Ten Lectures on Small Value Probabilities and Applications

L3: Probabilistic Techniques for Independent Sums

Wenbo V. Li

University of Delaware

http://www.math.udel.edu/~wli
wli@math.udel.edu

CBMS Lectures at UAH, June 4-8, 2012

Separate treatments are analyzed for exponential and power decay rates. Many estimates are non-asymptotic and hence they can be applied in the setting of conditional probability.

Direct Probabilistic Arguments and Partitions in Finer Scale

A basic and important starting point for many probability estimates is the study of the behavior of the sums of independent random variables in terms of individual terms. It can be viewed as a fundamental algebraic property. We will introduce some techniques such as partitions in finer scale, and also use some familiar arguments such as exponential Chebyshev's inequality.

- In the last lecture, we have found the small value probability for independent sums based on the behavior of its individual terms. Here we provide direct probabilistic arguments, often separated for upper and lower bounds. These arguments can be used in the conditioning setting where the Tauberian type results can not be applied. We will see such type of arguments for processes running on independent clocks.
- Once again, we try to emphasis various proofs of the same results based on different methods.

Simple but Rough Lower Bounds: Polynomial Rate

Lemma: Let $V_1 \geq 0$ and $V_2 \geq 0$ be two independent random variables such that

$$\mathbb{P}(V_i \leq t) \geq c_i t^{\alpha_i}, \quad i = 1, 2 \tag{0.1}$$

for all $t > 0$ small. Then for all $t > 0$ small

$$\mathbb{P}(V_1 + V_2 \leq t) \geq \frac{c_1 c_2 \alpha_1^{\alpha_1} \alpha_2^{\alpha_2}}{(\alpha_1 + \alpha_2)^{\alpha_1 + \alpha_2}} t^{\alpha_1 + \alpha_2}. \tag{0.2}$$

Proof: For any $0 < \lambda < 1$, we have

$$\begin{aligned} \mathbb{P}(V_1 + V_2 \leq t) &\geq \mathbb{P}(V_1 \leq \lambda t, V_2 \leq (1-\lambda)t) \\ &= \mathbb{P}(V_1 \leq \lambda t) \cdot \mathbb{P}(V_2 \leq (1-\lambda)t) \\ &\geq c_1 c_2 \lambda^{\alpha_1} (1-\lambda)^{\alpha_2} t^{\alpha_1 + \alpha_2}. \end{aligned}$$

Pick the best parameter $\lambda = \alpha_1/(\alpha_1 + \alpha_2)$ and the estimate follows from

$$\max_{0<\lambda<1} \lambda^{\alpha_1} (1-\lambda)^{\alpha_2} = \frac{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2}}{(\alpha_1 + \alpha_2)^{\alpha_1 + \alpha_2}}$$

which can be checked easily by calculus.

Simple and Sharp Lower Bounds: Exponential Rate

Lemma: Let $V_1 \geq 0$ and $V_2 \geq 0$ be two independent random variables such that

$$\log \mathbb{P}(V_i \leq t) \geq -c_i t^{-\alpha_i}, \quad i = 1, 2$$

for all $t > 0$ small. Then for all $t > 0$ small

$$\log \mathbb{P}(V_1 + V_2 \leq t) \geq \begin{cases} -\left(c_1^{1/(1+\alpha)} + c_2^{1/(1+\alpha)}\right)^{1+\alpha} t^{-\alpha} & \text{if} \quad \alpha_1 = \alpha_2 \\ \text{?HW?} & \text{if} \quad \alpha_1 \neq \alpha_2 \end{cases}$$

Proof: The case $\alpha_1 = \alpha_2 = \alpha > 0$ follows from

$$\begin{aligned} \log \mathbb{P}(V_1 + V_2 \leq t) &\geq \log \mathbb{P}(V_1 \leq \lambda t) + \log \mathbb{P}(V_2 \leq (1-\lambda)t) \\ &\geq -(c_1 \lambda^{-\alpha} + c_2 (1-\lambda)^{-\alpha}) t^{-\alpha}. \end{aligned}$$

Pick the best parameter $\lambda = c_1^{1/(1+\alpha)} / (c_1^{1/(1+\alpha)} + c_2^{1/(1+\alpha)})$ and the estimate follows from

$$\min_{0<\lambda<1} (c_1 \lambda^{-\alpha} + c_2 (1-\lambda)^{-\alpha}) = \left(c_1^{1/(1+\alpha)} + c_2^{1/(1+\alpha)}\right)^{1+\alpha}$$

which can be checked easily by calculus. $\square$

HW: Find the estimate in the above theorem for $\alpha_1 \neq \alpha_2$.

Simple but Rough Upper Bounds: Polynomial Rate

Lemma: Let $V_1 \geq 0$ and $V_2 \geq 0$ be two independent random variables such that

$$\mathbb{P}(V_i \leq t) \leq c_i t^{\alpha_i} \quad i = 1, 2, \tag{0.3}$$

for all $t > 0$ small. Then for all $t > 0$ small

$$\mathbb{P}(V_1 + V_2 \leq t) \leq c_1 c_2 t^{\alpha_1 + \alpha_2} \tag{0.4}$$

Proof: The result follows easily from

$$\mathbb{P}(V_1 + V_2 \leq t) \leq \mathbb{P}(\max(V_1, V_2) \leq t) = \mathbb{P}(V_1 \leq t) \cdot \mathbb{P}(V_2 \leq t) \leq c_1 c_2 t^{\alpha_1 +}$$

•A relation between sum and max:

$$\mathbb{P}(\max(V_1, V_2) \leq t/2) \leq \mathbb{P}(V_1 + V_2 \leq t) \leq \mathbb{P}(\max(V_1, V_2) \leq t).$$

Simple Upper Bounds: Exponential Rate

Lemma: Let $V_1 \geq 0$ and $V_2 \geq 0$ be two independent random variables such that

$$\log \mathbb{P}(V_i \leq t) \leq -c_i t^{-\alpha_i}, \quad i = 1, 2 \tag{0.5}$$

for all $t > 0$ small. Then for all $t > 0$ small

$$\log \mathbb{P}(V_1 + V_2 \leq t) \leq \begin{cases} -(c_1 + c_2)t^{-\alpha} & \text{if} \quad \alpha_1 = \alpha_2 = \alpha > 0 \\ -c_2 t^{-\alpha} & \text{if} \quad \alpha_1 < \alpha_2 = \alpha \end{cases} \tag{0.6}$$

Proof: Exact the same argument as the above proof. □

•This is sharp under the assumption $\alpha_1 \neq \alpha_2$ and not sharp under the assumption $\alpha_1 = \alpha_2 = \alpha$. We can tighten the bound by the argument given below.

Thm: Let V_1 and V_2 be independent random variables such that $\mathbb{P}(V_i \leq t) \sim c_i t^{\alpha_i}$ as $t \to 0$. Then as $t \to 0$

$$\mathbb{P}(V_1 + V_2 \leq t) \sim c_1 c_2 \frac{\Gamma(\alpha_1 + 1)\Gamma(\alpha_2 + 1)}{\Gamma(\alpha_1 + \alpha_2 + 1)} t^{\alpha_1 + \alpha_2} \tag{0.7}$$

•Even though this result follows from Tauberian's theorem given in the previous lecture, it is still useful and important to give a direct arguments based on probabilistic arguments. The arguments are applicable to the conditional setting.

Proof of the Lower Bound

Let m be a large positive integer. First we consider the lower bound. It is natural to exam the probability in finer scale to obtain

$$\mathbb{P}(V_1 + V_2 \leq t) = \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m} t < V_1 \leq \frac{i}{m} t, V_1 + V_2 \leq t\right)$$

Proof of the Lower Bound

Let m be a large positive integer. First we consider the lower bound. It is natural to exam the probability in finer scale to obtain

$$\begin{aligned}
\mathbb{P}(V_1 + V_2 \le t) &= \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m}t < V_1 \le \frac{i}{m}t, V_1 + V_2 \le t\right) \\
&\ge \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m}t < V_1 \le \frac{i}{m}t, V_2 \le \frac{m-i}{m}t\right) \\
&= \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m}t < V_1 \le \frac{i}{m}t\right) \cdot \mathbb{P}\left(V_2 \le \frac{m-i}{m}t\right) \\
&= \sum_{i=1}^{m} \left(\mathbb{P}\left(V_1 \le \frac{i}{m}t\right) - \mathbb{P}\left(V_1 \le \frac{i-1}{m}t\right)\right) \\
&\qquad \cdot \mathbb{P}\left(V_2 \le \frac{m-i}{m}t\right)
\end{aligned}$$

Thus we have

$$\begin{aligned}
& \liminf_{t\to 0} t^{-(\alpha_1+\alpha_2)}\mathbb{P}(V_1+V_2\le t) \\
\ge \;& \sum_{i=1}^{m}\left(c_1(\frac{i}{m})^{\alpha_1}-c_1(\frac{i-1}{m})^{\alpha_1}\right)\cdot c_2(\frac{m-i}{m})^{\alpha_2} \\
= \;& c_1c_2\frac{1}{m}\sum_{i=1}^{m}\alpha_1\left(\frac{q_i}{m}\right)^{\alpha_1-1}\cdot\left(1-\frac{i}{m}\right)^{\alpha_2}
\end{aligned}$$

where $i-1\le q_i\le i$ is determined by the mean value theorem. Taking $m\to\infty$, we obtain by the approximation of Rimann sums.

$$\begin{aligned}
\liminf_{t\to 0} t^{-(\alpha_1+\alpha_2)}\mathbb{P}(V_1+V_2\le t) &\ge c_1c_2\int_0^1\alpha_1 x^{\alpha_1-1}(1-x)^{\alpha_2}dx \\
&= c_1c_2\frac{\Gamma(\alpha_1+1)\Gamma(\alpha_2+1)}{\Gamma(\alpha_1+\alpha_2+1)}.
\end{aligned}$$

Proof of the Upper Bound

Next we consider the upper bound in the similar way as the lower bound. It follows that in the finer scale,

$$
\begin{aligned}
\mathbb{P}(V_1 + V_2 \le t) &= \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m}t < V_1 \le \frac{i}{m}t, V_1 + V_2 \le t\right) \\
&\le \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m}t < V_1 \le \frac{i}{m}t, V_2 \le \frac{m-i+1}{m}t\right) \\
&= \sum_{i=1}^{m} \left(\mathbb{P}\left(V_1 \le \frac{i}{m}t\right) - \mathbb{P}\left(V_1 \le \frac{i-1}{m}t\right)\right) \\
&\qquad \cdot \mathbb{P}\left(V_2 \le \frac{m-i+1}{m}t\right)
\end{aligned}
$$

Thus we have

$$\begin{aligned} & \limsup_{t\to 0} t^{-(\alpha_1+\alpha_2)}\mathbb{P}(V_1+V_2\le t) \\ \le\ & \sum_{i=1}^{m}\left(c_1(\frac{i}{m})^{\alpha_1}-c_1(\frac{i-1}{m})^{\alpha_1}\right)\cdot c_2(\frac{m-i+1}{m})^{\alpha_2} \\ =\ & c_1c_2\frac{1}{m}\sum_{i=1}^{m}\alpha_1\left(\frac{q_i}{m}\right)^{\alpha_1-1}\cdot\left(1-\frac{i-1}{m}\right)^{\alpha_2} \end{aligned}$$

where $i-1\le q_i\le i$ is determined by the mean value theorem. Taking $m\to\infty$, we obtain by the approximation of Rimann sums.

$$\begin{aligned} \limsup_{t\to 0} t^{-(\alpha_1+\alpha_2)}\mathbb{P}(V_1+V_2\le t) &\le c_1c_2\int_0^1\alpha_1x^{\alpha_1-1}(1-x)^{\alpha_2}dx \\ &= c_1c_2\frac{\Gamma(\alpha_1+1)\Gamma(\alpha_2+1)}{\Gamma(\alpha_1+\alpha_2+1)}. \end{aligned}$$

This finishes the entire proof. $\square$

Probabilistic Arguments for Independent Sums

Similarly, we reprove the following result.

Thm: Let V_1 and V_2 be independent random variables such that $\log \mathbb{P}(V_i \le t) \sim -c_i t^{-\alpha_i}$ as $t \to 0$. Then as $t \to 0$

$$\log \mathbb{P}(V_1 + V_2 \le t) \sim \begin{cases} -\left(c_1^{1/(1+\alpha)} + c_2^{1/(1+\alpha)}\right)^{1+\alpha} t^{-\alpha} & \text{if} \quad \alpha_1 = \alpha_2 = \alpha > 0 \\ -c_2 t^{-\alpha} & \text{if} \quad \alpha_1 < \alpha_2 = \alpha \end{cases}$$

Pf: The lower estimates is easy, and also the upper estimate for $\alpha_2 > \alpha_1 = \alpha$. For the upper estimate in the case $\alpha_1 = \alpha_2 = \alpha > 0$, we have

$$\begin{aligned}
\mathbb{P}(V_1 + V_2 \le t) &= \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m}t < V_1 \le \frac{i}{m}t, V_2 \le \frac{m-i+1}{m}t\right) \\
&\le \sum_{i=1}^{m} \mathbb{P}\left(\frac{i-1}{m}t < V_1 \le \frac{i}{m}t\right) \cdot \mathbb{P}\left(V_2 \le \frac{m-i+1}{m}t\right) \\
&\le \sum_{i=1}^{m} \mathbb{P}\left(V_1 \le \frac{i}{m}t\right) \cdot \mathbb{P}\left(V_2 \le \frac{m-i+1}{m}t\right)
\end{aligned}$$

For any given $\varepsilon > 0$ small, there is a $\delta_\varepsilon > 0$ such that for all $0 < t \le \delta_\varepsilon$,

$$\log \mathbb{P}(V_i \le t) \le -(1-\varepsilon)c_i t^{-\alpha}, \quad i = 1, 2.$$

Thus

$$\begin{aligned}
& \mathbb{P}\left(V_1 \leq \frac{i}{m} t\right) \cdot \mathbb{P}\left(V_2 \leq \frac{m-i+1}{m} t\right) \\
\leq\ & \exp\{-(1-\varepsilon) \cdot \left(c_1 (i/m)^{-\alpha} + c_2(1-(i/m)+(1/m))^{-\alpha}\right) \cdot t^{-\alpha}\} \\
\leq\ & \exp\{-(1-\varepsilon) \cdot \min_{0 \leq x \leq 1} \left(c_1 x^{-\alpha} + c_2(1-x+(1/m))^{-\alpha}\right) \cdot t^{-\alpha}\}
\end{aligned}$$

Hence we have

$$\limsup_{t \to 0} t^{-\alpha} \log \mathbb{P}(V_1 + V_2 \leq t) \quad \leq \quad ...$$

Taking $m \to \infty$ first and then $\varepsilon \to 0$ finishes the proof. Note that there are m terms in the sum and the number of terms disappeared by taking $t \to 0$. $\square$

•Note that at the exponential rate, $\mathbb{P}(V \leq i/m)$ is an order of magnitude different from $\mathbb{P}(V \leq (i-1)/m)$. Hence dropping the much smaller term does not change the overall asymptotic. This is not the case for the polynomial rate.

•The above argument shows that we do have a non-asymptotic upper estimate in terms of m.

More generally, we have the following results.
Thm: Let $V_1 \geq 0$ and $V_2 \geq 0$ be independent random variables such that

$$\mathbb{P}(V_i \leq t) \sim b_i t^{\beta_i} \exp\{-c_i t^{-\alpha_i}\}$$

as $t \to 0$. Then as $t \to 0$

$$\mathbb{P}(V_1+V_2 \leq t) \sim \begin{cases} b_{12} t^{\beta_1+\beta_2-\alpha/2} \cdot \exp\{c_{12} t^{-\alpha}\} & \text{if} \quad \alpha_1 = \alpha_2 = \alpha > \\ \text{?HW?} & \text{if} \quad \alpha_1 < \alpha_2 = \alpha \end{cases}$$

where

$$\begin{aligned} b_{12} &= b_1 b_2 \left(2\pi\alpha/(1+\alpha)\right)^{1/2} \left(c_1^{1/(1+\alpha)} + c_2^{1/(1+\alpha)}\right)^{(1+\alpha)/2-\beta_1-\beta_2} \cdot \\ & \quad \cdot c_1^{(1+2\beta_1)/2(1+\alpha)} c_2^{(1+2\beta_2)/2(1+\alpha)} \end{aligned}$$

and

$$c_{12} = \left(c_1^{1/(1+\alpha)} + c_2^{1/(1+\alpha)}\right)^{1+\alpha}.$$

•We omit the detailed proofs based on general Tauberian theorems.

Sums, Single Term and Ind. Binomial Sums

The following argument changes the small value estimate into a simple, but useful upper tail Markov bound and/or bounds on Ind. Binomial Sums. The idea can also be used in several other situations.

Lemma: Let V_i be i.i.d non-negative random variables. Then for any $\varepsilon > 0$ and $0 < \lambda < 1$,

$$\mathbb{P}(\sum_{i=1}^{n} V_i \leq n\varepsilon) \leq (1-\lambda)^{-1}\mathbb{P}(V_1 \leq \lambda^{-1}\varepsilon).$$

In particular, for $\lambda = 1/2$,

$$\mathbb{P}(\sum_{i=1}^{n} V_i \leq n\varepsilon) \leq 2\mathbb{P}(V_1 \leq 2\varepsilon).$$

•Various refinement are possible.

Sums, Single Term and Ind. Binomial Sums

The following argument changes the small value estimate into a simple, but useful upper tail Markov bound and/or bounds on Ind. Binomial Sums. The idea can also be used in several other situations.

Lemma: Let V_i be i.i.d non-negative random variables. Then for any $\varepsilon > 0$ and $0 < \lambda < 1$,

$$\mathbb{P}(\sum_{i=1}^{n} V_i \leq n\varepsilon) \leq (1-\lambda)^{-1}\mathbb{P}(V_1 \leq \lambda^{-1}\varepsilon).$$

In particular, for $\lambda = 1/2$,

$$\mathbb{P}(\sum_{i=1}^{n} V_i \leq n\varepsilon) \leq 2\mathbb{P}(V_1 \leq 2\varepsilon).$$

•Various refinement are possible. **Proof:** Note that

$$\sum_{i=1}^{n} V_i \geq \sum_{i=1}^{n} V_i \mathbb{I}(V_i \geq \lambda^{-1}\varepsilon) \geq \lambda^{-1}\varepsilon \sum_{i=1}^{n} \mathbb{I}(V_i \geq \lambda^{-1}\varepsilon).$$

Hence by Markov's inequality,

$$
\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} V_i \le n\varepsilon\right) &\le \mathbb{P}\left(\lambda^{-1}\varepsilon \sum_{i=1}^{n} \mathbb{I}(V_i \ge \lambda^{-1}\varepsilon) \le n\varepsilon\right) \\
&= \mathbb{P}\left(\sum_{i=1}^{n} \mathbb{I}(V_i < \lambda^{-1}\varepsilon) \ge (1-\lambda)n\right) \\
&\le \frac{n\mathbb{P}(V_1 < \lambda^{-1}\varepsilon)}{(1-\lambda)n}
\end{aligned}
$$

which finished the proof. $\square$

•Note that exponential type bounds can be given based on sum of ind. Bernoulli's via large deviation. Of course, if we know the Laplace transform of $\mathbb{E}\exp(-\delta V)$ for $\delta > 0$, then we can use exponential Chebyshev's inequality.

Large deviation: Cramer's Upper Bound

Let $\{X_i, i \geq 1\}$ be a sequence of i.i.d. random variables. Then, for all $a \leq \mathbb{E}\, X_1$,

$$\mathbb{P}(S_n \leq na) \leq e^{-nI(a)},$$

where $I(a) = \sup_{\lambda \geq 0}(-\lambda a - \log \mathbb{E}\, e^{-\lambda X_1})$.

•See e.g., Dembo and Zeitouni (1998) [Theorem 2.2.3] for a more precise version of Cramer's Theorem

EX: For i.i.d. p-Bernoulli variables X_i, and for $\beta < p$,

$$\mathbb{P}(S_n \leq \beta n) \leq e^{-n[\beta \log(\beta/p) + (1-\beta)\log((1-\beta)/(1-p))]}.$$

Various simpler bounds can be found in Shorack and Wellner (1986, book) and Janson, Luczak and Rucinski (2001, book on Random Graphs).

•We can also consider Chernoff type bound for ind. Bernoulli sums (not identical). Set $T = \sum_{j=1}^{n} X_i$ where each X_i is a random indicator variable with $\mathbb{P}(X_i = 1) = p_i = 1 - \mathbb{P}(X_i = 0)$. It is nature to compare T with $S \sim \mathrm{Bi}(n, \bar{p})$ with $\mathbb{E}\, T = \mathbb{E}\, S = \sum_{i=1}^{n} p_i = n\bar{p}$. In fact, it follows from Jensen's inequality after taking logarithm that for $\lambda \in \mathbb{R}$,

$$\mathbb{E}\, e^{\lambda T} = \prod_{i=1}^{n} (1 + p_i(e^{\lambda} - 1)) \leq (1 + \bar{p}(e^{\lambda} - 1))^n = \mathbb{E}\, e^{\lambda S}.$$

Consequently, every bound derived from exponential Chebyshev's inequality applies to T.

•More general cases, such that independent X_i with $0 \leq X_i \leq 1$, can be found in Bennett (1962), Hoeffding (1963), etc.

SVP for the Martingale Limit of a Galton-Watson Tree

Consider the Galton-Watson branching process $(Z_n)_{n\geq 0}$ with offspring distribution $(p_k)_{k\geq 0}$ starting with $Z_0 = 1$. In any subsequent generation individuals independently produce a random number of offspring according to $\mathbb{P}(X = k) = p_k$. Suppose $m = \mathbb{E}\, X > 1$ and $\mathbb{E}\, X \log X < \infty$. Then by Kesten-Stigum theorem, the martingale limit (a.s and in L^1)

$$W = \lim_{n\to\infty} \frac{Z_n}{m^n}$$

exists and is nontrivial almost surely with $\mathbb{E}\, W = 1$. WOLG, assume $p_0 = 0$ and $p_k < 1$ for all $k \geq 1$. Then in the case $p_1 > 0$, there exist constants $0 < c < C < \infty$ such that for all $0 < \varepsilon < 1$

$$c\varepsilon^\tau \leq \mathbb{P}(W \leq \varepsilon) \leq C\varepsilon^\tau, \quad \tau = -\log p_1/\log m$$

and in the case $p_1 = 0$, there exist constants $0 < c < C < \infty$ such that for all $0 < \varepsilon < 1$

$$c\varepsilon^{-\beta/(1-\beta)} \leq -\log \mathbb{P}(W \leq \varepsilon) \leq C\varepsilon^{-\beta/(1-\beta)}.$$

with $\nu = \min\{k \geq 2 : p_k \neq 0\}$ and $\beta = \log \nu/\log m < 1$.

•These results are due to Dubuc (1971a,b) in the $p_1 > 0$ case, and up to a Tauberian theorem also in the $p_1 = 0$ case, see Bingham (1988). The proofs are relying on nontrivial complex analysis and are therefore difficult to generalize, for example to processes with immigration and/or dependent offsprings.
•Examples, near-constancy phenomena and various refinements, see Harris (1948), Karlin and McGregor (1968 a,b), Dubuc (1982), Barlow and Perkins (1987), Goldstein (1987) and Kusuoka (1987), Bingham (1988), Biggins and Bingham (1991), Biggins and Bingham (1993), Biggins and Nadarajah (1994), Hambly (1995), Fleischman and Wachtel (2007, 2009).
•A probabilistic argument is given in Mörters and Ortgiese (2008).

SVP for supercritical branching processes with Immigration

Consider the supercritical branching process with immigration, denoted by $(\mathcal{Z}_n, n \geq 0)$. That is

$$\mathcal{Z}_0 = Y_0, \quad \mathcal{Z}_{n+1} = X_1^n + X_2^n + \cdots + X_{\mathcal{Z}_n}^n + Y_{n+1}, \quad n \geq 0,$$

where $X_1^n, X_2^n, \cdots$ are independent and identically distributed with the same offspring distribution as X, the $Y_0, Y_1, \cdots$ are i.i.d. with the same immigration distribution $\{q_k, k \geq 0\}$ and the $X's$ and $Y's$ are independent. It is classic result, see Seneta (1970), for example, that

$$\lim_{n\to\infty} \mathcal{Z}_n/m^n = \mathcal{W}$$

exists and is finite a.s. if and only if

$$\mathbb{E} \log^+ Y < \infty \quad \text{and} \quad \mathbb{E}\,(X \log X) < \infty.$$

where here and throughout, $\log^+ x = \log \max(x, 1) \geq 0$.

Thm: (Chu, Li and Ren (2012))

Assume the $X \log X$ and $\log Y$ conditions and $p_0 = 0$.
(a) If $0 < q_0 < 1$, then

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \asymp \varepsilon^{|\log q_0|/\log m} \quad \text{as} \quad \varepsilon \to 0^+.$$

(b) If $q_0 = 0$ and $p_1 > 0$, then

$$\log \mathbb{P}(\mathcal{W} \le \varepsilon) \sim -\frac{\kappa |\log p_1|}{2(\log m)^2} \cdot |\log \varepsilon|^2, \quad \text{as} \quad \varepsilon \to 0^+,$$

with $\kappa = \inf\{n : q_n > 0\}$.
(c) If $q_0 = 0$ and $p_1 = 0$, then

$$\log \mathbb{P}(\mathcal{W} \le \varepsilon) \asymp -\varepsilon^{-\beta/(1-\beta)}, \quad \text{as} \quad \varepsilon \to 0^+,$$

with β being defined as in the case without immigration.
(d) If $p_0 > 0$, then

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \asymp \varepsilon^{|\log h(\rho)|/\log m}, \quad \text{as} \quad \varepsilon \to 0^+,$$

where ρ is the solution of $f(s) = s$ between $(0, 1)$, and h is the generating function of immigration.

•The asymptotic $\asymp$ is best possible in the sense that it can not be improved into the more precise asymptotic $\sim$.

•The oscillation occurs with immigration even there is no oscillation without immigration. This is quite unexpected and demonstrates the significant effects of the immigration.

Basic Ideas of Proofs

•Built on Dubuc's result without extending the involved analytic method used.
•Start with the very useful probabilistic approach of Mörters and Ortgiese (2008), the branching tree heuristic.
•Develop additional powerful arguments to overcome difficulties of immigration effects.
•Decomposition as infinite sums in distribution, truncation, exponential Chebyshev's inequality, estimates of Laplace transform, Tauberian and exponential Tauberian Theorems.

Fundamental Decomposition

For fixed integer $r \geq 0$ and $l \geq 1$, let $\xi_r(1), \cdots, \xi_r(\mathcal{Z}_r)$ be the individuals in generation r, and $\eta_l(j), j = 1, \cdots, Y_l$ be the individuals of immigration in generation l. Then for any $r \geq 0$ and $n \geq r+1$,

$$\mathcal{Z}_n = \sum_{i=1}^{\mathcal{Z}_r} Z_{n-r}(\xi_r(i)) + \sum_{l=r+1}^{n} \sum_{j=1}^{Y_l} Z_{n-l}(\eta_l(j)).$$

Here $(Z_n(v), n \geq 0)$ is a supercritical G-W branching process initiated with one individual v and $W(v)$ is the limit of positive martingale $m^{-n} Z_n(v)$.

Divided by m^n on both sides, then let $n \to \infty$, we get

$$\mathcal{W} = m^{-r} \sum_{i=1}^{\mathcal{Z}_r} W(\xi_r(i)) + \sum_{l=r+1}^{\infty} m^{-l} \sum_{j=1}^{Y_l} W(\eta_l(j)). \qquad (0.8)$$

For simplicity, we rewrite (0.8) as

$$\mathcal{W} = m^{-r} \sum^{\mathcal{Z}_r} W_i + \sum^{\infty} m^{-l} \sum^{Y_l} W_l^j. \qquad (0.9)$$

Review

Lemma: Assume V is a positive random variable and $\alpha > 0$ is a constant.
(i) For constant $C > 0$,

$$\mathbb{E}\, e^{-\lambda V} \sim C\lambda^{-\alpha} \qquad \textit{as } \lambda \to \infty,$$

if and only if

$$\mathbb{P}(V \le t) \sim \frac{C}{\Gamma(1+\alpha)} t^{\alpha} \qquad \textit{as } t \to 0^+.$$

(ii) The one-sided relation

$$\mathbb{P}(V \le t) \le C_1 t^{\alpha} \quad \text{for some constant } C_1 > 0 \text{ and all } t > 0$$

is equivalent to

$$\mathbb{E}\, e^{-\lambda V} \le C_2 \lambda^{-\alpha} \quad \text{for some constant } C_2 > 0 \text{ and all } \lambda > 0.$$

Review

Lemma: Assume V is a positive random variable and $\alpha > 0, \theta \in \mathbb{R}$, or $\alpha = 0, \theta > 0$ are constants.
(i) For constant $C > 0$,

$$\log \mathbb{P}(V \leq t) \sim -Ct^{-\alpha} |\log t|^{\theta} \qquad as\ t \to 0^+,$$

if and only if

$$\log \mathbb{E}\, e^{-\lambda V} \sim -(1+\alpha)^{1-\theta/(1+\alpha)} \alpha^{-\alpha/(1+\alpha)} C^{1/(1+\alpha)} \lambda^{\alpha/(1+\alpha)} (\log \lambda)^{\theta/(1+\alpha)}$$

(ii) The one-sided relation

$$\log \mathbb{P}(V \leq t) \leq -C_1 t^{-\alpha} |\log t|^{\theta} \quad \text{for some constant } C_1 > 0 \text{ and all } t > 0$$

is equivalent to

$$\log \mathbb{E}\, e^{-\lambda V} \leq -C_2 \lambda^{\alpha/(1+\alpha)} (\log \lambda)^{\theta/(1+\alpha)} \quad \text{for some constant } C_2 > 0 \text{ and}$$

Lemma: Under condition $\mathbb{E} \log^+ Y < \infty$, for any fixed constant $\delta > 0$, there exists integer l such that

$$\mathbb{P}(\max_{i \geq l+1} Y_i e^{-\delta i} \leq 1) \geq e^{-1}.$$

Basic ideas for Lower bounds

For any $\varepsilon > 0$, let $k = k_\varepsilon$ be the integer such that

$$m^{-k} \leq \varepsilon < m^{-k+1}, \tag{0.10}$$

which is equivalent to say that

$$k - 1 < |\log \varepsilon| / \log m \leq k, \quad \text{or} \quad k = \lceil |\log \varepsilon| / \log m \rceil. \tag{0.11}$$

Using the fundamental distribution identity (0.9) with $r = 0$, we have for a fixed integer l to be chosen later,

$$\begin{aligned} \mathbb{P}(\mathcal{W} \leq \varepsilon) &= \mathbb{P}\Big(\sum_{i=0}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \varepsilon \Big) \\ &\geq \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2} \Big) \cdot \mathbb{P}\Big(\sum_{i=k+l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2} \Big) \end{aligned}$$

For the second term, we have by using $\varepsilon \geq m^{-k}$ in (0.10),

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=k+l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2} \Big) &\geq \mathbb{P}\Big(\sum_{i=k+l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{m^{-k}}{2} \Big) \\
&= \mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{1}{2} \Big).
\end{aligned}$$

Note that the last equality follows from the independence and identical distribution of all W_i^j's and Y_i's.
Next we have by controlling the size of Y_i, $i \geq l+1$.

$$\begin{aligned}
&\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{1}{2} \Big) \\
\geq\ &\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{1}{2}, \max_{i \geq l+1} Y_i e^{-\delta i} \leq 1 \Big) \\
\geq\ &\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i} \sum_{j=1}^{e^{\delta i}} W_i^j \leq \frac{1}{2} \Big) \cdot \mathbb{P}\Big(\max_{i \geq l+1} Y_i e^{-\delta i} \leq 1 \Big).
\end{aligned}$$

Using Chebyshev's inequality for the first part, we get

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{e^{\delta i}} W_i^j \leq \frac{1}{2}\Big) &\geq 1-2\mathbb{E}\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{e^{\delta i}} W_i^j \\ &= 1-\frac{2e^{\delta(l+1)}}{(m-e^{\delta})m^l}.\end{aligned}$$

We can now choose δ such that $e^{\delta} < m$, and then find large enough integer l so that

$$\frac{2e^{\delta(l+1)}}{(m-e^{\delta})m^l} < \frac{1}{2}.$$

Combining all estimates together, we obtain that

$$\mathbb{P}\Big(\sum_{i=k+l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2}\Big) \geq \mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \leq \frac{1}{2}\Big) \geq \frac{1}{2e}.$$

Now back to the first part, we have to handle it under conditions (a) and (b) separately. In the case (a) with $q_0 > 0$, we have the simple estimate

$$\mathbb{P}\left(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2} \right) \geq \mathbb{P}\big(Y_0 = \cdots = Y_{k+l} = 0 \big) = q_0^{k+l+1}.$$

Using $k - 1 < |\log \varepsilon| / \log m$, it's easy to deduce that

$$q_0^k \quad \geq \quad q_0 \cdot q_0^{|\log \varepsilon| / \log m} = q_0 \varepsilon^{|\log q_0| / \log m}.$$

which finishes the lower bound under condition $q_0 > 0$.

For the case (b) with $q_0 = 0$, we have, recalling the definition of $K = \inf\{n : q_n > 0\}$,

$$\begin{aligned} & \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{Y_i} W_i^j \le \frac{\varepsilon}{2} \Big) \\ \ge \ & \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{Y_i} W_i^j \le \frac{\varepsilon}{2}, Y_0 = \cdots = Y_{k+l} = K \Big) \\ = \ & \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{K} W_i^j \le \frac{\varepsilon}{2} \Big) \cdot q_K^{k+l+1}. \end{aligned}$$

The above probability of sums can be bounded termwise, and thus

$$
\begin{aligned}
&\mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{K} W_i^j \le \frac{\varepsilon}{2} \Big) \\
\ge\ & \mathbb{P}\Big(\max_{0\le i\le k+l} \max_{1\le j\le K} m^{-i} W_i^j \le \frac{\varepsilon/2}{K(k+l+1)} \Big) \\
=\ & \prod_{i=0}^{k+l} \mathbb{P}^K \Big(m^{-i} W \le \frac{\varepsilon/2}{K(k+l+1)} \Big) \\
\ge\ & \prod_{i=0}^{k+l} \mathbb{P}^K \Big(W \le \frac{m^{i-k}/2}{K(k+l+1)} \Big).
\end{aligned}
$$

where we used the independence of all W_i^j's in the last equality and $\varepsilon \ge m^{-k}$ in the last inequality.

Basic ideas for upper bounds

As we can see from the arguments before, only the finite terms are contributing to the small value probabilities of $\mathcal{W}$. Hence we take only $r = 0$ in (0.9), choose suitable cut off k, and focus on properties of $\sum_{l=0}^{k} m^{-l} \sum_{j=1}^{Y_l} W_l^j$.

Let $k = k_\varepsilon$ be the integer defined as in (0.10). Using the fundamental distribution identity (0.9) with $r = 0$ and exponential Chebyshev's inequality, we have for any $\lambda > 0$,

$$
\begin{aligned}
\mathbb{P}(\mathcal{W} \le \varepsilon) &\le \mathbb{P}\Big(\sum_{i=0}^{k} m^{-i} \sum_{j=1}^{Y_i} W_i^j \le \varepsilon \Big) \\
&\le e^{\lambda\varepsilon} \cdot \mathbb{E} \exp\Big(-\lambda \sum_{i=0}^{k} m^{-i} \sum_{j=1}^{Y_i} W_i^j \Big).
\end{aligned}
$$

Notice that all the $(W_i^j, Y_i, i=0,\cdots,k, j=1,\cdots)$ are independent, we have

$$\mathbb{E}\exp\Big(-\lambda\sum_{i=0}^{k} m^{-i}\sum_{j=1}^{Y_i} W_i^j\Big) = \prod_{i=0}^{k}\mathbb{E}\exp\Big(-\lambda m^{-i}\sum_{j=1}^{Y_i} W_i^j\Big).$$

Conditioning on $Y_i = 0$ or $Y_i \geq 1$, we have

$$\begin{aligned} & \mathbb{E}\exp\Big(-\lambda m^{-i}\sum_{j=1}^{Y_i} W_i^j\Big) \\ \leq\ & q_0 + (1-q_0)\mathbb{E}\exp\big(-\lambda m^{-i} W_i^1\big) \leq q_0(1+\delta_i), \end{aligned}$$

where

$$\delta_i = q_0^{-1}\mathbb{E}\exp\big(-\lambda m^{-i} W_i^1\big) = q_0^{-1}\mathbb{E}\exp\big(-\lambda m^{-i} W\big),\ \ i=0,\cdots,k.$$

Substituting things in and letting $\lambda = \varepsilon^{-1}$, we obtain

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \le e q_0^{k+1} \prod_{i=0}^{k} (1 + \delta_i).$$

Since $k \ge |\log \varepsilon| / \log m$ in (0.11), we have

$$q_0^k \le \varepsilon^{|\log q_0| / \log m}.$$

So we finish the proof by showing

$$\sum_{i=0}^{k} \log(1 + \delta_i) \le \sum_{i=0}^{k} \delta_i \le M \tag{0.12}$$

where $M > 0$ is a constant independent of ε (noticing that the k depends on ε).

Next we turn to consider a slightly different type of supercritical branching process with immigration, which is denoted by $(\widetilde{\mathcal{Z}}_n, n \geq 0)$. The only difference is to assume $\widetilde{\mathcal{Z}}_0 = 1$. The corresponding limit of $\widetilde{\mathcal{Z}}_n/m^n$ is denoted by $\widetilde{\mathcal{W}}$. Then by simple computation we get that

$$\widetilde{\mathcal{W}} =^d W + \frac{\mathcal{W}}{m} \tag{0.13}$$

in distribution, as denoted by $=^d$ throughout this paper. Due to (0.13) and the fact that

$$\begin{aligned}
\mathbb{P}(W + \mathcal{W}/m \leq \varepsilon) &\geq \mathbb{P}(W \leq \varepsilon/2) \cdot \mathbb{P}(\mathcal{W}/m \leq \varepsilon/2), \\
\mathbb{P}(W + \mathcal{W}/m \leq \varepsilon) &\leq \mathbb{P}(W \leq \varepsilon) \cdot \mathbb{P}(\mathcal{W}/m \leq \varepsilon),
\end{aligned} \tag{0.14}$$

we can obtain the following result as a consequence of combining two early results.

Assume $p_0 = 0$.

(a) If $p_1 > 0$ and $q_0 > 0$, then

$$\mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \asymp \varepsilon^{|\log(p_1 q_0)|/\log m} \qquad \textit{as } \varepsilon \to 0^+.$$

(b) If $p_1 > 0$ and $q_0 = 0$, then

$$\log \mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \sim -\frac{\kappa|\log p_1|}{2(\log m)^2}|\log \varepsilon|^2 \qquad \textit{as } \varepsilon \to 0^+,$$

(c) If $p_1 = 0$, then

$$\log \mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \asymp -\varepsilon^{-\beta/(1-\beta)} \qquad \textit{as } \varepsilon \to 0^+,$$

(d) If $p_0 > 0$, then

$$\mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \asymp \varepsilon^{|\log h(\rho)|/\log m}, \qquad \textit{as } \varepsilon \to 0^+.$$