Upper bound for the tail functions of the growth rate for supercritical branching processes in random environment

Yinna Ye*

Abstract. Suppose that \((Z_n)_{n \geq 0}\) is a supercritical branching process in independent and identically distributed random environment. We study the right tail function of the scaled growth rate for \((Z_n)_{n \geq 0}\) and establish an Hoeffding type inequality.

Keywords: Branching processes; Random environment; Hoeffding’s inequality

2000 MSC: Primary 60J80; 60K37; 60G42

1 Introduction and setting

The branching process in random environment (BPRE) has become a hot topic since 1970’s. One of the best-known works since then are those developed by Athreya and Karlin in 1971 ([1], [2]), based on the Smith and Wilkinson’s model ([15]) where the random environment process is supposed to be independent and identically distributed (i.i.d.) random variables (r.v.s), they extend their model to more general situations of the random environment like e.g. stationary and metrically transitive random process or Markov chain. They found ‘extinction or explosion result’ for BPRE and delimited complete criteria for certainty or noncertainty of extinction and clarified the three classes: subcritical, critical and supercritical BPRE for the first time. Then Tanny [16] improved the conditions of ‘extinction or explosion result’ and studied the rate of growth of the BPRE especially when the branching process (BP) is supercritical. Later on, based on these results, a series of papers appeared and studied especially BPs in i.i.d. random environment. Their focus are mainly in two directions: one is on the asymptotic of the survival probability for subcritical BPs, the typical papers are [12], [13], [14] and [15]; another one is on the large deviations of supercritical BPs, the typical papers are [15], [16], [14] and [13].

However, so far the Hoeffding type inequalities for BPRE has not been studied yet in the literature. If \((Y_i)_{1 \leq i \leq n}\) is a sequence of centered \((\mathbb{E}(Y_i) = 0)\) r.v.s with finite variance. The Hoeffding type inequalities provide upper bounds for the right tail function of \(\sum_{i=1}^{n} Y_i\). Particularly, if \((Y_i)_{1 \leq i \leq n}\) are independent r.v.s satisfying \(Y_i \leq 1\), the classical Hoeffding inequality, initially developed by Bennett [3] and Hoeffding [10], is in the following form: for any \(t > 0\),

\[
P \left( \sum_{i=1}^{n} Y_i \geq nt \right) \leq \left( \frac{\tau^2}{t + \tau^2} \right)^n e^{nt},
\]

*Department of Mathematical Sciences, Xi’an Jiaotong-Liverpool University, China. E-mail: yinna.ye@xjtlu.edu.cn
where $\tau^2 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(Y_i^2)$. More recently, Fan et al. \cite{5} improved the classical upper bound and extended their results to the case when $(\sum_{i=1}^{n} Y_i)_n$ is a supermatingale with differences bounded above.

In this paper, we study the right tail function of the growth rate for supercritical BPs in i.i.d. random environment and give an upper bound, by using the Hoeffding type inequality developed by Fan et al. \cite{5}. The paper is organized as follows. In the rest of this section, the BPs in i.i.d. random environment and the assumptions will be introduced. The main result Theorem 2.4 and its proof are given in Section 2. In Section 3 some examples such as binary branching and binomial branching in i.i.d. random environment will be discussed to show the feasibility.

Consider a branching process (BP), denoted by $(Z_n)_{n\geq 0}$, evolving in i.i.d. random environment $\xi = (\xi_0, \xi_1, \cdots)$ with common distribution $\nu$. The BPs in i.i.d. environment can be defined as follows. For instance, one may also refer to \cite{11} and \cite{8} for the definition. Let $Z_n$ denote the number of individuals at the $n$th generation in a family tree and $Z_n$ satisfies the following recursive form:

$$Z_0 = 1, \quad Z_{n+1} = \sum_{i=1}^{Z_n} N_{n,i}, \text{ for } n \geq 0,$$

where $N_{n,i}$ represents the family size of the $i$th father individual located in the $n$th generation. Given the environment $\xi_n$ in the $n$th generation, $(N_{n,i})_{i \geq 1}$ is a sequence of i.i.d. r.v.s with the conditional probability mass function (p.m.f.) $(p_k(\xi_n))_{k \geq 0}$. Because the one-to-one correspondence between p.m.f. and probability distribution, sometimes people also call $(p_k(\xi_n))_{k \geq 0}$ the (conditional) offspring distribution in the ($n+1$)th generation. Suppose that given the environment sequence $\xi$, every family size $(N_{n,i})_{n \geq 0}$ are conditionally independent from generation to generation. Let

$$m_n := \sum_{k=0}^{+\infty} kp_k(\xi_n), \text{ for } n \geq 0;$$

$$\Pi_n := m_0 \cdots m_{n-1}, \text{ for } n \geq 1 \text{ and } \Pi_0 = 1;$$

$$S_n := \log \Pi_n = \sum_{i=1}^{n} X_i, \text{ for } n \geq 1,$$

where for $i \geq 1$, $X_i := \log m_{i-1}$ is the log-conditional mean of the family size of a father individual located in the ($i-1$)th generation given the environment $\xi_{i-1}$ in that generation. Then we have

$$\log Z_n = S_n + \log W_n,$$

from which it can be seen that $\log Z_n$ is decomposed into two components: $S_n$ and $\log W_n$. It is well known (see also \cite{11} and \cite{8} for instance) that $(S_n)_{n \geq 1}$ is a random walk with i.i.d increments and $(W_n)_{n \geq 1}$ is a non-negative martingale under both quenched and annealed laws.
denoted by $\mathbb{P}_\xi$ and $\mathbb{P}$ respectively, with respect to (w.r.t) the filtration $(\mathcal{F}_n)_{n \geq 1}$, where $\mathcal{F}_n$ is given by

$$\mathcal{F}_n = \sigma(\xi, N_{k,i}, 0 \leq k \leq n - 1, i = 1, 2, \ldots).$$

Moreover,

$$W = \lim_{n \to +\infty} W_n \quad \text{exists} \quad \mathbb{P} - \text{a.s.}$$

and

$$\mathbb{E}(W) \leq 1.$$

Let $\mu = \mathbb{E}(X_1)$ and $\sigma^2 = \text{Var}(X_1)$. Throughout the paper, we will use $\nu(\cdot)$, $\mathbb{E}_\xi$ and $\mathbb{E}$ to present the expectations w.r.t the probability distributions $\nu$, $\mathbb{P}_\xi$ and $\mathbb{P}$ respectively. We denote by $C$ an absolute constant whose value may differ from line to line. And we assume the following basic assumptions:

1) $0 < \mu < +\infty$ and $\mathbb{E}|\log(1 - p_0(\xi_0))| < +\infty$,

which implies the population size tends to $+\infty$ with positive probability.

2) $0 < \sigma^2 < +\infty$,

which implies that $\mathbb{P}(Z_1 = 1) = \mathbb{E}(p_1(\xi_0)) < 1$.

3) $\mathbb{E}\left(\frac{Z_1 \log^+ Z_1}{m_0}\right) < +\infty,$

which is a necessary and sufficient condition (by Theorem 2 in [16]) to imply that $W_n \to W$ in $L^1$, as $n \to +\infty$; and

$$\mathbb{P}(W > 0) = \mathbb{P}(Z_n \to +\infty) = \lim_{n \to +\infty} \mathbb{P}(Z_n > 0) > 0.$$

We need the following two more hypothesis:

**H1)** Suppose $\exists M > 0, \text{ s.t. } \forall i \geq 1,$

$$\frac{X_i - \mu}{M} \leq 1 \quad \text{a.s.}$$

in other words, $X_i$ is almost surely a bounded random variable.

**H2)** There exist $p > 1$ and $q > 2$, s.t.

$$\mathbb{E}\left(\frac{Z_1}{m_0}\right)^p < +\infty \quad \text{and} \quad \mathbb{E}(|\log m_0|^q) < +\infty.$$

For BPs in i.i.d random environment, it is easy to arrange $(p_k(\xi_0))_k$ to satisfy the assumption 1), as long as the conditional offspring distribution is not concentrated on $k = 0, 1$ and its conditional variance exists, for almost all environments $\xi_0$’s in the 0th (initial) generation. In other words, if the conditions below hold, then the assumptions 1) and 2) will be both satisfied.

(C1) $\sum_{k \geq 1} k^2 p_k(\xi_0) < +\infty, \quad \exists k \geq 2, \text{ s.t. } p_k(\xi_0) > 0 \text{ and } m_0 > 1, \nu - \text{a.s.}$

Under (C1), a sufficient condition for the assumption 3) is the following

(C2) $\sum_{k \geq 2} (k \log k) \nu |p_k(\xi_0)| < +\infty,
because $m_0 > 1$, $\nu - a.s.$ and
\[
\mathbb{E}\left[\frac{Z_1 \log Z_1}{m_0}\right] = \sum_{k \geq 2} (k \log k) \nu \left[\frac{p_k(\xi_0)}{m_0}\right] \leq \sum_{k \geq 2} (k \log k) \nu [p_k(\xi_0)].
\]

The assumption H1) can be attained if and only if for any $i \geq 1$,
\[
\sum_{k=0}^{+\infty} k p_k(\xi_{i-1}) \leq C, \quad \nu - a.s.
\]
From above, if (C1) and (C2) hold, then all the assumptions 1) - 3), H1) and H2) will be satisfied. One may find further discussion in Section 3 for two particular examples.

Let
\[
\tilde{Z}_n := \log Z_n - n\mu \sqrt{n}M = \tilde{X}_n + V_n,
\]
where $\tilde{X}_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{M}\right)$ constitutes a centered random walk with i.i.d. increments,
\[
V_m := \log W_m \sqrt{n}M \quad \text{for } 1 \leq m \leq n,
\]
and the last equality above is obtained by \ref{eq:4}. Throughout the paper, we denote by $C$ an absolute constant whose value may differ from line to line.

2 Main results and proofs

We denote by $\mathbb{1}$ the indicator function. Applying the Hoeffding type inequality in Theorem 2.1, \cite{5} to $\tilde{X}_n$, we have

\textbf{Theorem 2.1} (\cite{5}). Under the assumption H1), for any $x \geq 0$ and $n \geq 1$,
\[
\mathbb{P}\left[\sum_{i=1}^{n} \left(\frac{X_i - \mu}{M}\right) \geq x\right] \leq H_n(x, v_n),
\]
where $v_n = \sqrt{\sum_{i=1}^{n} \mathbb{E} \left(\frac{X_i - \mu}{M}\right)^2} = \frac{\sqrt{n\sigma}}{M}$ and $H_n$ is a function defined by
\[
H_n(x, v) := \left[\left(\frac{v^2}{x + v^2}\right)^{x + v^2} \left(\frac{n}{n - x}\right)^{n - x}\right]^{\frac{n}{x + v^2}} \mathbb{1}\{x \leq n\},
\]
for any $x \geq 0$ and $v > 0$.

\textbf{Corollary 2.2}. Under the assumption H1), for any $x \geq 0$,
\[
\mathbb{P}(\tilde{X}_n \geq x) \leq H_n \left(\sqrt{nx}, \frac{\sqrt{n\sigma}}{M}\right).
\]
And we have the following elementary lemma about the function $H_n$ which is useful to prove the main result Theorem 2.4.
Lemma 2.3. For any $x \geq 0$ and $v > 0$, the function $H_n(x, v)$ is non-increasing in $x$.

Proof. By the definition of $H_n$, when $0 \leq x \leq n$,
\[
\log H_n(x, v) = \frac{n}{n + v^2} \left\{ (x + v^2)[2 \log v - \log(x + v^2)] + (n - x)[\log n - \log(n - x)] \right\}.
\]

Taking partial derivative in both sides w.r.t $x$, we can obtain
\[
\frac{\partial}{\partial x} H_n(x, v) = \frac{n}{n + v^2} H_n(x, v) \log \left[ \frac{v^2(n - x)}{n(v^2 + x)} \right] \leq 0,
\]

since for any $0 \leq x \leq n$ and $v > 0$, $H_n(x, v) > 0$ and
\[
0 < \frac{v^2(n - x)}{n(v^2 + x)} \leq \frac{nv^2}{nv^2 + nx} \leq 1.
\]

And from the inequality above, the equality in (2) is attained only when $x = 0$. In the case when $x > n$, we have $H_n(x, v) \equiv 0$. We thus prove the desired result.

We want to find the upper bound for the tail function $P(\tilde{Z}_n \geq x)$. Indeed, we have

Theorem 2.4. Under the assumptions H1) and H2), there exist constants $C > 0$, $\delta \in (0, 1)$ and $x_1 \geq 0$ s.t. for $1 \leq m \leq n$ and $x \in [x_1, +\infty)$,
\[
P\left( \frac{\log Z_n - n\mu}{\sqrt{nM}} \geq x \right) \leq H_n\left( \sqrt{nx}, \sqrt{n\sigma M} \right) + 4H_n\left( \sqrt{n - m} (x - x_1), v_{m,n} \right) + C\delta^m,
\]

where $v_{m,n} = \frac{\sqrt{n - m} \sigma}{M}$.

In particular, if $m = np$ with $p \in (0, 1)$, then
\[
P\left( \frac{\log Z_n - n\mu}{\sqrt{nM}} \geq x \right) \leq H_n\left( \sqrt{nx}, \sqrt{n\sigma M} \right) + 4H_n\left( \sqrt{n - np} (x - x_1), \frac{\sqrt{n - np} \sigma}{M} \right) + C\delta^{np}.
\]

We can replace $\sqrt{n}$ by $n$ in the denominator of the tail functions $P\left( \frac{\log Z_n - n\mu}{\sqrt{nM}} \geq \cdot \right)$ above, on contrary to what is usually done by the others (see \[11\] and \[8\] for instance), subject to the scale level or the subjects of interest.

From the perspective on the rate of convergence in the central limit theorem for log $Z_n$, the Theorem 1.1 in \[8\] about the Berry-Esseen’s bound provides both lower and upper bounds for $|P[ (\log Z_n - n\mu)/\sigma \sqrt{n} - \Phi(x)]|$, where $\Phi$ is the standard normal distribution function. This theorem implies that there exists a constant $\delta_2 \in (0, 1)$, s.t. $\forall x \in [\frac{xn}{M}, +\infty)$,
\[
P\left[ \frac{\log Z_n - n\mu}{\sqrt{nM}} \geq x \right] \leq \overline{F}\left( \frac{Mx}{\sigma} \right) + \frac{C}{n^{\delta_2/2}},
\]

where $\overline{F}(x) := 1 - \Phi(x)$ is the right tail function of the standard normal distribution. In comparison between \[3\] and \[4\], it can be seen that $\delta^{np} = o\left( n^{-\delta_2/2} \right)$, as $n \to +\infty$, given the facts
\( F(x) \leq 1 \) and \( \lim_{n \to +\infty} H_n(x, v) = \left( \frac{v^2}{x + v^2} \right)^{x+v^2} \exp(x) \leq \exp \left[ -\frac{x^2}{2(v^2 + \frac{1}{3} x)} \right] \leq 1, \) for any \( x \geq 0 \) and \( v > 0 \) (see also Remark 2.1 in [5] for the last inequalities). Therefore, our upper bound in (3) is sharper and improves the one implied by Berry-Esseen’s in [3].

The proof of Theorem 2.4 is inspired from that of Theorem 1.1 in [8].

Since for any \( x \geq 0 \) and \( n \geq 1 \),
\[
\mathbb{P}(\tilde{Z}_n \geq x) = \mathbb{P}(\tilde{X}_n \geq x) + \mathbb{P}(\tilde{Z}_n \geq x, \tilde{X}_n < x) - \mathbb{P}(\tilde{Z}_n < x, \tilde{X}_n \geq x).
\]
Therefore,
\[
\mathbb{P}(\tilde{Z}_n \geq x) \leq \mathbb{P}(\tilde{X}_n \geq x) + \left| \mathbb{P}(\tilde{Z}_n \geq x, \tilde{X}_n < x) - \mathbb{P}(\tilde{Z}_n < x, \tilde{X}_n \geq x) \right|.
\]
(5)
So we need to study the joint distribution of \( (\tilde{Z}_n, \tilde{X}_n) \). We have the following lemma.

**Lemma 2.5.** Under assumptions H1) and H2), there exist constants \( C > 0, \delta \in (0, 1) \), and \( x_1 \geq x_0 \), s.t. for \( 1 \leq m \leq n \),

1) \( \mathbb{P}(\tilde{Z}_n < x, \tilde{X}_n \geq x) \leq 2H_{n-m} \left( \sqrt{n-m}(x-x_0), v_{m,n} \right) + C\delta^m, \ \forall x \in [x_0, +\infty), \)

2) \( \mathbb{P}(\tilde{Z}_n \geq x, \tilde{X}_n < x) \leq 2H_{n-m} \left( \sqrt{n-m}(x-x_1), v_{m,n} \right) + C\delta^m, \ \forall x \in [x_1, +\infty). \)

**Proof of Theorem 2.4.** From (5), Corollary 2.2 and Lemma 2.5 we have for \( x \geq x_1 \),
\[
\mathbb{P}(\tilde{Z}_n \geq x) \leq H_n \left( \sqrt{n}x, \frac{\sqrt{n}\sigma}{M} \right) + 2 \left[ H_{n-m} \left( \sqrt{n-m}(x-x_0), v_{m,n} \right) + H_{n-m} \left( \sqrt{n-m}(x-x_1), v_{m,n} \right) \right] + C\delta^m.
\]
By Lemma 2.3 and Remark 2.1 in [5], \( H_n(x, v) \) is \( \downarrow \) in \( x \) and \( \uparrow \) in \( n \), so
\[
\mathbb{P}(\tilde{Z}_n \geq x) \leq H_n \left( \sqrt{n}x, \frac{\sqrt{n}\sigma}{M} \right) + 2 \left[ H_n \left( \sqrt{n-m}(x-x_0), v_{m,n} \right) + H_n \left( \sqrt{n-m}(x-x_1), v_{m,n} \right) \right] + C\delta^m
\]
\[
\leq H_n \left( \sqrt{n}x, \frac{\sqrt{n}\sigma}{M} \right) + 4H_n \left( \sqrt{n-m}(x-x_1), v_{m,n} \right) + C\delta^m.
\]
Letting \( m = n^p \), we can obtain the second inequality in this theorem.

Henceforth our main task is to prove Lemma 2.5. Before proving it, let’s introduce some notations. Set
\[
Y_{m,n} := \frac{1}{\sqrt{n}} \sum_{i=m+1}^{n} \frac{X_i - \mu}{M}, \text{ for } 0 \leq m \leq n - 1;
\]
\[
Y_n := Y_{0,n};
\]
and
\[
D_{m,n} := V_n - V_m, \text{ for } 1 \leq m \leq n - 1.
\]
Then we have

$$\tilde{X}_n = Y_n$$

and

$$\tilde{Z}_n = Y_n + V_n.$$  

**Proof of Lemma 2.5.** Using the notations above, we have

$$P(\tilde{Z}_n < x, \tilde{X}_n \geq x) = P(Y_n + V_n < x, Y_n \geq x)$$

$$\leq P(Y_n + V_m < x + \frac{1}{\sqrt{n}}, Y_n \geq x) + P(|D_{m,n}| > \frac{1}{\sqrt{n}})$$

$$:= I_1 + I_2.$$  

(6)

Since $Y_n = Y_m + Y_{m,n}$,

$$I_1 = P(Y_m + Y_{m,n} + V_m < x + \frac{1}{\sqrt{n}}, Y_m + Y_{m,n} \geq x)$$

$$= \int P(Y_{m,n} + s + t < x + \frac{1}{\sqrt{n}}, Y_{m,n} + s \geq x) \nu_m(ds, dt),$$

where $\nu_m(ds, dt)$ is the joint distribution of $(Y_m, V_m)$. By conditioning on $(Y_m, V_m)$ and the independence between $Y_{m,n}$ and $(Y_m, V_m)$, for $0 \leq m \leq n - 1$, the equality above becomes

$$I_1 = \int \mathbb{1}_{\{t \leq \frac{1}{\sqrt{n}}\}} \left[ P(Y_{m,n} \geq x - s) - P(Y_{m,n} \geq x - s - t + \frac{1}{\sqrt{n}}) \right] \nu_m(ds, dt).$$  

(7)

By the identical distribution of the random environment, $Y_{m,n}$ follows the same distribution as $Y_{n-m}$, for $0 \leq m \leq n - 1$. Moreover, by the assumption H1), we have

$$Y_{n-m} \leq \sqrt{\frac{n-m}{n}} \ a.s.$$  

Therefore, we can replace $Y_{m,n}$ in (7) by $Y_{n-m}$ to obtain

$$I_1 = \int \mathbb{1}_{\{t \leq \frac{1}{\sqrt{n}}\}} \left[ P(Y_{n-m} \geq x - s) - P(Y_{n-m} \geq x - s - t + \frac{1}{\sqrt{n}}) \right] \nu_m(ds, dt).$$

Applying Theorem 2.4 to $Y_{n-m}$, we have for $x \in [x_0, +\infty)$, with $x_0 := \max(s, t + s - \frac{1}{\sqrt{n}})$,

$$P(Y_{n-m} \geq x - s) \leq H_{n-m} \left( \sqrt{n-m} (x-s), \frac{\sqrt{n-m} \sigma}{M} \right)$$

and

$$P(Y_{n-m} \geq x - s - t + \frac{1}{\sqrt{n}}) \leq H_{n-m} \left( \sqrt{n-m} (x-s - t + \frac{1}{\sqrt{n}}), \frac{\sqrt{n-m} \sigma}{M} \right).$$

By Lemma 2.3 and Remark 2.1 in [5], $H_n(x, v)$ is ↓ in $x$ and ↑ in $n$, we thus have

$$P(\tilde{Z}_n < x, \tilde{X}_n \geq x) \leq 2H_{n-m} \left( \sqrt{n-m} (x-x_0), \frac{\sqrt{n-m} \sigma}{M} \right) + I_2$$

$$\leq 2H_n \left( \sqrt{n-m} (x-x_0), \frac{\sqrt{n-m} \sigma}{M} \right) + I_2.$$  

7
The next step is to find an upper bound for $I_2$. By Markov’s inequality,

$$I_2 = P(|V_n - V_m| > \frac{1}{\sqrt{n}}) = P\left(\frac{1}{M} |\log W_n - \log W_m| > 1\right) \leq \frac{1}{M} E |\log W_n - \log W_m|.$$ 

We need the following Lemma.

**Lemma 2.6** ([8]). Under assumption H2), there exist constants $c > 0$ and $\delta \in (0, 1)$, s.t. for $n \geq 0$,

$$E |\log W_{n+1} - \log W_n| \leq c \delta^n.$$

Using Lemma above, we obtain that there exist constants $\delta \in (0, 1)$ and $c > 0$, s.t. for $0 \leq m \leq n - 1$,

$$E |\log W_n - \log W_m| = \sum_{k=m}^{n-1} E |\log W_{k+1} - \log W_k|$$

$$\leq c \sum_{k=m}^{n-1} \delta^k = \frac{c \delta^m (1 - \delta^{n-m})}{1 - \delta} \leq c' \delta^m,$$

where $c' = \frac{c}{1-\delta}$. The first inequality in Lemma 2.5 is therefore proven with $C = \frac{c'}{M}$. The second one can be proven similarly, with $x_1 := \max(s, t + s + \frac{1}{\sqrt{n}}) \geq x_0$. 

\[\square\]

3 Examples

3.1 Binary branching

A typical example is the binary BP in i.i.d random environment $\bar{\xi}$ with common distribution supported in a finite set. Consider $\nu = \{\nu_k\}_{1 \leq k \leq k_0}$ be a probability distribution with support in the finite set $\{a_1, \ldots, a_{k_0}\} \in (0, 1)^{k_0}$, where $\nu_k$ is the mass on $a_k$ for each $k$. Suppose the random environment of the $(n + 1)$th generation $\xi_n$ depends solely on one single random parameter $P_n \in (0, 1)$, the success rate of a Bernoulli trial, so that we can denote the entire random environment process by $\bar{\xi} = (P_0, P_1, \cdots)$. Suppose $\{P_0, P_1, \cdots\}$ is a sequence of i.i.d r.v.s following the common probability distribution $\nu$. Given the environment, the offspring distribution in each family is Bernoulli. More precisely, given $\bar{\xi}$, the conditional p.m.f. of the family size $N_{k,i}$ is given by

$$\mathbb{P}(N_{k,i} = 1) = P_k, \quad \text{and} \quad \mathbb{P}(N_{k,i} = 2) = 1 - P_k,$$

for any $k \geq 0$ and $i \geq 1$, that is, every family size is either 1 or 2.

Since

$$m_0 = \mathbb{E}(N_{0,1}|P_0) = 1 \times P_0 + 2 \times (1 - P_0) = 2 - P_0 > 1, \quad \nu - a.s.$$ 

and $p_0(\xi_0) = 0$, we have

$$0 < \mu = \nu[\log(2 - P_0)] < +\infty,$$

$$\mathbb{E}[\log(1 - p_0(\xi_0))] = 0,$$

8
\[
\mathbb{E} \left( \frac{Z_1 \log^+ Z_1}{m_0} \right) = \nu \left[ \mathbb{E}_{\xi_0} \left( \frac{Z_1 \log^+ Z_1}{m_0} \right) \right] = 2(\log 2) \nu \left[ \frac{1 - P_0}{2 - P_0} \right] < 2(\log 2) \nu (1 - P_0) < 2 \log 2,
\]
for any \( i \geq 1 \), \( X_i = \log (2 - P_{i-1}) < \log 2 \), \( \nu \) - a.s.,
\[
\mathbb{E} \left( \frac{Z_1}{m_0} \right)^p = \nu \left[ \mathbb{E}_{\xi_0} \left( \frac{Z_1}{m_0} \right)^p \right] = 2^p + (1 - 2^p) \nu (P_0) < 1, \quad \text{for } p > 1
\]
and
\[
\mathbb{E} | \log m_0|^q = \nu | \log (2 - P_0)|^q < (\log 2)^q, \quad \text{for } q > 2.
\]
So the assumptions 1) - 3), H1) and H2) are all fulfilled.

### 3.2 Binomial branching

Consider a BP living in the same kind of random environment as in the example above, with a Binomial(\( N, P \)) offspring distribution in the \((i + 1)\)th generation, where \( P \) is a r.v. parameter and \( N > 1 \) be a given determinist number s.t. \( N \cdot \min\{a_1, \cdots, a_k\} > 1 \). Similar to the example above, let \( \xi = (P_0, P_1, \cdots) \) be the random environment depending sorely on the i.i.d random parameters \((P_i)_{i \geq 0} \in (0, 1)^\mathbb{N})\). Given \( \xi \), the conditional p.m.f. of the family size \( N_{k,i} \) is given by
\[
\mathbb{P}_\xi(N_{k,i} = j) = \binom{N}{j} P_k^j (1 - P_k)^{N - j},
\]
for any \( 0 \leq j \leq N, k \geq 0 \) and \( i \geq 1 \). Then in this case, \( m_0 = NP_0 > 1 \) \( \nu \) - a.s. and we can verify that all the assumptions 1) - 3), H1) and H2) are fulfilled. Indeed, since the environment distribution \( \nu \) is with support in a finite set, it is evident that all the summations in the expectation w.r.t \( \nu \) below are finite summations of bounded terms and more precisely,
\[
0 < \mu = \log N + \nu (\log P_0) < \log N,
\]
\[
\mathbb{E} | \log (1 - \rho_0(\xi_0))| = \nu | \log [1 - (1 - P_0)^N]| < +\infty,
\]
\[
0 < \sigma^2 = \nu | \log P_0 - \nu (\log P_0)|^2 < +\infty,
\]
\[
\mathbb{E} \left( \frac{Z_1 \log^+ Z_1}{m_0} \right) = N^{-1} \sum_{k=2}^{N} (k \log k) \binom{N}{k} \nu \left[ P_0^{k-1} (1 - P_0)^{N-k} \right] < +\infty,
\]
\( \forall i \geq 1 \), \( X_i = \log N + \log P_{i-1} < \log N \), \( \nu \) - a.s.,
\[
\mathbb{E} \left( \frac{Z_1}{m_0} \right)^p = N^{-p} \sum_{k=0}^{N} k^p \binom{N}{k} \nu \left[ P_0^{k-p} (1 - P_0)^{N-k} \right] < +\infty, \quad \text{for } p > 1
\]
and
\[
\mathbb{E} | \log m_0|^q = \nu | \log N + \log P_0|^q < +\infty, \quad \text{for } q > 2.
\]

### Acknowledgements

The author is grateful for Prof. Xiequan Fan from Center for Applied Mathematics, Tianjin University, for a couple of discussions on the classical Hoeffding inequality and especially the recent development of this kind of inequalities.
References

[1] Athreya, K. B. and Karlin, S. On branching processes with random environments I. *Ann. Math. Stat.*, 42 (5) (1971), 1499-1520.

[2] Athreya, K. B. and Karlin, S. On branching processes with random environments II. *Ann. Math. Stat.*, 42 (6) (1971), 1843-1858.

[3] Bennett, G. Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.*, 57 (297) (1962), 33-45.

[4] Boinghoff, C. and Kersting, G. Upper large deviations of branching processes in a random environment - Offspring distributions with geometrically bounded tails. *Stoch. Pro. and Their Appli.*, 120 (2010), 2064-2077.

[5] Fan, X., Gramma, I., Liu, Q. Hoeffding’s inequality for supermartingales. *Stoch. Pro. and Their Appli.*, 122 (10) (2012), 3545-3559.

[6] Geiger, J. and Kersting, G. The survival probability of a critical branching process in random environment, *Ther. Verojatnost. iPrimenen*, 45 (2000), 607-615.

[7] Geiger, J., Kersting, G. and Vatutin, V. A. Limit theorems for subcritical branching process in random environment, *Ann. I. H. Poincaré*, 39 (4) (2003), 593-620.

[8] Gramma, I., Liu, Q. Miqueu, E. Berry-Esseen’s bound and Cramèr’s large deviation expansion for a supercritical branching processes in a random environment. *Stoch. Pro. and Their Appli.*, 127 (4) (2017), 1255-1281.

[9] Guivarc’h, Y. and Liu, Q. Asymptotic properties of braching processes in random environ-ment, *C. R. Mathematique*, 332 (4) (2001), 339-344.

[10] Hoeffding, W. Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.*, bf 58 (1963), 13-30.

[11] Huang, C., Liu, Q. Moments, moderate and large deviations for a branching process in a random environment. *Stoch. Pro. and Their Appli.*, 122 (2) (2012), 522-545.

[12] Kozlov, M.V. On the asymptotic behavior of the probability of non-extinction for critical branching processes in a random environment. *Theory Probab*, 21 (4) (1976), 791-804.

[13] Le Page, E., Ye, Y. The survival probability of a critical branching process in a Markovian random environment, *C. R. Mathematique*, 348 (5-6) (2011), 301-304.

[14] Nakashima, M. Lower deviation of branching process in random environment with geometrical offspring distributions. *Stoch. Pros. and their Appli.*, 123 (2013), 3560-3587.

[15] Smith, W. L., Wilkinson, W. E., On Branching Processes in Random Environments. *Ann. Math. Statist.*, 40 (3) (1969), 814-827.

[16] Tanny, D. A necessary and sufficient condition for a branching process in a random environment to grow like the product of its means. *Stoch. Pro. and Their Appli.*, 28 (1) (1988), 123-139.