Zipfp’s Law and the Growth of Cities

Xavier Gabaix

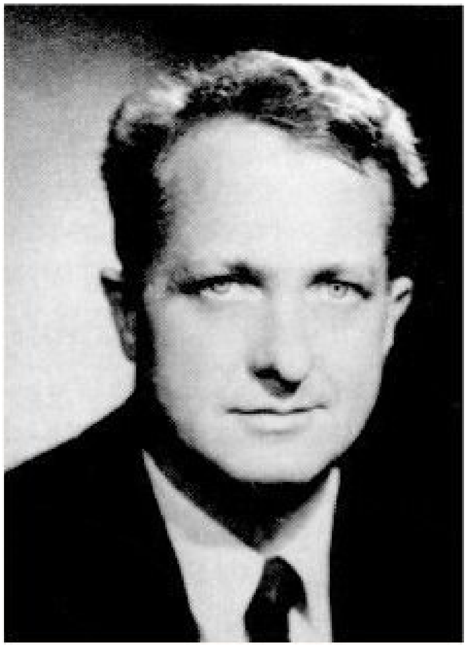
The American Economic Review Vol. 89, No. 2, Papers and Proceeding of the One Hundred Eleven Annual Meeting of the American Economic Association (May, 1999), pp. 129-132

Zipf's Law for Cities: An Explanation

Xavier Gabaix

The Quarterly Journal of Economics, Vol. 114, No. 3 (Aug., 1999), pp. 739-767

Закон Ципфа



George Kingsley Zipf
1902-1950

- Закон Ципфа — эмпирическая закономерность распределения частоты слов языка: если все слова (или просто достаточно длинного текста) упорядочить по убыванию частоты их использования, то частота n-го слова в таком списке окажется приблизительно обратно пропорциональной его порядковому номеру n (рангу этого слова).
- Например второе по используемости слово встречается примерно в два раза реже, чем первое, третье — в три раза реже, чем первое, и т. д.

Основные разделы

- Введение
- Закон Ципфа как стационарное распределение закона Gibrat
- Простые модели роста городов
- Модель случайного роста
- Эмпирические исследования отклонения от закона Ципфа

Закон Ципфа для численности населения городов

$$R = a \cdot S^{\xi}$$

- где R – ранг города по размеру (№ 1 – самый большой город), S – население городов, ξ – наклон прямой, близкий к -1.

Вероятность того, что размер города превзойдет величину S пропорциональна $1/S$:

$$P(Size > S) = a/S^{\xi}, \ \ \xi \approx 1$$

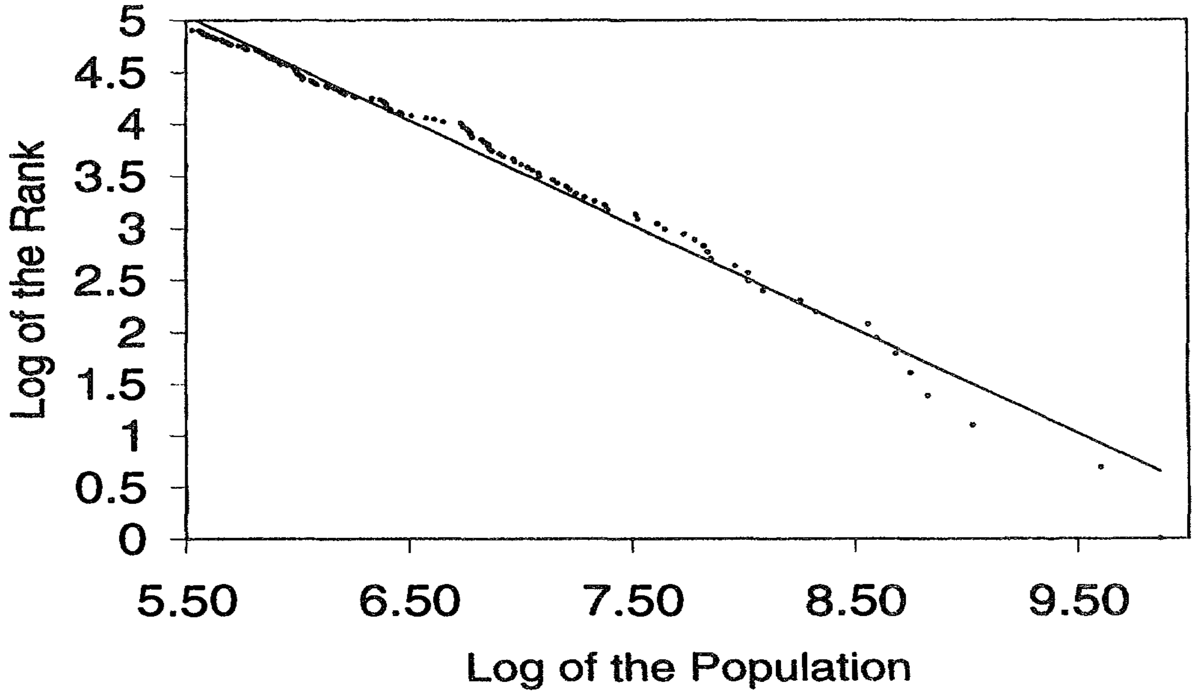


FIGURE 1. A PLOT OF LOG RANK (BY POPULATION) AGAINST LOG POPULATION FOR U.S. CITIES, 1991

$$Ln(Rank)=\beta 0+\beta 1 ln(Size)$$

$$Ln(Rank)=10.53-1.005\ ln(Size)$$

$$(0.01)$$

$$R^2=0,98$$

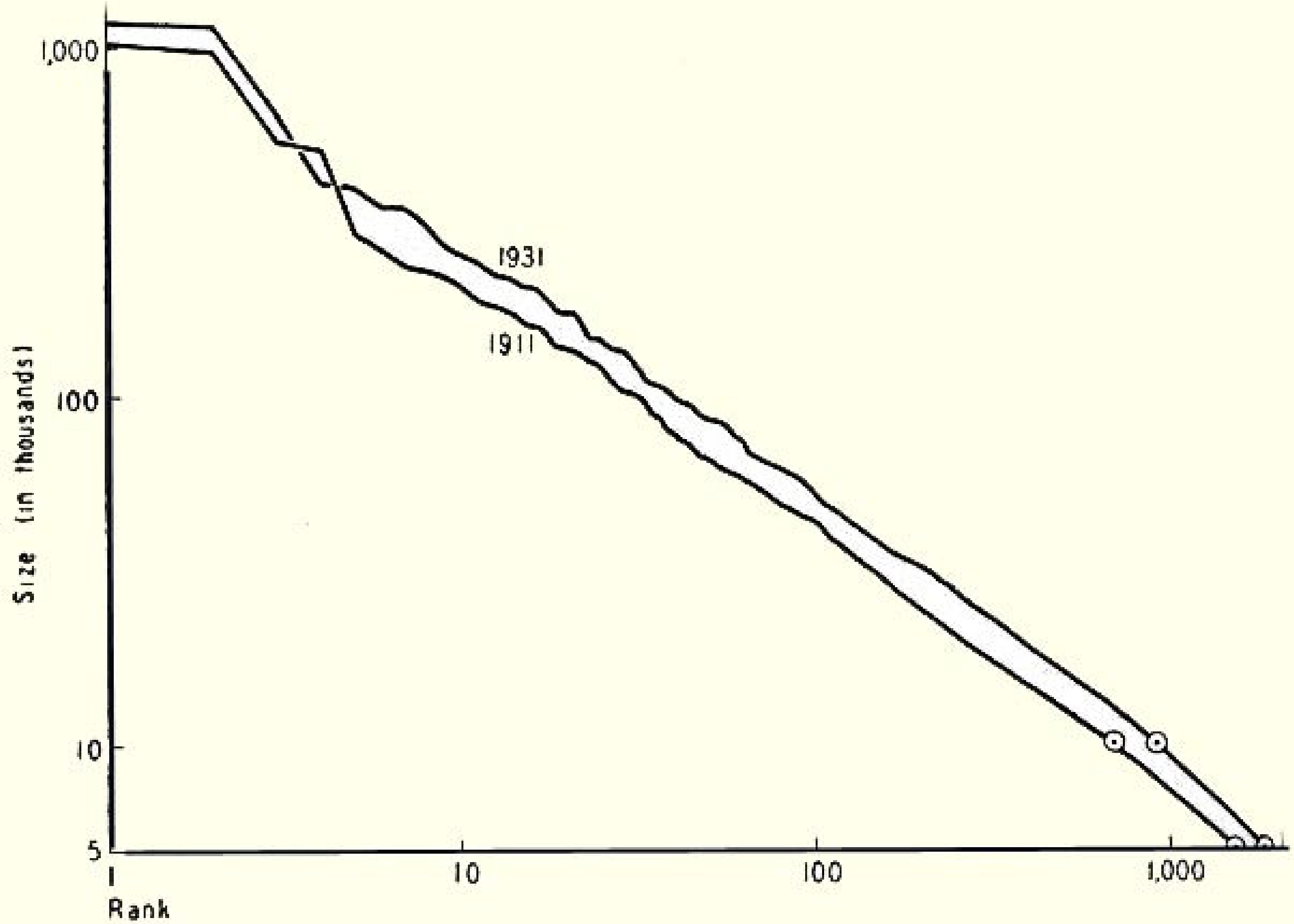


Fig. 10–7. India in 1911 and 1931. Communities ranked in the decreasing order of population size.

China's different economic zones

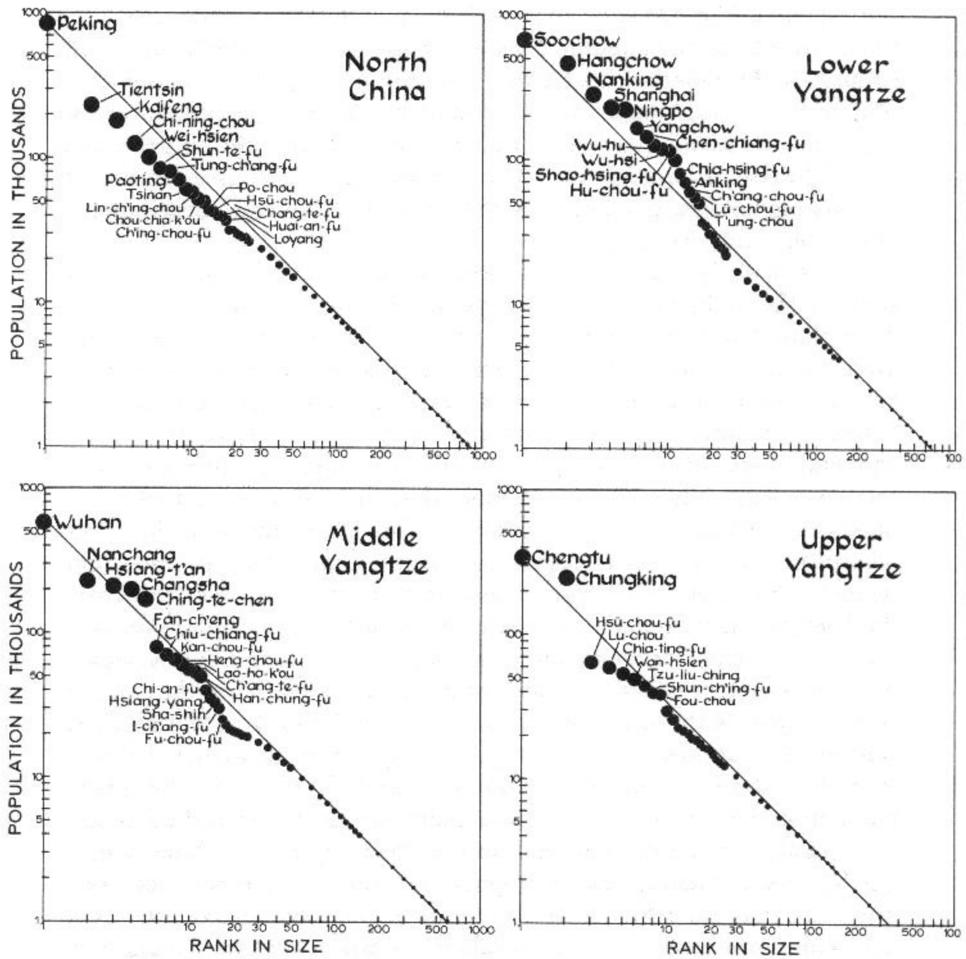


FIG. 1. RANK-SIZE DISTRIBUTION OF URBAN CENTRAL PLACES, BY REGION, 1843

for other periods in U. S. history [Dobkins and Ioannides 1998a; Krugman 1996a, p. 41; 1996b; Zipf 1949, p. 420],
for most countries in the modern period [Rosen and Resnick 1980]
and for even India in 1911 [Zipf 1949, p. 432],
and China in the mid-nineteenth century [Rozman 1990, p. 68].

Основные положения X.Gabaix

- Если города будут расти случайным образом с одинаковым средним значением темпов роста и одинаковым среднеквадратическим отклонением, то в пределе распределение будет сходиться к закону Ципфа.
- Независимо от особенностей роста городов, их экономической роли и т.д., как только выполняются предположения (по крайней мере в определенном диапазоне) закона Gibrat, распределение городов по численности населения будет сходиться к закону Ципфа.

Закон Gibrat

- Закон R. Gibrat (правило пропорционального роста): размер фирмы и ее темпы роста являются независимыми

- Распределение $\gamma_{t+1}=S_{t+1}/S_t$ не зависит от размера городов S_t

Закон Ципфа как стационарное распределение при предположениях закона Gibrat

- Пусть имеется N городов
- Начальное распределение городов произвольно
- Средний темп роста городов не зависит от времени и темп роста города в будущем не зависит от численности населения сегодня и прошлом.
- наблюдаются скачки роста (шоки)
- Перейдем к нормализованным величинам численности городов S^i ($i = 1, \ldots, N$).
- S^i – население города i, деленное на общее городское население всех городов $\sum_{i=1}^{N} S_t^i = 1$

Закон Ципфа
как стационарное распределение

- Пусть хотя бы в верхней области распределения (для крупных городов) процесс имеет вид $S^i_{t+1} = \gamma_{t+1} S^i_t$,
- где γ_t - независимые одинаково распределенные факторы роста городов, распределенные по $f(\gamma)$.
- $(\gamma_{t+1} - 1)$ – темп роста города i.
- Средняя нормированная величина
- остается постоянной ($\gamma = 1$), что означает, что $E[\gamma] = 1$ или

$$\int_0^\infty \gamma f(\gamma)\, d\gamma = 1.$$

Закон Ципфа
как стационарное распределение

- Введем функцию $G_t(S) = P(S_t > S)$ - *функция распределения в верхней области* для величины города в момент t.
- Уравнение динамики для G_t :

$$
\begin{aligned}
G_{t+1}(S) &= P(S_{t+1} > S) = P(\gamma_{t+1} S_t > S) = E[1_{S_t > S/\gamma_{t+1}}] \\
&= E[E[1_{S_t > S/\gamma_{t+1}} | \gamma_{t+1}]] = E[G_t(S/\gamma_{t+1})] \\
&= \int_0^\infty G_t\left(\frac{S}{\gamma}\right) f(\gamma)\, d\gamma.
\end{aligned}
$$

Закон Ципфа
как стационарное распределение

- Предположим, что существует стационарное распределение $G_t = G$
- Тогда должно выполняться равенство

$$G(S) = \int_0^\infty G\left(\frac{S}{\gamma}\right) f(\gamma)\, d\gamma.$$

- Распределение $G(S) = a/S$ обладает данным свойством

Закон Ципфа как стационарное распределение (пример)

- Пусть в момент времени t город А удваивается, а В – уменьшается вдвое.
- Тогда, вероятность удвоения 1/3,
- вероятность сокращения в два раза 2/3 (ожидаемый рост $1/3*2+2/3*1/2=1$)
- число $2S$ размеров городов должно быть вдвое меньше S-размерных городов, а количество $S/2$- размерных городов должно быть вдвое больше

Модель роста городов

- Перекрывающиеся поколения агентов

- N^0_{it} - старые агенты
- δ - вероятность смерти агента
- N^y_{it} - новые агенты
- $N^0_{it} >> N^y_{it}$
- w_{ij} - заработная плата в городе i

Модель роста городов

- Производственная функция

$$F(N_i^0, N_i^y) = A_i \pi(N_i^0) N_i^0 f(\frac{N_i^y}{N_i^0}),$$

- A_i - общий выпуск в городе,
- f - возрастающая вогнутая функция.
- $\pi(N^0{}_i)$ - технологический спиловер.
- Номинальная заработная плата новых агентов

$$w_i^y = A_i \pi(N_i^0) f'(\frac{N_i^y}{N_i^0}).$$

Модель роста городов

- Общая полезность: $u_t = a_{ij}\lambda(N^o{}_i)w_{ij}$
- $a_{ij}\lambda(N^o{}_i)$ – общий уровень amenities
- $\lambda(N^o{}_i)$ - преимущества и недостатки от наличия большого количества людей в своем городе (социальные контакты, пробки, высокая аренда)
- a_{ij} – уровень amenities (качество образования, полиция)

Модель роста городов

- Функция полезности: $u_t = a_{it} A_{it} \lambda(N_{it}^0) \pi(N_{it}^0) f'(\frac{N_{it}^y}{N_{it}^0}),$
- Равновесие

$$\frac{u_t}{a_{it} f'(\frac{N_{it}^y}{N_{it}^0})} = A_{it} \lambda(N_{it}^0) \pi(N_{it}^0).$$

Два направления анализа закона Ципфа:

- Первый - закон роста Gibrat и CRS;
- второй - закон Gibrat с неограниченными внешними эффектами и и различиями в производительности.

Модель роста городов

1. Диапазон изменений $\lambda(N_{it}^0)\pi(N_{it}^0)$ ограничен.

Т.е. асимптотически отдача от масштаба города постоянна.

Тогда выпуск A_{it} имеет ограниченную дисперсию между городами.

Процесс приводит к закону Gibrat: случайные шоки удобств a_{it} и выпуска A_{it} вызывают шоки темпов роста населения, которые не зависят от размера города для больших N^0_{it} поскольку экстерналии $\lambda\pi$ достигли своих пределов.

Модель роста городов

2. Диапазон изменений $\lambda(N_{it}^{0})\pi(N_{it}^{0})$ не ограничен.

Для больших городов возможно существование экстремально плохих удобств или социальных экстерналий. В пределе различия в условиях жизни между большими и малыми городами бесконечны.

В этом случае *зарплаты* должны иметь неограниченную дисперсию, чтобы компенсировать плохие жизненные условия (за счет большей производительности).

Модель роста городов

- Перекрывающиеся поколения агентов

- N^{0}_{it} - старые агенты
- δ - вероятность смерти агента
- N^{y}_{it} - новые агенты
- $N^{0}_{it} >> N^{y}_{it}$
- w_{ij} - заработная плата в городе i

Модель роста городов

- Если город i имеет уровень общественных благ a_{it}, не зависящий от размера города, то житель этого города с уровнем потребления c получит услуги в количестве $u(c) = a_{it}c$.
- a_{it} - независимые и одинаково распределенные случайные величины.
- Полезность $a_{it}w_{it} = u_t$,
- $\max u_t = \max_i a_{i\,t} w_{i\,t}$

Модель роста городов

- Производственная функция $F(N_i^o, N_i^y) = N_i^o f\,(N_i^y / N_i^o\,)$.
- Заработки молодежи $w_i = w_i^y = f\,'(N_i^y / N_i^o)$.
- Полезность $a_{it} w_{it} = u_t,$

$$N_i^y = N_i^o f\,'^{-1}(u_t / a_{it}).$$

- Увеличение численности города

$$\Delta N_{it} = N^y{}_{it} - \delta\, N^0{}_i$$

- темп роста городов: $\gamma_{ij} = \dfrac{N_{it}^y}{N_{it}^0} - \delta.$

$$\gamma_{it} := \Delta N_{it} / N_{it} = f\,'^{-1}(u_t / a_{it}) - \delta.$$

Предложения

$$\gamma^{it} = \overline{\gamma} + \gamma^{it}_{\text{policy}} + \gamma^{it}_{\text{region}} + \gamma^{it}_{\text{industries}},$$

$$\sigma^2(S) = \sigma^2_{\text{policy}}(S) + \sigma^2_{\text{region}}(S) + \sigma^2_{\text{industries}}(S).$$

$$\sigma^2_{\text{policy}}(S) = \sigma^2_{\text{policy}}.\ \gamma^{it}_{\text{region}}$$

$$\sigma^2_{\text{region}}(S) = \sigma^2_{\text{region}}.$$

$$\sigma^2_{\text{industries}}(S) = \sigma^2_{\text{industries}}/S.$$

$$\sigma^2(S) = [\sigma^2_{\text{policy}} + \sigma^2_{\text{region}}] + \sigma^2_{\text{industries}}/S.$$

$$(\sigma^2_{\text{policy}} + \sigma^2_{\text{region}} \gg \sigma^2_{\text{industries}}/S),$$

Модели случайного роста

- $dP_{it}/P_{it} = gdt + \sigma dB_{it}$
- где B_{it} - броуновское движение, для некоторых γ, которое зависит от дискретных параметров модели (u_t, a_{it}, δ).
- $S_{it} = P_{it}$ / (общая численность городского населения ожидаемая на момент t)

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

- ожидаемый рост нормированного размера

$$\mu = \gamma(S) - \bar{\gamma}$$

Модели случайного роста

- Есть ли механизмы, предотвращающие исчезновение малых городов?
- Известно, что геометрическое броуновское движение имеет логнормальное распределение и большинство городов будут иметь бесконечно малые размеры.
- Поэтому вводится S_{min} - нижняя граница размера города, для которой
- $dS_t = S_t max(\mu dt + \sigma dB_t ,0)$

Модели случайного роста Утверждение 1.

- Предположим, что нормированные размеры S следуют геометрическому броуновскому движению.
- Процесс $dS_t/S_t=\mu dt+\sigma dB_t$ и
- где B_t - броуновское движение,
- $\mu<0$ - отрицательный дрейф
- $dS_t=S_t max(\mu dt+\sigma dB_t ,0)$ для $S_t<=Smin$,
- *Smin* - барьер процесса, т. е. минимальный нормированный размер города.
- Тогда распределение сходится к распределению Ципфа с показателем
$$\xi = \frac{1}{1-\frac{S_{\min}}{\bar{S}}}$$
- Где $\bar{S}$ - средний размер города.
- Таким образом, в пределе, когда минимальный допустимый размер *Smin* города стремится к 0, показатель ξ стремится к 1.

Модели случайного роста Утверждение 2.

- Предположим, что страна состоит из R регионов, и что в каждом регионе в условиях предложения 1 выполнены, так что закон Ципфа с показателем ξ выполняется в каждом регионе. В частности, рост процессов (в верхней части хвоста) идентичны в каждом регионе, но не обязательно во всех.
- Тогда асимптотическое распределение национальных городов существует и удовлетворяет закону Ципфа, с тем же показателем ξ.

Модели случайного роста Утверждение 3.

- Предположим, что число городов увеличивается со скоростью v и $v \le \gamma$. Тогда существует стационарное распределение, удовлетворяющее закону Ципфа с показателем 1 в верхнем хвосте.
- При $v > \gamma$, в непрерывном случае, стационарное распределение имеет показатель ξ, который является положительным корнем
- $\xi^2-(1-\gamma/\sigma^2)\xi-2v/\sigma^2= 0$ и имеет значение больше, чем 1.

Модели случайного роста
Утверждение 4. Правило ранг-размер

- Упорядочим города по размеру ($S_{(1)}$>= $S_{(2)}$>= ...), и предположим, что стационарное распределение удовлетворяет закону Ципфа с показателем 1.
- Тогда для *i*<*j*, среднее $S_{(j)}/S_{(i)}$ есть *i/j* , стандартное отклонение
- $$\sqrt{\frac{\left(1-{}^{i}/_{j}\right){}^{i}/_{j}}{j+1}}$$
- Медиана $B^{-1}(1/2,\ i,\ j-i)$ обратная *1/2* регуляризованная бета-функция.
- Ожидаемая средняя $S_{(i)}/\ S_{(j)}$ является $(j-1)/(i-1)$.
- Эти результат справедлив и для конечного числа городов (то есть, он является точным, не асимптотическим).

Модели случайного роста
Утверждение 4. Правило ранг-размер

TABLE I

STATISTICS OF THE SIZE RATIO $S_{(j)}/S_{(i)}$ BETWEEN THE jTH LARGEST CITY AND THE iTH LARGEST CITY ($i < j$)

	$S_{(j)}/S_{(i)}$	$S_{(2)}/S_{(1)}$	$S_{(3)}/S_{(1)}$	$S_{(3)}/S_{(2)}$	$S_{(10)}/S_{(5)}$	$S_{(100)}/S_{(10)}$
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[a]	$B^{-1}(1/2, i, j-i)$	.5	.2929	.7071	.5	.0973

a. $B^{-1}(1/2, i, j-i)$ denotes the inverse at 1/2 of the regularized beta function with parameters i and $i-j$. (See Appendix 2 for a definition.)

Модели роста городов

- *Модели роста Steindl* [1965, 1968]. Новые города рождаются со скоростью v, а существующие города растут со скоростью γ.
- Результат - распределение новых городов - степенной закон с показателем $\xi = v / \gamma$.
- Однако, это не дает степень 1. Она обеспечивает его только тем, что исторически $v = \gamma$. Это совершенно неправдоподобно эмпирически, особенно для зрелых городских систем, для которых $v < \gamma$.

Модели роста городов

- *Модели стохастического роста*. Наиболее успешная модель стохастического роста Саймона [1955]. В этой модели новый мигранты приезжают в каждом периоде, и с вероятностью π они образуют новый город, а с вероятностью $1 - \pi$ они уедут в существующие города. Вероятность того, что они выберут данный город пропорционально его населению.
- Тогда эта модель генерирует степенной закон с показателем
- $\xi = 1 / (1 - \pi)$. Таким образом, показатель 1 имеет очень естественное объяснение: вероятность π новых городов мало. Это кажется вполне понятным. Однако такой подход имеет как минимум два существенных недостатка.
- Во-первых, Кругман [1996a, стр. 96-97, 1996b], показывает, существует вырождение при $\pi = 0$. Чтобы получить показатель $\xi = 1$, необходимо, чтобы π было очень близко к 0, а затем процесс сходится бесконечно медленно. Второй проблемой является то, что имеет важное значение для этой модели, что темпы роста количества городов должно быть больше, чем темпы роста населения существующих городов

Эмпирические исследования отклонения от закона Ципфа

- 1. Столицы
- 2. При экстраполяции с верхнего хвоста до середины в соответствии с законом Ципфа - слишком мало средних и малых городов (100000 жителей или менее в Соединенных Штатах) - степень Ципфа ξ меньше, чем 1 [Dobkins и Ioannidis 1998a].
- Объяснение: малые города имеют более высокую дисперсию темпов роста, чем большие

Замечания

- 1. Нормированный размер $S_i = Pop_i / \sum Pop_j$
- 2. В непрерывном случае - как связаны g и μ
- 3. Эмпирические исследования

- Благодарю за внимание !