Modulated Branching Processes, Origins of Power Laws and Queueing Duality

Predrag R. Jelenković Jian Tan

Dept. of Electrical Engineering
Columbia University , NY 10027, USA
{predrag, jiantan}@ee.columbia.edu

October 2007

Outline

100+ years of repeated observations of power laws

Socioeconomic area

- Incomes, Pareto (1897)
- Population of cities Arrherbach (1913) & Zipf (1949)

Biological area

- Species-area relationship, Arrhenius (1921)
- Gene family sizes Huynen & Nimwegen (1998)

Technological area: Internet

- Ethernet LAN traffic Leland, Willinger et al. (1993), Scenes in MPEG video streams Jelenković et al. (1997), WWW traffic Crovella & Bestavros (1997)
- Page requests Cunha et al. (1995), pages and visitors per Web site Adamic & Huberman (1999, 2000)

These are just a few examples, there are many, many more. . .

100+ years of repeated observations of power laws

Socioeconomic area

- Incomes, Pareto (1897)
- Population of cities Arrherbach (1913) & Zipf (1949)

Biological area

- Species-area relationship, Arrhenius (1921)
- Gene family sizes Huynen & Nimwegen (1998)

Technological area: Internet

- Ethernet LAN traffic Leland, Willinger et al. (1993), Scenes in MPEG video streams Jelenković et al. (1997), WWW traffic Crovella & Bestavros (1997)
- Page requests Cunha et al. (1995), pages and visitors per Web site Adamic & Huberman (1999, 2000)

These are just a few examples, there are many, many more. . .

100+ years of repeated observations of power laws

Socioeconomic area

- Incomes, Pareto (1897)
- Population of cities Arrherbach (1913) & Zipf (1949)

Biological area

- Species-area relationship, Arrhenius (1921)
- Gene family sizes Huynen & Nimwegen (1998)

Technological area: Internet

- Ethernet LAN traffic Leland, Willinger et al. (1993), Scenes in MPEG video streams Jelenković et al. (1997), WWW traffic Crovella & Bestavros (1997)
- Page requests Cunha et al. (1995), pages and visitors per Web site Adamic & Huberman (1999, 2000)

These are just a few examples, there are many, many more. . .

100+ years of repeated observations of power laws

Socioeconomic area

- Incomes, Pareto (1897)
- Population of cities Arrherbach (1913) & Zipf (1949)

Biological area

- Species-area relationship, Arrhenius (1921)
- Gene family sizes Huynen & Nimwegen (1998)

Technological area: Internet

- Ethernet LAN traffic Leland, Willinger et al. (1993), Scenes in MPEG video streams Jelenković et al. (1997), WWW traffic Crovella & Bestavros (1997)
- Page requests Cunha et al. (1995), pages and visitors per Web site Adamic & Huberman (1999, 2000)

These are just a few examples, there are many, many more...

Are these observations merely a big coincidence?

Are there universal mathematical laws governing these phenomena?

Power law distribution

Roughly speaking, a random variable X has a power law tail if

there exists $\alpha > 0$, such that

$$\mathbb{P}[X > x] \sim \frac{H}{x^{\alpha}},$$

or, more generally,

$$\frac{\log \mathbb{P}[X > x]}{\log x} \to -\alpha$$

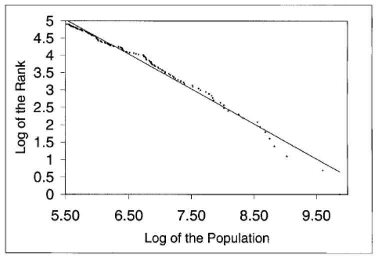


Figure: Size versus rank of the 135 largest U.S. Metropolitan Areas in 1991 [cited from Gabaix (1999)]

Are these observations merely a big coincidence?

Are there universal mathematical laws governing these phenomena?

Power law distribution

Roughly speaking, a random variable X has a power law tail if

there exists $\alpha > 0$, such that

$$\mathbb{P}[X > x] \sim \frac{H}{x^{\alpha}},$$

or, more generally,

$$\frac{\log \mathbb{P}[X > x]}{\log x} \to -\alpha$$

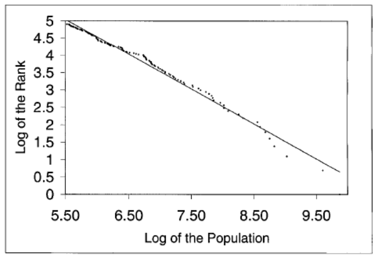


Figure: Size versus rank of the 135 largest U.S. Metropolitan Areas in 1991 [cited from Gabaix (1999)]

Power law is memoryless

How is that possible?

In the multiplicative world

- If Y has a power law distribution ($x > 1, \alpha > 0$)

$$\mathbb{P}[Y > x] = \frac{H}{x^\alpha}$$

- then, for $x, y > 1$

$$\mathbb{P}[Y > xy | Y > y] = \mathbb{P}[Y > x]$$

- Should the power law distributions play the same role in the multiplicative world as do the exponential ones in the additive?

Power law is memoryless

How is that possible?

In the multiplicative world

- If Y has a power law distribution ($x > 1, \alpha > 0$)

$$\mathbb{P}[Y > x] = \frac{H}{x^\alpha}$$

- then, for $x, y > 1$

$$\mathbb{P}[Y > xy | Y > y] = \mathbb{P}[Y > x]$$

- Should the power law distributions play the same role in the multiplicative world as do the exponential ones in the additive?

Power law is memoryless

How is that possible?

In the multiplicative world

- If Y has a power law distribution ($x > 1, \alpha > 0$)

$$\mathbb{P}[Y > x] = \frac{H}{x^{\alpha}}$$

- then, for $x, y > 1$

$$\mathbb{P}[Y > xy | Y > y] = \mathbb{P}[Y > x]$$

- Should the power law distributions play the same role in the multiplicative world as do the exponential ones in the additive?

Power law is memoryless

How is that possible?

In the multiplicative world

- If Y has a power law distribution ($x > 1, \alpha > 0$)

$$\mathbb{P}[Y > x] = \frac{H}{x^\alpha}$$

- then, for $x, y > 1$

$$\mathbb{P}[Y > xy | Y > y] = \mathbb{P}[Y > x]$$

- Should the power law distributions play the same role in the multiplicative world as do the exponential ones in the additive?

Outline

Motivation for our model

Features of many existing power law observations

1. Replication of independent components
2. Replication is modulated by dynamic environments, causing periods of expansions and contractions , examples $\cdots$
3. Have reflective lower barriers or porous/absorbing lower barriers, examples $\cdots$

We believe that this model provides a basic structure that explains the origins of power laws in proportional growth situations

Motivation for our model

Features of many existing power law observations

To capture all these features,
we propose modulated branching process
with a reflecting or absorbing lower barrier

We believe that this model provides a basic structure that explains the origins of power laws in proportional growth situations

Outline

Reflected Modulated Branching Processes (RMBP)

- $\{J_n \in \mathbb{N}\}_{n>-\infty}$ models the environment dynamics
  - stationary and ergodic
- $\{B_n^i(j) \in \mathbb{N}\}$ - number of children of object i at time n when $J_n = j$
  - i.i.d. for fixed j and mutually independent, and independent of $\{J_n\}$
- l is the lower barrier

Modulated Branching Processes

For $Z_0 \in \mathbb{N}$, define

$$Z_{n+1} = \sum_{i=1}^{Z_n} B_n^i(J_n)$$

Reflected Modulated Branching Processes (RMBP)

For $l, \Lambda_0 \in \mathbb{N}$,

$$\Lambda_{n+1} = \max\left(\sum_{i=1}^{\Lambda_n} B_n^i(J_n),\ l\right)$$

Reflected Modulated Branching Processes (RMBP)

- $\{J_n \in \mathbb{N}\}_{n>-\infty}$ models the environment dynamics
  - stationary and ergodic
- $\{B_n^i(j) \in \mathbb{N}\}$ - number of children of object i at time n when $J_n = j$
  - i.i.d. for fixed j and mutually independent, and independent of $\{J_n\}$
- l is the lower barrier

Modulated Branching Processes

For $Z_0 \in \mathbb{N}$, define

$$Z_{n+1} = \sum_{i=1}^{Z_n} B_n^i(J_n)$$

Reflected Modulated Branching Processes (RMBP)

For $l, \Lambda_0 \in \mathbb{N}$,

$$\Lambda_{n+1} = \max\left(\sum_{i=1}^{\Lambda_n} B_n^i(J_n),\ l\right)$$

Reflected Modulated Branching Processes (RMBP)

- $\{J_n \in \mathbb{N}\}_{n>-\infty}$ models the environment dynamics
  - stationary and ergodic
- $\{B_n^i(j) \in \mathbb{N}\}$ - number of children of object i at time n when $J_n = j$
  - i.i.d. for fixed j and mutually independent, and independent of $\{J_n\}$
- l is the lower barrier

Modulated Branching Processes

For $Z_0 \in \mathbb{N}$, define

$$Z_{n+1} = \sum_{i=1}^{Z_n} B_n^i(J_n)$$

Reflected Modulated Branching Processes (RMBP)

For $l, \Lambda_0 \in \mathbb{N}$,

$$\Lambda_{n+1} = \max\left(\sum_{i=1}^{\Lambda_n} B_n^i(J_n),\ l\right)$$

Reflected Modulated Branching Processes (RMBP)

- $\{J_n \in \mathbb{N}\}_{n>-\infty}$ models the environment dynamics
  - stationary and ergodic
- $\{B_n^i(j) \in \mathbb{N}\}$ - number of children of object i at time n when $J_n = j$
  - i.i.d. for fixed j and mutually independent, and independent of $\{J_n\}$
- l is the lower barrier

Modulated Branching Processes

For $Z_0 \in \mathbb{N}$, define

$$Z_{n+1} = \sum_{i=1}^{Z_n} B_n^i(J_n)$$

Reflected Modulated Branching Processes (RMBP)

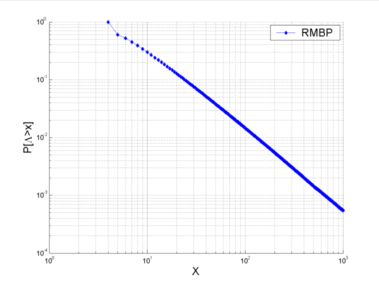
For $l, \Lambda_0 \in \mathbb{N}$,

$$\Lambda_{n+1} = \max\left(\sum_{i=1}^{\Lambda_n} B_n^i(J_n),\ l \right)$$

Reflected Modulated Branching Processes (RMBP)

Example

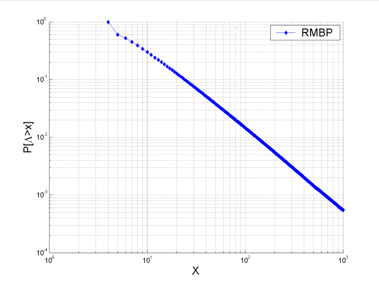
- $\{J_n\}_{n\geq 1}$ - Bernoulli with $\mathbb{P}[J_n = 1] = 0.4$
- $\{B_n^i(1)\}_{i\geq 1} \sim$ Poisson(1.5)
- $\{B_n^i(0)\}_{i\geq 1} \sim$ Poisson(0.6)
- $l = 4$



Even when $\{J_n\}$ is i.i.d., there is no known method to analyze RMBP

Example

- $\{J_n\}_{n\geq 1}$ - Bernoulli with $\mathbb{P}[J_n = 1] = 0.4$
- $\{B_n^i(1)\}_{i\geq 1} \sim$ Poisson(1.5)
- $\{B_n^i(0)\}_{i\geq 1} \sim$ Poisson(0.6)
- $l = 4$



- Where is the difficulty? (dependency $\cdots$)
- Need a new technique...

Outline

Before stating the results

Notation

$\{J_n\}$ - modulating process
$\mu(j) \triangleq \mathbb{E}[B_n^i(j)]$ - replication rate when in state j
$\Pi_n = \prod_{i=-n}^{-1} \mu(J_i),\ n \geq 1,\ \Pi_0 = I$ and $M = \sup_{n\geq 0} \Pi_n$

Polynomial Gärtner-Ellis conditions

1. $n^{-1} \log \mathbb{E}[(\Pi_n)^\alpha] \to \Psi(\alpha)$ as $n \to \infty$ for $\mid \alpha - \alpha^* \mid < \varepsilon^*$
2. Ψ is finite in a neighborhood of α^* and differentiable at α^* with $\Psi(\alpha^*) = 0,\ \Psi'(\alpha^*) > 0$

- When $\{J_n\}$ is i.i.d., $n^{-1} \log \mathbb{E}[(\Pi_n)^\alpha] = \Psi(\alpha) = \log \mathbb{E}[\mu(J_{-1})^\alpha]$
- Condition 1) allows for dependency in $\{J_n\} \cdots$

Before stating the results

Notation

$\{J_n\}$ - modulating process
$\mu(j) \triangleq \mathbb{E}[B_n^i(j)]$ - replication rate when in state j
$\Pi_n = \prod_{i=-n}^{-1} \mu(J_i),\ n \geq 1,\ \Pi_0 = I$ and $M = \sup_{n\geq 0} \Pi_n$

Polynomial Gärtner-Ellis conditions

1. $n^{-1} \log \mathbb{E}[(\Pi_n)^\alpha] \to \Psi(\alpha)$ as $n \to \infty$ for $|\ \alpha - \alpha^*\ | < \varepsilon^*$
2. Ψ is finite in a neighborhood of α^* and differentiable at α^* with $\Psi(\alpha^*) = 0,\ \Psi'(\alpha^*) > 0$

- When $\{J_n\}$ is i.i.d., $n^{-1} \log \mathbb{E}[(\Pi_n)^\alpha] = \Psi(\alpha) = \log \mathbb{E}[\mu(J_{-1})^\alpha]$
- Condition 1) allows for dependency in $\{J_n\} \cdots$

Before stating the results

Notation

$\{J_n\}$ - modulating process
$\mu(j) \triangleq \mathbb{E}[B_n^i(j)]$ - replication rate when in state j
$\Pi_n = \prod_{i=-n}^{-1} \mu(J_i),\ n \geq 1,\ \Pi_0 = I$ and $M = \sup_{n \geq 0} \Pi_n$

Polynomial Gärtner-Ellis conditions

1. $n^{-1} \log \mathbb{E}[(\Pi_n)^\alpha] \to \Psi(\alpha)$ as $n \to \infty$ for $| \alpha - \alpha^* | < \varepsilon^*$
2. Ψ is finite in a neighborhood of α^* and differentiable at α^* with $\Psi(\alpha^*) = 0,\ \Psi'(\alpha^*) > 0$

- When $\{J_n\}$ is i.i.d., $n^{-1} \log \mathbb{E}[(\Pi_n)^\alpha] = \Psi(\alpha) = \log \mathbb{E}[\mu(J_{-1})^\alpha]$
- Condition 1) allows for dependency in $\{J_n\} \cdots$

Main results

$\bar{B}_n^i \triangleq \sup_k B_n^i(k)$, $M = \sup_{n\geq 0} \Pi_n$.

Theorem

- *If* $\{\Pi_n\}$ *satisfies the polynomial Gärtner-Ellis conditions, and* $\mathbb{E}(\Pi_n)^{\alpha^*+\varepsilon} < \infty$, $\mathbb{E}[e^{\theta \bar{B}_n^i}] < \infty$ *(*$\varepsilon, \theta > 0$, $n \geq 1$*), then,*

$$\lim_{x\to\infty} \frac{\log \mathbb{P}[\Lambda > x]}{\log x} = \lim_{x\to\infty} \frac{\log \mathbb{P}[M > x]}{\log x} = -\alpha^*$$

- *If* $\sup_j \mu(j) < 1$ *and* $\mathbb{E}[e^{\theta \bar{B}_n^i}] < \infty$ *(*$\theta > 0$*), then,* $\mathbb{P}[\Lambda > x] = o(e^{-\xi x})$ *for some* $\xi > 0$*, implying*

$$\overline{\lim_{x\to\infty}} \frac{\log \mathbb{P}[\Lambda > x]}{\log x} = -\infty$$

What can we learn from the theorem?

What causes power laws? (expansions and contractions...)

Logarithmically asymptotic equivalence between the tails of M and Λ

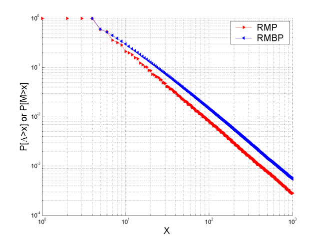
Reflected Multiplicative Process (RMP) → M

Define for $n \geq 0$ and $M_0 < \infty$

$$M_{n+1} = \max(M_n \cdot \mu(J_n), I),$$

If $\mathbb{E} \log J_n < 0$, then $M_n \xrightarrow{d} M$, $M = \sup_{n \geq 0} \prod_{i=-n}^{-1} \mu(J_i)$

Compare RMBP & RMP



What can we learn from the theorem?

What causes power laws? (expansions and contractions...)

Logarithmically asymptotic equivalence between the tails of M and Λ

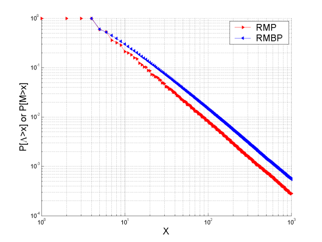
Reflected Multiplicative Process (RMP) $\rightarrow$ M

Define for $n \geq 0$ and $M_0 < \infty$

$$M_{n+1} = \max(M_n \cdot \mu(J_n), l),$$

If $\mathbb{E} \log J_n < 0$, then $M_n \xrightarrow{d} M$, $M = \sup_{n \geq 0} \prod_{i=-n}^{-1} \mu(J_i)$

Compare RMBP & RMP



What can we learn from the theorem?

What causes power laws? (expansions and contractions...)

Logarithmically asymptotic equivalence between the tails of M and Λ

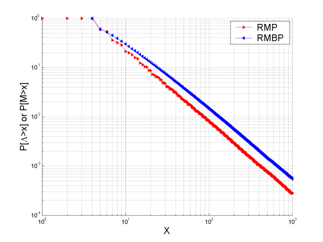
Reflected Multiplicative Process (RMP) $\rightarrow$ M

Define for $n \geq 0$ and $M_0 < \infty$

$$M_{n+1} = \max(M_n \cdot \mu(J_n), l),$$

If $\mathbb{E} \log J_n < 0$, then $M_n \xrightarrow{d} M$,
$M = \sup_{n \geq 0} \prod_{i=-n}^{-1} \mu(J_i)$

Compare RMBP & RMP



Outline

Recursion of RMP

For $I > 0$, $M_0 < \infty$ define RMP for all $n \geq 0$ as

$$M_{n+1} = \max(M_n \cdot J_{n+1}, I)$$

Waiting time of FIFO queue

Let $X_n = \log J_n$, $W_n = \log M_n$, $I = 1$,

$$W_{n+1} = \max(W_n + X_{n+1}, 0)$$

- Also, reflected random walk
- The connection between reflected i.i.d. random walks with negative drift and RMPs recognized and used by Cont and Sornette (1997)
- We emphasize that this connection extends beyond the i.i.d. random walks into the stationary and ergodic framework, and in general duality with queueing theory

Preliminaries - duality with queueing

Recursion of RMP

For $l > 0$, $M_0 < \infty$ define RMP for all $n \geq 0$ as

$$M_{n+1} = \max(M_n \cdot J_{n+1}, l)$$

Waiting time of FIFO queue

Let $X_n = \log J_n$, $W_n = \log M_n$, $l = 1$,

$$W_{n+1} = \max(W_n + X_{n+1}, 0)$$

- Also, reflected random walk
- The connection between reflected i.i.d. random walks with negative drift and RMPs recognized and used by Cont and Sornette (1997)
- We emphasize that this connection extends beyond the i.i.d. random walks into the stationary and ergodic framework, and in general duality with queueing theory

Preliminaries - duality with queueing

Recursion of RMP

For $l > 0$, $M_0 < \infty$ define RMP for all $n \geq 0$ as

$$M_{n+1} = \max(M_n \cdot J_{n+1}, l)$$

Waiting time of FIFO queue

Let $X_n = \log J_n$, $W_n = \log M_n$, $l = 1$,

$$W_{n+1} = \max(W_n + X_{n+1}, 0)$$

- Also, reflected random walk
- The connection between reflected i.i.d. random walks with negative drift and RMPs recognized and used by Cont and Sornette (1997)
- We emphasize that this connection extends beyond the i.i.d. random walks into the stationary and ergodic framework, and in general duality with queueing theory

Preliminaries - duality with queueing

Recursion of RMP

For $l > 0$, $M_0 < \infty$ define RMP for all $n \geq 0$ as

$$M_{n+1} = \max(M_n \cdot J_{n+1}, l)$$

Waiting time of FIFO queue

Let $X_n = \log J_n$, $W_n = \log M_n$, $l = 1$,

$$W_{n+1} = \max(W_n + X_{n+1}, 0)$$

- Also, reflected random walk
- The connection between reflected i.i.d. random walks with negative drift and RMPs recognized and used by Cont and Sornette (1997)
- We emphasize that this connection extends beyond the i.i.d. random walks into the stationary and ergodic framework, and in general duality with queueing theory

Preliminaries - duality with queueing

Exponentiated queueing processes

Assume that $J_n = e^{A_n - C_n}$ where $\{A_i\}, \{C_i\}$ are two mutually independent sequences. Then, $W_n = \log M_n$ satisfies

$$W_{n+1} = (W_n + A_n - C_n)^+$$

Example (Exponentiated M/M/1 queue)

If $\mathbb{P}[C_i > x] = e^{-\mu x}$, $\mathbb{P}[A_i > x] = e^{-\lambda x}$ and $\lambda < \mu$, then W_n represents the waiting time in a $M/M/1$ queue.

$$\mathbb{P}[W > x] = \frac{\lambda}{\mu} e^{-(\mu-\lambda)x}, \; x \geq 0,$$

equivalently it yields

$$\mathbb{P}[M > x] = \frac{\lambda}{\mu x^{(\mu-\lambda)}}, \; x \geq 1.$$

Preliminaries - duality with queueing

Exponentiated queueing processes

Assume that $J_n = e^{A_n - C_n}$ where $\{A_i\}, \{C_i\}$ are two mutually independent sequences. Then, $W_n = \log M_n$ satisfies

$$W_{n+1} = (W_n + A_n - C_n)^+$$

Example (Exponentiated M/M/1 queue)

If $\mathbb{P}[C_i > x] = e^{-\mu x}$, $\mathbb{P}[A_i > x] = e^{-\lambda x}$ and $\lambda < \mu$, then W_n represents the waiting time in a $M/M/1$ queue.

$$\mathbb{P}[W > x] = \frac{\lambda}{\mu} e^{-(\mu-\lambda)x}, \; x \geq 0,$$

equivalently it yields

$$\mathbb{P}[M > x] = \frac{\lambda}{\mu x^{(\mu-\lambda)}}, \; x \geq 1.$$

Preliminaries - duality with queueing

Translating queueing results

Basically, every result ever obtained in queueing theory (on exponential asymptotics) for single queue implies a corresponding result for RMPs (on power laws), including the results on:

- Loynes' stability (for RMPs logarithmic moment $\mathbb{E} \log J < 0$)
- Ladder heights, Cycle maximum (no reflection needed!)
- Diffusion (heavy traffic) limits, Exponential large deviations, etc

Theorem (Logarithmic asymptotics, Glynn & Whitt (1994))

Under the conditions of our main result, we have

$$\lim_{x \to \infty} \frac{\log \mathbb{P}[M > x]}{\log x} = -\alpha^*$$

Theorem allows for exponential correlations in $\{J_n\}$. But, if $\{J_n\}$ has subexponential correlation, then M can be very heavy $\mathbb{P}[M > x] \approx 1/l(x)$, $l(x)$ being slowly varying, e.g., even if J_n takes only two values, by exponentiating Thm 9 from Jelenković & Lazar, Adv. Appl. Prob. 31(2) 1999. Thus, power law distributions are on the lighter end of the spectrum for multiplicative (proportional growth) models, similarly as the exponential ones are for reflected additive models (queueing).

Preliminaries - duality with queueing

Translating queueing results

Basically, every result ever obtained in queueing theory (on exponential asymptotics) for single queue implies a corresponding result for RMPs (on power laws), including the results on:

- Loynes' stability (for RMPs logarithmic moment $\mathbb{E} \log J < 0$)
- Ladder heights, Cycle maximum (no reflection needed!)
- Diffusion (heavy traffic) limits, Exponential large deviations, etc

Theorem (Logarithmic asymptotics, Glynn & Whitt (1994))

Under the conditions of our main result, we have

$$\lim_{x \to \infty} \frac{\log \mathbb{P}[M > x]}{\log x} = -\alpha^*$$

Theorem allows for exponential correlations in $\{J_n\}$. But, if $\{J_n\}$ has subexponential correlation, then M can be very heavy $\mathbb{P}[M > x] \approx 1/l(x)$, $l(x)$ being slowly varying, e.g., even if J_n takes only two values, by exponentiating Thm 9 from Jelenković & Lazar, Adv. Appl. Prob. 31(2) 1999. Thus, power law distributions are on the lighter end of the spectrum for multiplicative (proportional growth) models, similarly as the exponential ones are for reflected additive models (queueing).

Preliminaries - duality with queueing

Translating queueing results

Basically, every result ever obtained in queueing theory (on exponential asymptotics) for single queue implies a corresponding result for RMPs (on power laws), including the results on:

- Loynes' stability (for RMPs logarithmic moment $\mathbb{E} \log J < 0$)
- Ladder heights, Cycle maximum (no reflection needed!)
- Diffusion (heavy traffic) limits, Exponential large deviations, etc

Theorem (Logarithmic asymptotics, Glynn & Whitt (1994))

Under the conditions of our main result, we have

$$\lim_{x \to \infty} \frac{\log \mathbb{P}[M > x]}{\log x} = -\alpha^*$$

Theorem allows for exponential correlations in $\{J_n\}$. But, if $\{J_n\}$ has subexponential correlation, then M can be very heavy $\mathbb{P}[M > x] \approx 1/l(x)$, $l(x)$ being slowly varying, e.g., even if J_n takes only two values, by exponentiating Thm 9 from Jelenković & Lazar, Adv. Appl. Prob. 31(2) 1999. Thus, power law distributions are on the lighter end of the spectrum for multiplicative (proportional growth) models, similarly as the exponential ones are for reflected additive models (queueing).

Preliminaries - duality with queueing

Translating queueing results

Basically, every result ever obtained in queueing theory (on exponential asymptotics) for single queue implies a corresponding result for RMPs (on power laws), including the results on:

- Loynes' stability (for RMPs logarithmic moment $\mathbb{E} \log J < 0$)
- Ladder heights, Cycle maximum (no reflection needed!)
- Diffusion (heavy traffic) limits, Exponential large deviations, etc

Theorem (Logarithmic asymptotics, Glynn & Whitt (1994))

Under the conditions of our main result, we have

$$\lim_{x \to \infty} \frac{\log \mathbb{P}[M > x]}{\log x} = -\alpha^*$$

Theorem allows for exponential correlations in $\{J_n\}$. But, if $\{J_n\}$ has subexponential correlation, then M can be very heavy $\mathbb{P}[M > x] \approx 1/l(x)$, $l(x)$ being slowly varying, e.g., even if J_n takes only two values, by exponentiating Thm 9 from Jelenković & Lazar, Adv. Appl. Prob. 31(2) 1999. Thus, power law distributions are on the lighter end of the spectrum for multiplicative (proportional growth) models, similarly as the exponential ones are for reflected additive models (queueing).

Outline

Sketch of the proof

General points

- Based on sample path arguments
- Representation lemma:

$$\Lambda_n \stackrel{d}{=} \max_{0 \le i \le n} Z_{-i} \to \max_{i \ge 0} Z_{-i} \stackrel{d}{=} \Lambda$$

- Identify the critical time scale within which $\max_{n \ge 0} Z_{-n}$ reaches a big value. For all $\beta > 0$,

$$\sum_{n > x}^{\infty} \mathbb{P}[Z_n^I > x] = o\left(\frac{1}{x^\beta}\right)$$

$$\Rightarrow \mathbb{P}[\Lambda > x] \sim \mathbb{P}[\Lambda_{\lfloor x \rfloor} > x]$$

Sketch of the proof

General points

- Based on sample path arguments
- Representation lemma:

$$\Lambda_n \stackrel{d}{=} \max_{0 \le i \le n} Z_{-i} \to \max_{i \ge 0} Z_{-i} \stackrel{d}{=} \Lambda$$

- Identify the critical time scale within which $\max_{n\geq 0} Z_{-n}$ reaches a big value. For all $\beta > 0$,

$$\sum_{n>x}^{\infty} \mathbb{P}[Z_n^l > x] = o\left(\frac{1}{x^{\beta}}\right)$$

$$\Rightarrow \mathbb{P}[\Lambda > x] \sim \mathbb{P}[\Lambda_{\lfloor x \rfloor} > x]$$

Sketch of the proof

General points

- Based on sample path arguments
- Representation lemma:

$$\Lambda_n \stackrel{d}{=} \max_{0 \le i \le n} Z_{-i} \to \max_{i \ge 0} Z_{-i} \stackrel{d}{=} \Lambda$$

- Identify the critical time scale within which $\max_{n \ge 0} Z_{-n}$ reaches a big value. For all $\beta > 0$,

$$\sum_{n > x}^{\infty} \mathbb{P}[Z_n^l > x] = o\left(\frac{1}{x^{\beta}}\right)$$

$$\Rightarrow \mathbb{P}[\Lambda > x] \sim \mathbb{P}[\Lambda_{\lfloor x \rfloor} > x]$$

Sketch of the proof for the upper bound

Increase the lower barrier to get an upper bound

Choose $l_x = \lfloor x^{\epsilon} \rfloor \geq l$, then, $\mathbb{P}[\Lambda^{l_x} > x] \geq \mathbb{P}[\Lambda^{l} > x]$, and

$$\mathbb{P}[\Lambda^l > x] = \mathbb{P}\left[\sup_{j\geq 1} Z^l_{-j} > x\right] \leq \mathbb{P}\left[\Lambda^{l_x}_{\lfloor x \rfloor} > x\right] + \sum_{j>x} \mathbb{P}[Z^l_j > x]$$

$\curvearrowleft$ now, l_x large $\Rightarrow Z^{l_x}_i \approx \Pi^{l_x}_i$

$$\leq \mathbb{P}\left[\sup_{j\geq 1} \Pi_j (1+\epsilon)^j > x^{1-\epsilon}\right] + x\mathbb{P}[\mathcal{B}^{l_x,\epsilon}_1] + \sum_{j>x} \mathbb{P}[Z^l_j > x],$$

where $\mathcal{B}^{l_x,\epsilon}_1 = \bigcup_{j\geq l_x} \{ \sum_{i=1}^{j} B^i_1(J_1) > j\mu(J_1)(1+\epsilon) \}$, $\Pi_j = \prod_{i=-1}^{-j} \mu(J_i)$

Sketch of the proof for the lower bound

How to increase the lower barrier but still obtain a lower bound?

- If $\{\Lambda_n^{y_1}\}$ and $\{\Lambda_n^{y_2}\}$ are defined on the same $\{J_n\}_{n\geq 0}$, but independent copies of $\{B_n^i(j)\}$, then

$$\Lambda_n^{y_1+y_2} \overset{d}{\leq} \Lambda_n^{y_1} + \Lambda_n^{y_2}$$

- $\mathbb{P}[\Lambda_n^l > x] \geq \mathbb{P}[\Lambda_n^1 > x] = \frac{y \cdot \mathbb{P}[\Lambda_n^1 > x]}{y} \geq \frac{\mathbb{P}[\sum_{j=1}^{y} \Lambda_{n,j}^1 > y \cdot x]}{y} \geq \frac{\mathbb{P}[\Lambda_n^y > yx]}{y}$
- Choose $y = \lfloor x^\delta \rfloor$, $n = \lfloor x \rfloor$
- Then, use similar arguments for the upper bound $\cdots$
- More details are in the paper...

Sketch of the proof for the lower bound

How to increase the lower barrier but still obtain a lower bound?

- If $\{\Lambda_n^{y_1}\}$ and $\{\Lambda_n^{y_2}\}$ are defined on the same $\{J_n\}_{n\geq 0}$, but independent copies of $\{B_n^i(j)\}$, then

$$\Lambda_n^{y_1+y_2} \overset{d}{\leq} \Lambda_n^{y_1} + \Lambda_n^{y_2}$$

- $\mathbb{P}[\Lambda_n^l > x] \geq \mathbb{P}[\Lambda_n^1 > x] = \frac{y \cdot \mathbb{P}[\Lambda_n^1 > x]}{y} \geq \frac{\mathbb{P}[\sum_{j=1}^{y} \Lambda_{n,j}^1 > y \cdot x]}{y} \geq \frac{\mathbb{P}[\Lambda_n^y > yx]}{y}$
- Choose $y = \lfloor x^\delta \rfloor$, $n = \lfloor x \rfloor$
- Then, use similar arguments for the upper bound $\cdots$
- More details are in the paper...

Sketch of the proof for the lower bound

How to increase the lower barrier but still obtain a lower bound?

- If $\{\Lambda_n^{y_1}\}$ and $\{\Lambda_n^{y_2}\}$ are defined on the same $\{J_n\}_{n\geq 0}$, but independent copies of $\{B_n^i(j)\}$, then

$$\Lambda_n^{y_1+y_2} \overset{d}{\leq} \Lambda_n^{y_1} + \Lambda_n^{y_2}$$

- $\mathbb{P}[\Lambda_n^l > x] \geq \mathbb{P}[\Lambda_n^1 > x] = \frac{y \cdot \mathbb{P}[\Lambda_n^1 > x]}{y} \geq \frac{\mathbb{P}[\sum_{j=1}^{y} \Lambda_{n,j}^1 > y \cdot x]}{y} \geq \frac{\mathbb{P}[\Lambda_n^y > yx]}{y}$
- Choose $y = \lfloor x^{\delta} \rfloor$, $n = \lfloor x \rfloor$
- Then, use similar arguments for the upper bound $\cdots$
- More details are in the paper...

Sketch of the proof for the lower bound

How to increase the lower barrier but still obtain a lower bound?

- If $\{\Lambda_n^{y_1}\}$ and $\{\Lambda_n^{y_2}\}$ are defined on the same $\{J_n\}_{n\geq 0}$, but independent copies of $\{B_n^i(j)\}$, then

$$\Lambda_n^{y_1+y_2} \overset{d}{\leq} \Lambda_n^{y_1} + \Lambda_n^{y_2}$$

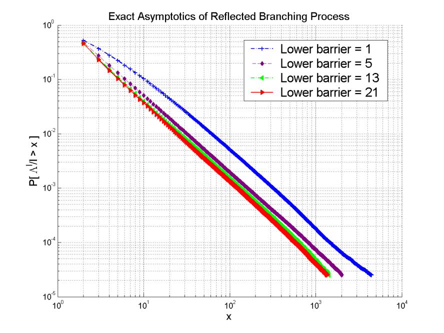
- $\mathbb{P}[\Lambda_n^l > x] \geq \mathbb{P}[\Lambda_n^1 > x] = \frac{y \cdot \mathbb{P}[\Lambda_n^1 > x]}{y} \geq \frac{\mathbb{P}[\sum_{j=1}^{y} \Lambda_{n,j}^1 > y \cdot x]}{y} \geq \frac{\mathbb{P}[\Lambda_n^y > yx]}{y}$
- Choose $y = \lfloor x^\delta \rfloor$, $n = \lfloor x \rfloor$
- Then, use similar arguments for the upper bound $\cdots$
- More details are in the paper...

Outline

Exact Asymptotics of RMBP

Difficult! But, in the scaling region, when barrier grows $\cdots$

$\{J_n : n \geq 1\}$ i.i.d.
barrier $l \geq (\log x)^{3+\epsilon}$



Outline

Modulated BP with an Absorbing Barrier

New Model of Hotspot Visitors

Hotspot dynamics for Web sites.
A tells *B*, *C* to visit the web,
and later *B* may tell *D*.

Model

- $\{J_n\}$ - i.i.d. modulating process; A_t - triggering events.
- Define stopping time $P \triangleq \inf\{n > 0 : Z_n^l \leq l\}$ for l; after P the process is killed/absorbed.
- At time t, Poisson A_t $\#$ of objects are created, and each evolves according to an i.i.d. copy of Z_P.

Modulated BP with an Absorbing Barrier

New Model of Hotspot Visitors

Hotspot dynamics for Web sites.
A tells *B*, *C* to visit the web,
and later *B* may tell *D*.

Theorem

Assume that $\mathbb{E}[J_n^\alpha] < \infty$, $\alpha - \epsilon < \alpha^* < \alpha + \epsilon$ *and* $\mathbb{E}[e^{\theta \bar{B}_n^i}] < \infty$ *for some* $\theta > 0$*, then,*

$$\lim_{x \to \infty} \frac{\log \mathbb{P}[Z_s > x]}{\log x} = -\alpha^*.$$

Outline

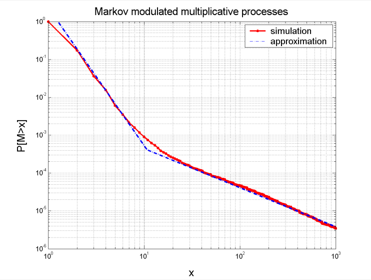
New explanations of double Pareto

Double Pareto - frequent empirical observations

- Transition from heavy-traffic region to large deviation region
- Multiple time scales resulting in double Pareto

Example (Jelenković & Lazar (1995), queueing context)

$J_n \equiv J(X(n))$,
$X(n)$- Markov chain
$p_{12} = 1/5000$,
$p_{21} = 1/10$,
$\mathbb{P}[J(1) = 1.2] = 0.5$,
$\mathbb{P}[J(1) = 0.6] = 0.5$,
$\mathbb{P}[J(2) = 1.7] = 0.6$,
$\mathbb{P}[J(2) = 0.25] = 0.4$.



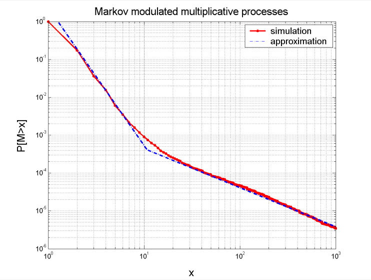
New explanations of double Pareto

Double Pareto - frequent empirical observations

- Transition from heavy-traffic region to large deviation region
- Multiple time scales resulting in double Pareto

Example (Jelenković & Lazar (1995), queueing context)

$J_n \equiv J(X(n))$,
$X(n)$- Markov chain
$p_{12} = 1/5000$,
$p_{21} = 1/10$,
$\mathbb{P}[J(1) = 1.2] = 0.5$,
$\mathbb{P}[J(1) = 0.6] = 0.5$,
$\mathbb{P}[J(2) = 1.7] = 0.6$,
$\mathbb{P}[J(2) = 0.25] = 0.4$.



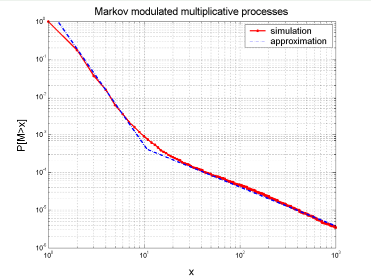
New explanations of double Pareto

Double Pareto - frequent empirical observations

- Transition from heavy-traffic region to large deviation region
- Multiple time scales resulting in double Pareto

Example (Jelenković & Lazar (1995), queueing context)

$J_n \equiv J(X(n))$,
$X(n)$- Markov chain
$p_{12} = 1/5000$,
$p_{21} = 1/10$,
$\mathbb{P}[J(1) = 1.2] = 0.5$,
$\mathbb{P}[J(1) = 0.6] = 0.5$,
$\mathbb{P}[J(2) = 1.7] = 0.6$,
$\mathbb{P}[J(2) = 0.25] = 0.4$.



Outline

Related phenomena and models

- Truncated power laws $\cdots$
  M(B)P with both lower and upper barriers
  Similarly as obtaining truncated geometric distributions in finite buffer queue, e.g., $M/M/1/B$
  Cont and Sornette (1997)
- Randomly stopped multiplicative processes $\rightleftharpoons$ RMP
  ladder height representation + Pollaczek-Khintchine formula
- Randomly stopped MBP

Concluding remarks

- RMBP - new general model of proportional growth
- Under the general polynomial Gärtner-Ellis conditions,

  RMBP $\Rightarrow$ power laws $\Rightarrow$ ubiquitous nature of power law
- Emphasize the duality:

  additive processes (queueing theory) $\Leftrightarrow$ proportional growth

Amusing question

What is more frequent, power law or exponential distributions? $\Longleftrightarrow$ What is more frequent, proportional growth or additive growth?

Another new phenomenon: Restarts and Retransmissions
(Papers: http://www.ee.columbia.edu/~predrag)
Thanks! Questions?

Concluding remarks

- RMBP - new general model of proportional growth
- Under the general polynomial Gärtner-Ellis conditions,

  RMBP $\Rightarrow$ power laws $\Rightarrow$ ubiquitous nature of power law
- Emphasize the duality:

  additive processes (queueing theory) $\Leftrightarrow$ proportional growth

Amusing question

What is more frequent, power law or exponential distributions? $\Longleftrightarrow$ What is more frequent, proportional growth or additive growth?

Another new phenomenon: Restarts and Retransmissions
(Papers: http://www.ee.columbia.edu/~predrag)
Thanks! Questions?

Concluding remarks

- RMBP - new general model of proportional growth
- Under the general polynomial Gärtner-Ellis conditions,

  RMBP $\Rightarrow$ power laws $\Rightarrow$ ubiquitous nature of power law
- Emphasize the duality:

  additive processes (queueing theory) $\Leftrightarrow$ proportional growth

Amusing question

What is more frequent, power law or exponential distributions?	$\Longleftrightarrow$	What is more frequent, proportional growth or additive growth?

Another new phenomenon: Restarts and Retransmissions
(Papers: http://www.ee.columbia.edu/∼predrag)
Thanks! Questions?

Concluding remarks

- RMBP - new general model of proportional growth
- Under the general polynomial Gärtner-Ellis conditions,

  RMBP $\Rightarrow$ power laws $\Rightarrow$ ubiquitous nature of power law
- Emphasize the duality:

  additive processes (queueing theory) $\Leftrightarrow$ proportional growth

Amusing question

What is more frequent, power law or exponential distributions?	$\Longleftrightarrow$	What is more frequent, proportional growth or additive growth?

Another new phenomenon: Restarts and Retransmissions
(Papers: http://www.ee.columbia.edu/$\sim$predrag)
Thanks! Questions?