dans l'estimation du coefficient de hiérarchisation.

L'estimateur de Hill est celui de la méthode du maximum de vraisemblance, Lorsque la distribution étudiée suit parfaitement une loi de Pareto (et donc la loi De Zipf est vérifiée).

Pour un échantillon de n villes avec des tailles  $T_1 > \dots > T_j > \dots > T_n$

L'estimateur de Hill égal à :

$$\hat{\zeta} = \frac{n-1}{\sum_{j=1}^{n-1} \ln T_j - \ln T_n}$$

L'écart-type de  $\frac{1}{\hat{\zeta}}$

$$\sigma_n\left(\frac{1}{\hat{\zeta}}\right) = \left( \sum_{j=1}^{n-1} j \cdot \frac{(\ln T_j - \ln T_{j+1})^2}{n-1} - \frac{1}{\hat{\zeta}^2} \right)^{1/2} (n-1)^{-1/2}$$

Si  $\frac{1}{\hat{\zeta}} > \sigma_n\left(\frac{1}{\hat{\zeta}}\right)$ , l'écart-type de l'estimation  $\hat{\zeta}$  est égale :

$$\sigma_n(\hat{\zeta}) = \hat{\zeta}^2 \left( \sum_{j=1}^{n-1} j \cdot \frac{(\ln T_j - \ln T_{j+1})^2}{n-1} - \frac{1}{\hat{\zeta}^2} \right)^{1/2} (n-1)^{-1/2}$$

L'estimateur de Hill décrit mieux la partie basse de la distribution rang-taille des villes que le coefficient obtenu par les MCO c'est que réduit le rejet de Zipf

## 4-Méthode des moindres carrés généralisée

Pour améliorer l'estimation du coefficient de hiérarchisation et en travaillant sur le modèle de Lotka plusieurs chercheurs (Osada, Morimune, Sato, Nishiyama) proposition une correction de l'estimateur en le multipliant par une constante qui ne dépend que du nombre de ville dans l'échantillon

$$\bar{\zeta} = \frac{n \sum (\ln i)^2 - \sum (\ln i)^2}{n \sum \ln\left(\frac{1}{n} + \dots + \frac{1}{i} - 1\right)} \hat{\zeta}$$

En valeur absolue cette constante est inférieure à 1 c'est qui réduit le biais de l'estimation du coefficient par MCO

Lorsque les erreurs sont indépendantes et leur distribution suit une loi normale le MCG Est identique à celle du Maximum de vraisemblance

MCG appliquée à la distribution rang-taille des villes :

$$Y = [\ln T_1, \ln T_2, \dots, \ln T_n]$$

$$X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \ln 1 & \ln 2 & \dots & \ln n \end{bmatrix}$$

$$\Omega = \text{var}(Y)$$

Les estimateurs de  $\zeta$  et  $A$  (constante qui dépend de la queue de distribution) par MCG sont donnés par :

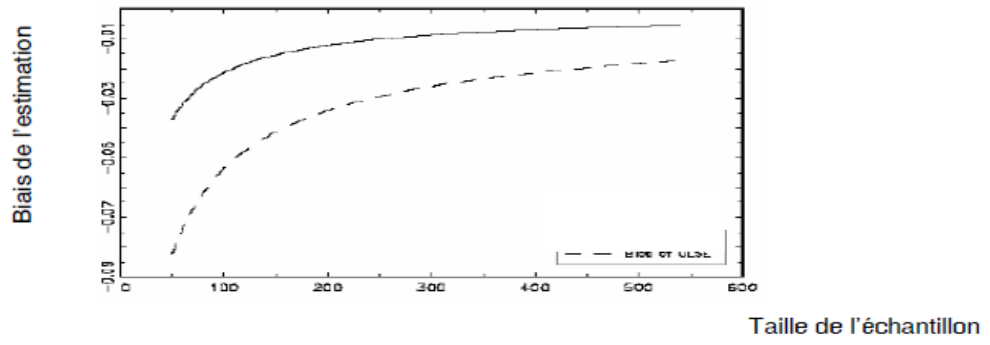
$$\begin{bmatrix} \hat{A} \\ \hat{\zeta} \end{bmatrix} = (X \Omega^{-1} X)^{-1} X \Omega^{-1} Y$$

Et leur variance  $V \begin{bmatrix} \hat{A} \\ \hat{\zeta} \end{bmatrix} = (X \Omega^{-1} X)^{-1}$

La méthode des MCG réduit la variance du coefficient  $\hat{\zeta}$

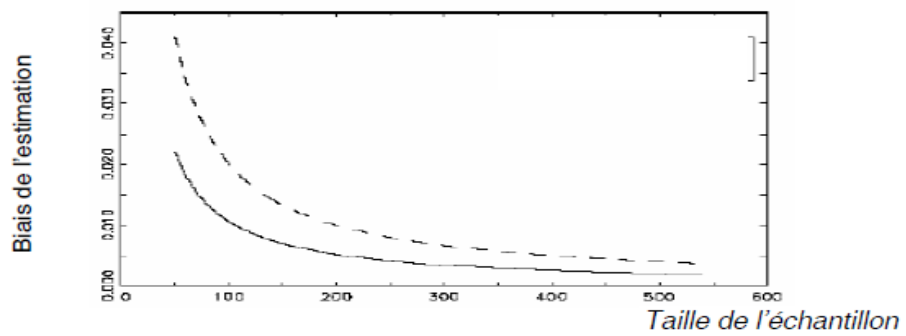
**Remarque :**

Si les erreurs sont indépendantes et suivent la loi normal l'estimation du coefficient par les MCG est égale à l'inverse de l'estimation de Hill



Source : Nishiyama, Osada et Morimune (2004)

Figure 4 : Comparaison du biais de l'estimation du coefficient de hiérarchisation ( $\gamma$ ) par la méthode des MCO et des MCG



Source : Nishiyama, Osada et Morimune (2004)

Figure 5 : Comparaison du biais de l'estimation de la variance du coefficient de hiérarchisation par la méthode des MCO et des MCG

Taille échantillon	20	50	100	200	500
$\beta$ (MCO)					
Moyenne	0,9018	0,9225	0,9437	0,9613	0,9782
Ecart-type réel	0,2804	0,1813	0,1324	0,0962	0,0624
Moyenne des e-types calculés	0,0485	0,0235	0,0137	0,0078	0,0036
IC à 5%	[0,52 ; 1,42]	[0,65 ; 1,24]	[0,74 ; 1,17]	[0,81 ; 1,13]	[0,88 ; 1,08]
$\beta_{cor}$ (MCO corr)					
Moyenne	1,0504	1,0101	1,0016	0,9986	0,9984
Ecart-type réel	0,3203	0,1944	0,1378	0,0983	0,0630
Moyenne des e-types calculés	0,0588	0,0263	0,0147	0,0081	0,0038
Ecart type estimé	0,3162	0,2000	0,1414	0,1001	0,0632
IC à 5%	[0,61 ; 1,65]	[0,72 ; 1,35]	[0,79 ; 1,24]	[0,84 ; 1,16]	[0,90 ; 1,10]
$\beta$ (Hill)					
Moyenne	1,0586	1,0219	1,0112	1,0054	1,0021
Ecart-type réel	0,2570	0,1482	0,1019	0,0708	0,0449
Moyenne des e-types calculés	0,2324	0,1433	0,1008	0,0710	0,0448
IC à 5%	[0,71 ; 1,53]	[0,80 ; 1,29]	[0,85 ; 1,19]	[0,89 ; 1,13]	[0,93 ; 1,08]
$\gamma$ (MCO)					
Moyenne	1,1395	1,0823	1,0529	1,0339	1,0180
Ecart-type réel	0,3321	0,2013	0,1403	0,0991	0,0629
Moyenne des e-types calculés	0,0646	0,0291	0,0159	0,0086	0,0038
IC à 5%	[0,67 ; 1,75]	[0,78 ; 1,44]	[0,84 ; 1,30]	[0,88 ; 1,21]	[0,92 ; 1,13]
$\gamma_{cor}$ ( $\gamma$ corrigé)					
Moyenne	0,9987	1,0001	0,9995	0,9998	0,9997
Ecart-type réel	0,2910	0,1860	0,1332	0,0958	0,0618
Moyenne des e-types calculés	0,0566	0,0269	0,0151	0,0083	0,0037
IC à 5%	[0,58 ; 1,53]	[0,72 ; 1,33]	[0,80 ; 1,23]	[0,85 ; 1,17]	[0,90 ; 1,11]
$\gamma$ (MCG)					
Moyenne	1,1025	1,0523	1,0298	1,0169	1,0078
Ecart-type réel	0,2657	0,1543	0,1049	0,0723	0,0453
Moyenne des e-types calculés	0,3655	0,2708	0,2104	0,1609	0,1109
IC à 5%	[0,71 ; 1,58]	[0,81 ; 1,32]	[0,86 ; 1,21]	[0,90 ; 1,14]	[0,93 ; 1,08]

Tableau 1 : Estimations du coefficient de Pareto –Résultats de la simulation Monte Carlo (20000 estimations)

# Conclusion

Bibliographie :

\*ZIPF'S LAW FOR CITIES: AN EXPLANATION      XAVIER GABAIX