EURASIAN MATHEMATICAL JOURNAL
ISSN 2077-9879
Volume 2, Number 1 (2011), 128 – 144

ASYMPTOTIC BEHAVIOUR OF A BOOTSTRAP BRANCHING PROCESS

I. Rahimov

Communicated by S. Sagitov

Key words: random probability measure, weak convergence, Skorokhod space, branching process, time-dependent immigration, bootstrap, least squares estimator.

AMS Mathematics Subject Classification: 60J80, 62F12, 60G99.

Abstract. In this paper we study weak convergence of sequences of random probability measures generated by bootstrap branching processes. Let $\{Z(k), k \geq 0\}$ be a branching stochastic process with non-stationary immigration given by an offspring distribution $\{p_j(\theta), j \geq 0\}$ depending on the unknown parameter $\theta \in \Theta \subseteq \mathbb{R}$. We estimate θ by an estimator $\hat{\theta}_n$ based on a sample $\{Z(i), i = 1, ..., n\}$. Given $\hat{\theta}_n$, we generate the bootstrap branching process (BBP) $\{Z^{\hat{\theta}_n}(k), k \geq 0\}$ for each $n = 1, 2, ...$ with the offspring distribution $\{p_j(\hat{\theta}_n), j \geq 0\}$. We derive conditions on the estimator $\hat{\theta}_n$ which are sufficient and necessary for the bootstrap process to have the same asymptotic properties as the original process. These results allow us to investigate the validity of the bootstrap without using an explicit form of the estimator. In applications of branching processes obtaining samples of large sizes is difficult. Therefore, the bootstrap process can be used to generate multiple samples of large size.

1 Introduction

We consider a discrete time branching stochastic process $Z(k), k \geq 0, Z(0) = 0$. It can be defined by two families of independent, nonnegative integer-valued random variables $\{X_{ki}, k, i \geq 1\}$ and $\{\xi_k, k \geq 1\}$ recursively as

$$Z(k) = \sum_{i=1}^{Z(k-1)} X_{ki} + \xi_k, \quad k \geq 1. \tag{1.1}$$

Assume that X_{ki} have a common distribution for all k and i, and that the families $\{X_{ki}\}$ and $\{\xi_k\}$ are independent. The variable X_{ki} will be interpreted as the number of offspring of the ith individual in the $(k-1)$th generation and ξ_k as the number of immigrating individuals to the kth generation. Then $Z(k)$ can be considered as the size of kth generation of the population.

In this interpretation $a = EX_{ki}$ is the mean number of offspring of a single individual. Process $Z(k)$ is called *subcritical, critical* or *supercritical* depending on $a < 1, a = 1$

or $a > 1$ respectively. The independence assumption of families $\{X_{ki}\}$ and $\{\xi_k\}$ means that the reproduction and immigration processes are independent. However, in contrast of classical models, we do not assume that $\xi_k, k \geq 1$ are identically distributed, i. e. the immigration rate may depend on the time of immigration.

The process with time-dependent immigration is given by the offspring distribution of $X_{ki}, k, i \geq 1$, and by the family of distributions of the number of immigrating individuals $\xi_k, k \geq 1$. We assume that the offspring distribution has the probability mass function

$$p_j(\theta) = P\{X_{ki} = j\}, \; j = 0, 1, ... \tag{1.2}$$

depending on the unknown parameter θ, where $\theta \in \Theta \subseteq \mathbb{R}$. We also assume that ξ_k for any $k \geq 1$ follows a known distribution with the probability mass function

$$q_j(k) = P\{\xi_k = j\}, \; j = 0, 1, ...$$

We estimate θ by an estimator $\hat{\theta}_n$ based on a sample $\{Z(i), i = 1, ..., n\}$ and generate the BBP $\{Z^{\hat{\theta}_n}(k), k \geq 0\}$ for each $n = 1, 2, ...$ as follow. Given $\hat{\theta}_n$, let $\{X_{ki}^{\hat{\theta}_n}, k, i \geq 1\}$ be the family of i.i.d. random variables with the probability mass function $\{p_j(\hat{\theta}_n), j = 0, 1, ...\}$. Now we obtain the process $\{Z^{\hat{\theta}_n}(k), k \geq 0\}$ recursively from

$$Z^{\hat{\theta}_n}(k) = \sum_{i=1}^{Z^{\hat{\theta}_n}(k-1)} X_{ki}^{\hat{\theta}_n} + \xi_k, \quad n, k \geq 1, \tag{1.3}$$

with $Z^{\hat{\theta}_n}(0) = 0$. As in (1.1), $\xi_k, k \geq 1$, are independent random variables with the probability mass functions $\{q_j(k), j = 0, 1, ...\}$.

Related to the process $\{Z^{\hat{\theta}_n}(k), k \geq 0\}$ the following question is of interest. How good must be the estimator $\hat{\theta}_n$ in order that the BBP $\{Z^{\hat{\theta}_n}(k), k \geq 0\}$ has the same asymptotic properties as the process $\{Z(k), k \geq 0\}$? For example, if we denote $\mathcal{Z}_n(t) = Z([nt])/E(Z(n))$ and $\{\mathcal{Z}_n(t), t \in \mathbb{R}_+\}$ converges weakly as $n \to \infty$ to some process $\{\mathcal{Z}(t), t \in \mathbb{R}_+\}$, in Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$, will the same be true for $\mathcal{Z}_n^{\hat{\theta}_n}(t) = Z^{\hat{\theta}_n}([nt])/E[Z^{\hat{\theta}_n}(n)|\hat{\theta}_n]$? A similar question for the process of fluctuations of $\{Z^{\hat{\theta}_n}(k), \; k \geq 0\}$ can also be considered.

To answer these questions without concretization of the process in the sense of criticality is impossible, because it is well known that the asymptotic properties of the process strictly depend on whether the process is subcritical, critical or supercritical. As a result, there is no general limit theory for branching processes without a criticality assumption. In this paper we address the above question in the critical case. In applications the question on criticality of the process is crucial. To answer this question, one may test hypothesis $H_0 : a = 1$ against one of $a \neq 1, a > 1$ or $a < 1$. Since the distribution of a test statistic is computed under the null hypothesis, the results obtained in the critical case allow to develop rejection regions for these hypotheses based on observed population sizes. On the other hand, the methods and concepts developed in this paper may also be used in subcritical and supercritical cases.

It is clear that the problem, which we are going to consider, is closely related to the problem of validity of the bootstrap procedure. In particular, if the process preserves its asymptotic properties after "bootstrapping", it can be used to generate multiple bootstrap samples. These new samples can further be used in statistical inference for the process. This is very important in branching process models, since in statistics of branching processes, usually, the generation number plays the role of the sample size and, therefore, it is difficult to obtain samples of large size. On the other hand, sometimes, in applications (for example, in epidemic processes) one needs to make a decision on criticality of the processes when it is still at the early stages.

First efforts for justification of the validity of the parametric bootstrap [15] have shown that in the critical case the bootstrap procedure based on a maximum likelihood estimator (MLE) of the offspring mean is asymptotically invalid for the process with stationary immigration. Later, it was demonstrated [2] that for a modified version of the MLE the parametric bootstrap is valid. It has recently been shown that in the process with non-stationary immigration the validity of the parametric bootstrap based on the conditional least squares estimator (CLSE) depends on the relative rate of the immigration mean and variance.

In present article we obtain sufficient and necessary conditions on the estimator of the offspring mean for the bootstrap process and for the process of fluctuations to preserve asymptotic properties of the original process. These conditions will be formulated in the form of the rate of convergence of the estimator to the true value of the parameter when the sample size increases and does not require the explicit form of the estimator. Therefore, our results can be used in investigation of the validity of bootstrap procedure when an explicit form of the estimator is unknown.

Statistical problems related to branching processes with various applications can be seen in [7] and [8]. Investigation of the problems related to the bootstrap methods and their applications has been an active area of research since they were introduced by Efron [5]. As a result a large number of papers and monographs have been published. We note monographs [4], [6] and [14] and the most recent review articles [3] and [9] as important sources of the literature on bootstrap methods.

Standing assumptions, necessary definitions and main theorems are given in Section 2. In Section 3 we provide functional limit theorems for an array branching process, which are necessary to prove our main results. Proofs of main theorems are given in Section 4. We note that some of the results of this paper without proofs were announced at the Workshop on Branching Processes and their Applications [13].

We conclude this section with a list of main notation.

• $\{p_j(\theta), j \geq 0\}$ is the offspring distribution depending on the unknown parameter $\theta \in \Theta$, which is taken to be the same for all generations.
• a, b are respectively the offspring mean and variance, depending on θ.
• $\alpha(n), \beta(n)$ are respectively the mean and variance of the number of immigrants in generation n (assumed to be known).
• R_ρ is the class of all sequences regularly varying at infinity with exponent ρ.
• α, β are exponents of the sequences $(\alpha(k))_{k=1}^{\infty} \in R_\alpha, (\beta(k))_{k=1}^{\infty} \in R_\beta$ with $\alpha, \beta \geq 0$.
• $(Z(k))_{n=1}^{\infty}$ is the sequence of random generation sizes for the branching process with variable immigration starting from $Z(0) = 0$ particles.
• $\hat{\theta}_n$ is the point estimate of the unknown parameter θ based on the sample observation $\{Z(i), i = 1, ..., n\}$.

• $\{Z^{\hat{\theta}_n}(k), k \geq 0\}$ is the bootstrap branching process (BBP) generated using the estimated parameter value $\hat{\theta}_n$.

2 Main theorems and examples

It follows from (1.2) that the quantities $a := E_\theta X_{ki} = f_a(\theta)$ and $b = Var_\theta X_{ki} = f_b(\theta)$ are some functions of θ, when they do exist. Let the following assumptions be satisfied.
A1. *The function f_a is a one-to-one mapping of Θ to $[0, \infty)$ and is continuous with continuous inverse (i.e. a homeomorphism between its domain and range).*
A2. *The function f_b is continuous on its domain.*

We note that A1 and A2 are satisfied, for example, for distributions of the power series family. Given a sample $\{Z(i), i = 1, ..., n\}$, we now estimate the offspring mean a by an estimator $\hat{a}_n$ and derive the estimate of parameter θ as $\hat{\theta}_n = f_a^{-1}(\hat{a}_n)$. Let $\{Z^{\hat{\theta}_n}(k), k \geq 0\}$ be the BBP defined by (1.3). This construction reduces the problem stated above to finding conditions for estimator $\hat{a}_n$, which are sufficient to preserve asymptotic properties of the process. Since the weak convergence of the conditioned process $\{\mathcal{Z}_n^{\hat{\theta}_n}(t), t \in \mathbb{R}_+\}$ given $\hat{\theta}_n$ is equivalent to convergence of the conditional probability measures generated by $\mathcal{Z}_n^{\hat{\theta}_n}$, now we provide necessary definitions of convergence of random probability measures defined on Skorokhod space.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(D, \mathcal{B}(D))$ be a measurable Skorokhod space, where $\mathcal{B}(D)$ is the Borel field on D. A function $\mu : \Omega \times \mathcal{B} \mapsto [0, 1]$ is called a random probability measure on D, if
(a) for each $B \in \mathcal{B}(D), \mu(\cdot, \ B)$ is a random variable on $(\Omega, \mathcal{A})$;
(b) for each $\omega \in \Omega, \mu(\omega, \ \cdot)$ is a probability measure on $(D, \mathcal{B}(D))$.

Definition 2.1. Let $\mu^{(n)}$ for each n be a random probability measure on $(D, \mathcal{B}(D))$.

(a) We say that $\mu^{(n)}$ converges weakly to μ on a set $A \in \mathcal{A}$, if for each $\omega \in A$

$$\int_D g(x)\mu^{(n)}(\omega, dx) \to \int_D g(x)\mu(\omega, dx) \tag{2.1}$$

as $n \to \infty$ for any function $g = g(x)$ bounded and continuous in Skorokhod metric. If $P\{A\} = 1$, we say that $\mu^{(n)}$ converges weakly to μ almost surely.

(b) We say that $\mu^{(n)}$ converges weakly to μ in probability (in distribution), if as $n \to \infty$ convergence (2.1) holds in probability (in distribution).

Here and throughout the paper " $\xrightarrow{D}$ ", " $\xrightarrow{d}$ " and " $\xrightarrow{P}$ " will denote convergence of random functions in Skorokhod topology and convergence of random variables in distribution and in probability, respectively. Also $X \stackrel{d}{=} Y$ denotes equality of distributions. In [17], the authors discussed the weak convergence of distributions of random probability measures. We note that in the case of conditional distributions Definition 2.1 coincides with their definition of weak convergence in probability. For different modes of convergence of conditional probability distributions see also [16].

Let $X(t), X_n(t), n \geq 1$, be conditioned processes with paths on Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ and $\mu, \mu^{(n)}, n \geq 1$, be corresponding random probability measures. Convergence of conditioned processes can now be defined as follows.

Definition 2.2. We say that a sequence of conditioned processes $\{X_n, n \geq 1\}$ converges weakly as $n \to \infty$ in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ to X on a set A, in probability or in distribution, if the sequence of corresponding random probability measures $\{\mu^{(n)}, n \geq 1\}$ converges weakly to μ on the set A, in probability or in distribution, respectively.

If a sequence $(f(k))_{k=1}^{\infty}$ is regularly varying with exponent ρ, we will write $(f(k))_{k=1}^{\infty} \in R_\rho$. We assume that $a = EX_{ij}$ and $b = VarX_{ij}$ are finite. We also assume that $\alpha(k) := E\xi_k < \infty$, $\beta(k) := Var\xi_k < \infty$ for each $k \geq 1$ and $(\alpha(k))_{k=1}^{\infty} \in R_\alpha, (\beta(k))_{k=1}^{\infty} \in R_\beta$ with $\alpha, \beta \geq 0$. Then $A(a,n) = EZ(n)$ and $B^2(a,n) = VarZ(n)$ are finite for each $n \geq 1$, and by a standard technique we find that

$$A(a,n) = \sum_{i=1}^{n} \alpha(i) a^{n-i}, \quad B^2(a,n) = \Delta^2(a,n) + \sigma^2(a,n),$$

where

$$\Delta^2(a,n) = \sum_{i=1}^{n} \alpha(i) Var(X(n-i)), \quad \sigma^2(a,n) = \sum_{i=1}^{n} \beta(i) a^{2(n-i)},$$

$$Var(X(i)) = \frac{b}{1-a} a^{i-1}(1-a^i), \ a \neq 1.$$

Here $\{X(i), i \geq 0\}$ is the corresponding branching process without immigration with offspring distribution (1.2) and $X(0) = 1$.

In particular, we denote $A(n) = A(1,n)$, $B^2(n) = B^2(1,n)$, $\Delta^2(n) = \Delta^2(1,n)$, $\sigma^2(n) = \sigma^2(1,n)$ and put

$$\mathcal{Z}_n^{\hat{\theta}_n}(t) = \frac{Z^{\hat{\theta}_n}([nt])}{A(\hat{a}_n, n)}, \ \mathcal{Y}_n^{\hat{\theta}_n}(t) = \frac{Z^{\hat{\theta}_n}([nt]) - A(\hat{a}_n, [nt])}{B(\hat{a}_n, n)}.$$

Now we provide the first result for the bootstrap process. We denote $A = \{\omega \in \Omega : n(\hat{a}_n - 1) \to 0, n \to \infty\}$, $\mu_\alpha(t) = t^{1+\alpha}, t \in \mathbb{R}_+$.

Theorem 2.1. *Let A1 and A2 be satisfied and* $\alpha(n) \to \infty$, $\beta(n) = o(n\alpha^2(n))$ *as* $n \to \infty$.

(a) Conditioned process $\{\mathcal{Z}_n^{\hat{\theta}_n} | \ \hat{\theta}_n\}$ *as* $n \to \infty$ *converges weakly in Skorokhod space* $D(\mathbb{R}_+, \mathbb{R}_+)$ *to* μ_α *on the set* A.

(b) If $n(\hat{a}_n - 1) \xrightarrow{P} 0$, *then* $\{\mathcal{Z}_n^{\hat{\theta}_n} | \ \hat{\theta}_n\}$ *as* $n \to \infty$ *converges weakly in Skorokhod space* $D(\mathbb{R}_+, \mathbb{R}_+)$ *to* μ_α *in probability.*

The next result is related to the fluctuations of the bootstrap process. Let $\bar{\xi}_k = \xi_k - \alpha(k)$ and

$$\delta_n(\varepsilon) = \frac{1}{B^2(n)} \sum_{k=1}^{n} E[(\bar{\xi}_k)^2 \chi(|\bar{\xi}_k| > \varepsilon B(n))].$$

We also denote $\psi(t) = \gamma_1 t^{2+\alpha} + \gamma_2 t^{1+\beta}$, where

$$\gamma_1 = \lim_{n \to \infty} \frac{\Delta^2(n)}{B^2(n)}, \ \gamma_2 = \lim_{n \to \infty} \frac{\sigma^2(n)}{B^2(n)}, \ \gamma_1 + \gamma_2 = 1.$$

We need two more conditions to be satisfied.
A3. The moment $E_\theta[(X_{ki})^{2+l}]$ is a continuous function of θ for some $l > 0$.
A4. $\delta_n(\varepsilon) \to 0$ as $n \to \infty$ for each $\varepsilon > 0$.

Theorem 2.2. *Let A1-A4 be satisfied and $\alpha(n) \to \infty$ as $n \to \infty$.*

(a) The conditioned process $\{\mathcal{Y}_n^{\hat{\theta}_n} | \ \hat{\theta}_n\}$ converges weakly in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ to $\mathcal{Y}$ on the set A, where $\mathcal{Y}(t) = W(\psi(t))$ and $W(t)$ is the standard Wiener process.

(b) If $n(\hat{a}_n - 1) \xrightarrow{P} 0$, then $\{\mathcal{Y}_n^{\hat{\theta}_n} | \ \hat{\theta}_n\}$ as $n \to \infty$ converges weakly in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ to $\mathcal{Y}$ in probability.

Remark 2.1. (a) It follows from Theorem 2.1 (a) and Theorem 2.1 in [11] that the conditioned bootstrap process $\{Z^{\hat{\theta}_n}(k) | \hat{\theta}_n, k \geq 0\}$ generated by estimator $\hat{\theta}_n$ such that $n(\hat{a}_n - 1) \to 0$ a.s., under some conditions, a.s. has the same asymptotic behavior as the original process.
(b) If we compare Theorem 2.2 (a) with Theorems 1, 2 and 3 in [10], we see that the same is true for fluctuations of the bootstrap process. More precisely, in [10] for a single critical process exactly the same limit process was obtained, considering three cases of the relationship between the immigration mean and the variance separately.

Example 2.1. Now we provide an example of the estimator that satisfies conditions of parts (b) of the above theorems. Let $\hat{a}_n$ be the weighted conditional least squares estimator (WCLSE), derived in [11] from a "standardized" stochastic regression equation. If the sample $\{Z(i), i = 1, ..., n\}$ is available and the immigration mean is known, it is defined as

$$\hat{a}_n = \frac{\sum_{k=1}^{n}(Z(k) - \alpha(k))}{\sum_{k=1}^{n} Z(k-1)}. \tag{2.2}$$

To provide the asymptotic distribution for $\hat{a}_n$, we assume that there exists $c \in [0, \infty]$ such that

$$\lim_{n\to\infty} \frac{\beta(n)}{n\alpha(n)} = c. \tag{2.3}$$

As was proved in [11], if $a = 1$, $b \in (0, \infty)$, $\alpha(n) \to \infty$, $\beta(n) = o(n\alpha^2(n))$, condition (2.3) is satisfied and $\delta_n(\varepsilon) \to 0$ for each $\varepsilon > 0$, then as $n \to \infty$

$$\frac{nA(n)}{B(n)}(\hat{a}_n - a) \xrightarrow{d} (2 + \alpha)\mathcal{N}(0, 1) \tag{2.4}$$

as $n \to \infty$. Furthermore, under the above conditions, $A(n)/B(n) \to \infty$ as $n \to \infty$ and when $c = 0$ the condition $\delta_n(\varepsilon) \to 0$ is satisfied automatically. More detailed discussion and examples can be found in [11].

From (2.4) we immediately obtain that $n(\hat{a}_n - 1) \xrightarrow{P} 0$ as $n \to \infty$. Thus the following result holds.

Corollary 2.1. *Let $\hat{a}_n$ be the WCLSE defined in (2.2), $a = 1$, $\alpha(n) \to \infty$ and $\beta(n) = o(n\alpha^2(n))$ as $n \to \infty$.*

(a) If A1 and A2 are satisfied, then $\{\mathcal{Z}_n^{\hat{\theta}_n} |\ \hat{\theta}_n\}$ converges weakly as $n \to \infty$ in Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$ to μ_α in probability.

(b) If A1-A4 are satisfied, then $\{\mathcal{Y}_n^{\hat{\theta}_n} |\ \hat{\theta}_n\}$ converges weakly as $n \to \infty$ in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ to $\mathcal{Y}$ in probability, where $\mathcal{Y}(t) = W(\psi(t))$.

The next theorem is related to the case

$$n(\hat{a}_n - 1) \xrightarrow{d} W_0 \tag{2.5}$$

as $n \to \infty$, where W_0 is a random variable. We denote

$$\mu_\alpha(d,t) = \int_0^t u^\alpha e^{d(t-u)} du,\ \pi_\alpha(d,t) = \frac{\mu_\alpha(d,t)}{\mu_\alpha(d,1)}. \tag{2.6}$$

We note that $\mu_\alpha(0,t) = \mu_\alpha(t)$, the limiting "process" in Theorem 2.1.

Theorem 2.3. *If A1, A2 and (2.5) are satisfied and $\alpha(n) \to \infty$, $\beta(n) = o(n\alpha^2(n))$ as $n \to \infty$, then $\{\mathcal{Z}_n^{\hat{\theta}_n} |\ \hat{\theta}_n\}$ converges weakly as $n \to \infty$ in Skorokhod space $D(\mathbb{R}_+, \mathbb{R}_+)$ to $\pi_\alpha(W_0, \cdot)$ in distribution.*

Let $(a_n)_{n=1}^\infty$ be a sequence of positive numbers, such that $n(a_n - 1) \to d \in \mathbb{R}$ as $n \to \infty$. We assume that there exist limits

$$\lim_{n\to\infty} \frac{\Delta^2(a_n,n)}{B^2(a_n,n)} = \gamma_1(d),\ \lim_{n\to\infty} \frac{\sigma^2(a_n,n)}{B^2(a_n,n)} = \gamma_2(d). \tag{2.7}$$

Naturally $\gamma_1(d) + \gamma_2(d) = 1$ for each d.

To provide the next theorem, we need some additional notation. We denote

$$\nu_\alpha(d,t) = \int_0^t u^\alpha e^{d(t-u)}(1 - e^{d(t-u)}) du,\ \nabla_\beta(d,t) = \int_0^t u^\beta e^{2d(t-u)} du,$$

$$\psi(d,t) = \frac{\gamma_1(d) d}{\nu_\alpha(d,1)} \int_0^t \mu_\alpha(d,u) e^{2d(t-u)} du + \frac{\gamma_2(d)}{\nabla_\beta(d,1)} \int_0^t u^\beta e^{2d(t-u)} du. \tag{2.8}$$

It is clear that the limits in (2.7) do exist, if ratio $\sigma^2(a_n,n)/\Delta^2(a_n,n)$ has a (finite or infinite) limit as $n \to \infty$. Using Lemma 3.1, given below, we can show that

$$\lim_{n\to\infty} \frac{\sigma^2(a_n,n)}{\Delta^2(a_n,n)} = \frac{bc}{d} \nu(d,1) \nabla_\beta(d,1)$$

where c is defined in (2.3). In particular, it is also useful to note that $\mu_\alpha(d,t) = t^{\alpha+1}/(\alpha+1)$ and $\nabla_\beta(d,t) = t^{1+\beta}/(1+\beta)$ when $d = 0$, and $\lim_{d\to 0} \nu_\alpha(d,t)/a = t^{\alpha+2}/(\alpha+1)(\alpha+2)$.

Theorem 2.4. *If A1-A4 and (2.5) are satisfied and $\alpha(n) \to \infty$ as $n \to \infty$, then $\{\mathcal{Y}_n^{\hat{\theta}_n} | \hat{\theta}_n\}$ converges weakly as $n \to \infty$ in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ to $\mathcal{Y}(W_0, \cdot)$ in distribution, where $\mathcal{Y}(W_0, t) = W(\psi(W_0, t))$.*

Remark 2.2. Theorems 2.3 and 2.4 show that, when the estimator $\hat{\theta}_n$ is such that (2.5) holds with $P(W_0 = 0) < 1$, then the asymptotic behavior of the bootstrap process is different from the behavior of the original process. In other words, the condition $n(\hat{a}_n - 1) \to 0$ as $n \to \infty$ a.s. or in probability is necessary for the conditioned bootstrap process to have the same asymptotic behavior as the initial process in the sense of convergence a.s. or in probability, respectively.

We conclude this section with a result which is useful in the estimation theory. Let $\mathcal{F}^{\hat{\theta}_n}(k)$ for each k and n be the sigma-algebra generated by $\{Z^{\hat{\theta}_n}(i), i = 1, 2, ..., k\}$ and $M^{\hat{\theta}_n}(k) = Z^{\hat{\theta}_n}(k) - E[Z^{\hat{\theta}_n}(k) | \mathcal{F}^{\hat{\theta}_n}(k-1)]$. Then $\{M^{\hat{\theta}_n}(k), \mathcal{F}^{\hat{\theta}_n}(k)\}_{k=1}^{\infty}$ given $\hat{\theta}_n$ is a sequence of martingale differences. We define process

$$\mathcal{W}_n^{\hat{\theta}_n}(t) = \sum_{i=1}^{[nt]} \frac{M^{\hat{\theta}_n}(i)}{B(\hat{a}_n, n)}.$$

Theorem 2.5. *Let A1-A4 be satisfied and $\alpha(n) \to \infty$ as $n \to \infty$. then*

(a) conditioned process $\{\mathcal{W}_n^{\hat{\theta}_n} | \hat{\theta}_n\}$ converges weakly as $n \to \infty$ in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ to $\mathcal{Y}$ on the set A, where $\mathcal{Y}(t) = W(\psi(t))$, $W(t)$ is the standard Wiener process and $\psi(t)$ is defined just before Theorem 2.2;

(b) if $n(\hat{a}_n - 1) \xrightarrow{P} 0$, then $\{\mathcal{W}_n^{\hat{\theta}_n} | \hat{\theta}_n\}$ converges weakly as $n \to \infty$ in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$ to $\mathcal{Y}$ in probability.

Example 2.2. As it was mentioned before, the approximation theorems for the bootstrap process allow us to investigate the validity of the bootstrap without using an explicit form of the estimator. Here we demonstrate it for WCLSE. We use the indirect approach suggested by Ch. Jacob in a private discussion. We represent the bootstrap process as

$$Z^{\hat{\theta}_n}(k) = E[Z^{\hat{\theta}_n}(k) | \mathcal{F}^{\hat{\theta}_n}(k-1)] + M^{\hat{\theta}_n}(k).$$

Using $(Z^{\hat{\theta}_n}(k-1))^{1/2}$ as a "weight", we define

$$S_n(a) = \sum_{k=1}^{n} (T_k^{(n)} - g_{nk}(a))^2,$$

where $T_k^{(n)} = Z^{\hat{\theta}_n}(k)(Z^{\hat{\theta}_n}(k-1))^{-1/2}$ and $g_{nk}(a) = E[Z^{\hat{\theta}_n}(k) | \mathcal{F}^{\hat{\theta}_n}(k-1)](Z^{\hat{\theta}_n}(k-1))^{-1/2}$. Then it is clear that the bootstrap WCLSE is

$$\hat{a}_n^{\hat{\theta}_n} = arg \min_{a \in \mathbb{R}_+} S_n(a).$$

We use the Taylor expansion for $S_n'(a)$ as follows:

$$S_n'(\hat{a}_n^{\hat{\theta}_n}) = S_n'(\hat{a}_n) + S_n''(a_n)(\hat{a}_n^{\hat{\theta}_n} - \hat{a}_n),$$

where $\hat{a}_n$ is the initial WCLSE, $a_n = \hat{a}_n + \varepsilon(\hat{a}_n^{\hat{\theta}_n} - \hat{a}_n)$ and $\varepsilon \in (0,1)$. Since $S_n'(\hat{a}_n^{\hat{\theta}_n}) = 0$, we obtain that

$$\hat{a}_n^{\hat{\theta}_n} - \hat{a}_n = -\frac{S_n'(\hat{a}_n)}{S_n''(a_n)}. \tag{2.9}$$

If we take into account that

$$g_{nk}'(a) = (Z^{\hat{\theta}_n}(k-1))^{1/2},\ g_{nk}''(a) = 0,\ S_n'(a) = -2\sum_{k=1}^{n}(T_k^{(n)} - g_{nk}(a))g_{nk}'(a),$$

$$S_n''(a) = -2\sum_{k=1}^{n}(g_{nk}'(a))^2 = 2\sum_{k=1}^{n} Z^{\hat{\theta}_n}(k-1),$$

we obtain from (2.9) that

$$\hat{a}_n^{\hat{\theta}_n} - \hat{a}_n = \frac{\sum_{k=1}^{n} M^{\hat{\theta}_n}(k)}{\sum_{k=1}^{n} Z^{\hat{\theta}_n}(k-1)}.$$

Let $n(\hat{a}_n - 1) \xrightarrow{P} 0$ as $n \to \infty$. It follows from parts (b) of Theorems 2.1 and 2.5 that

$$\left\{\frac{nA(\hat{a}_n, n)}{B(\hat{a}_n, n)}(\hat{a}_n^{\hat{\theta}_n} - \hat{a}_n)|\hat{\theta}_n\right\} \xrightarrow{d} (2+\alpha)\mathcal{N}(0,1),$$

as in the proof of Theorem 3.1 in [11]. Using Lemma 3.1, which is given in the next section, we can show that $A(\hat{a}_n, n) \overset{P}{\sim} A(n)$ and $B(\hat{a}_n, n) \overset{P}{\sim} B(n)$ as $n \to \infty$. Thus, we have the following result.

Theorem 2.6. *Let A1-A4 be satisfied,* $n(\hat{a}_n - 1) \xrightarrow{P} 0$, $\alpha(n) \to \infty$ *and* $\beta(n) = o(n\alpha^2(n))$ *as* $n \to \infty$. *Then*

$$\left\{\frac{nA(n)}{B(n)}(\hat{a}_n^{\hat{\theta}_n} - \hat{a}_n)|\hat{\theta}_n\right\} \xrightarrow{d} (2+\alpha)\mathcal{N}(0,1).$$

In particular, Theorem 2.6 shows the validity of the bootstrap for the WCLSE defined in Example 2.1.

3 Array of processes

In this section we provide functional limit theorems for an array of branching processes, which will be used in the proof of our main theorems. Let $\{X_{ki}^{(n)}, k, i \geq 1\}$ and $\{\xi_k^{(n)}, k \geq 0\}$ be two families of independent, nonnegative and integer-valued random variables for each $n \in \mathbb{N}$. We consider a sequence of branching processes $(Z^{(n)}(k), k \geq 0)_{n\geq 1}$ defined recursively as

$$Z^{(n)}(k) = \sum_{i=1}^{Z^{(n)}(k-1)} X_{ki}^{(n)} + \xi_k^{(n)}, \quad k, n \geq 1, \tag{3.1}$$

with $Z^{(n)}(0) = 0, n \geq 1$. As before, we assume that $X_{ki}^{(n)}$ have a common distribution for all k and i, and families $\{X_{ki}^{(n)}\}$ and $\{\xi_k^{(n)}\}$ are independent. The variables $X_{ki}^{(n)}$ will be interpreted as the number of offspring of the ith individual in the $(k-1)$th generation and $\xi_k^{(n)}$ as the number of immigrating individuals in the kth generation. Then $Z^{(n)}(k)$ can be considered as the size of population of kth generation in nth process.

Let $a_n = EX_{ki}^{(n)}$ be the mean number of offspring of a single individual in the nth process. The process with non-stationary immigration is a natural generalization of the classical model. It turned out that the long run behavior of the process is largely influenced by the non-homogeneity of the immigration process in time. As a result certain new problems, regarding the asymptotic behavior of the process when the immigration rate increases, decreases or remains bounded, emerged in the literature. Therefore, in solving these problems one needs certain regularity assumptions for the parameters of the immigration process. The family of branching processes (3.1) is said to be *nearly critical* if $a_n \to 1$ as $n \to \infty$.

We assume that $a_n = EX_{ij}^{(n)}$ and $b_n = VarX_{ij}^{(n)}$ are finite for each $n \geq 1$ and $\alpha(n,i) = E\xi_i^{(n)} < \infty$, $\beta(n,i) = Var\xi_i^{(n)} < \infty$ for all $n \geq 1$ and $i \geq 0$. Furthermore, we assume that the following condition is satisfied.

C1. There are sequences $(\alpha(i))_{i=1}^{\infty} \in R_\alpha$ and $(\beta(i))_{i=1}^{\infty} \in R_\beta$ with $\alpha, \beta \geq 0$, such that for each $s \in \mathbb{R}_+$,

$$\max_{1 \leq k \leq ns} |\alpha(n,k) - \alpha(k)| = o(\alpha(n)), \ \max_{1 \leq k \leq ns} |\beta(n,k) - \beta(k)| = o(\beta(n))$$

as $n \to \infty$.

In the above assumptions $A_n(a_n, i) = EZ^{(n)}(i)$ and $B_n^2(a_n, i) = VarZ^{(n)}(i)$ are finite for each $n \geq 1$, $0 \leq i \leq n$, and one can find that $A_n(a_n, k) = \sum_{i=0}^{k} \alpha(n,i) a_n^{k-i}$ and $B_n^2(a_n, k) = \Delta_n^2(a_n, k) + \sigma_n^2(a_n, k)$, where

$$\Delta_n^2(a_n, k) = \sum_{i=1}^{k} \alpha(n,i) Var(X^{(n)}(k-i)), \quad \sigma_n^2(a_n, k) = \sum_{i=1}^{k} \beta(n,i) a_n^{2(k-i)},$$

$$Var(X^{(n)}(i)) = \frac{b_n}{1-a_n} a_n^{i-1}(1-a_n^i), \ a_n \neq 1.$$

Here $X^{(n)}(i)$ is the corresponding branching process without immigration and, as usual, it is defined by the relation

$$X^{(n)}(k) = \sum_{i=1}^{X^{(n)}(k-1)} X_{ki}^{(n)}, \quad X^{(n)}(0) = 1, \quad k, n \geq 1.$$

In particular, when $k = n$ we use also notation

$$A(a_n, n) = A_n(a_n, n), \ B^2(a_n, n) = B_n^2(a_n, n),$$

$$\Delta^2(a_n, n) = \Delta_n^2(a_n, n), \ \sigma^2(a_n, k) = \sigma_n^2(a_n, n),$$

which is consistent with (2.1).

We consider the following processes.

$$Z_n(t) = \frac{Z^{(n)}([nt])}{A(a_n, n)}, \; Y_n(t) = \frac{Z^{(n)}([nt]) - EZ^{(n)}([nt])}{B(a_n, n)}.$$

First we provide a convergence theorem for $Z_n(t)$. We obtain approximation of the sequence $\{Z_n(t), n \geq 1\}, t \in \mathbb{R}_+$, satisfying the following conditions:
C2. for some $d \in \mathbb{R}$ $a_n = 1 + n^{-1}d + o(n^{-1})$ as $n \to \infty$.
C3. $b_n \to b \in \mathbb{R}_+$ as $n \to \infty$.
C4. $\alpha(n) \to \infty$ and $\beta(n) = o(n\alpha^2(n))$ as $n \to \infty$.

Theorem 3.1. *If conditions C1-C4 are satisfied, then $Z_n \xrightarrow{D} \pi_\alpha$ as $n \to \infty$ weakly in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$, where $\pi_\alpha(t)$ and $\mu_\alpha(d, t)$, $t \in \mathbb{R}_+$ are defined in (2.6).*

Remark 3.1. The condition C2 is the same as in the study of an array of time-homogeneous processes. The second condition in C4 appeared in the proof of the functional limit theorems for a single branching process with a non-homogeneous immigration as well. What concerns C1, the first part, related to the immigration mean, is satisfied when $\alpha(n) \to \infty$, if just $\lim_{n\to\infty} \max_{1\leq k\leq ns} |\alpha(n, k) - \alpha(k)| < \infty$. In general, C1 is satisfied, for example, if there are $\varepsilon_i(n) \to 0$ as $n \to \infty$, $i = 1, 2$, such that $\alpha(n, k) = \alpha(k)(1 + \varepsilon_1(n))$ and $\beta(n, k) = \beta(k)(1 + \varepsilon_2(n))$.

The proof of Theorem 3.1 can be found in [12].

Next theorem is related to fluctuations of the process. We denote “centered" offspring and immigration variables as $\bar{X}_{ki}^{(n)} = X_{ki}^{(n)} - a_n$, $\bar{\xi}_k^{(n)} = \xi_k^{(n)} - \alpha(n, k)$ and put

$$\delta_n^{(1)}(\varepsilon) = \gamma_1(d) E[(\bar{X}_{ki}^{(n)})^2 \chi(|\bar{X}_{ki}^{(n)}| > \varepsilon B(a_n, n))],$$

$$\delta_n^{(2)}(\varepsilon) = \frac{1}{B^2(a_n, n)} \sum_{k=1}^{n} E[(\bar{\xi}_k^{(n)})^2 \chi(|\bar{\xi}_k^{(n)}| > \varepsilon B(a_n, n))],$$

where $\chi(A)$ stands for the indicator of event A and $\gamma_1(a)$ is defined in (2.7). We need the following condition to be satisfied:
C5. $\delta_n^{(i)}(\varepsilon) \to 0$ as $n \to \infty$ for each $\varepsilon > 0$ and $i = 1, 2$.

Theorem 3.2. *If conditions C1-C3 and C5 are satisfied, then $Y_n \xrightarrow{D} Y$ as $n \to \infty$ weakly in Skorokhod space $D(\mathbb{R}_+, \mathbb{R})$, where $Y(t) = W(\psi(d, t))$, $t \in \mathbb{R}_+$, $W(t)$ is a standard Brownian motion and $\psi(d, t)$ is defined in (2.8).*

Remark 3.2. (a) Note that the Lindeberg-type condition for the family $\{X_{ki}^{(n)}, k, i \geq 1\}$ is trivially satisfied, if $\gamma_1(d) = 0$. If $\gamma_1(d) \neq 0$ and $E(X_{ki}^{(n)})^{2+l} < \infty$ for all $n \in \mathbb{N}$ and some $l \in \mathbb{R}_+$, then

$$\frac{\delta_n^{(1)}(\varepsilon)}{\gamma_1(d)} \leq \frac{1}{\varepsilon^l B^l(a_n, n)} E|X_{ki}^{(n)} - a_n|^{2+l}.$$

Since $B^2(a_n, n) \geq \Delta^2(a_n, n) \sim Kn^2\alpha(n)$ as $n \to \infty$ due to Lemma 3.1 below, where K is a positive constant, the Lindeberg-type condition is satisfied, for example, if $E|X_{ki}^{(n)} - a_n|^3 = o(n\sqrt{\alpha(n)})$ and $\alpha(n) \to \infty$ as $n \to \infty$.

(b) What concerns the Lindeberg-type condition for the immigration variables, it is automatically satisfied when $\gamma_1(d) \neq 0$, since in this case $\sigma^2(a_n, n) = o(B^2(a_n, n))$ as $n \to \infty$. If $\gamma_1(d) = 0$, then it is equivalent to the Lindeberg condition for the array $\{\xi_k^{(n)}, k, n \geq 1\}$.

(c) When processes $\{Z^{(n)}(k), k \geq 0\}, n \geq 1$, are critical with the same offspring and immigration distributions, conditions C1-C3 are satisfied with $a = 0$. Therefore, from Theorem 3.3 we obtain assertions of Theorems 1, 2 and 3 in [10] in the cases $\gamma_1(0) = 1,\ \gamma_2(0) = 1$ and $0 < \gamma_i(0) < 1,\ i = 1, 2$, respectively.

Theorem 3.2 can be proved using the same approach which was used in the proof of Theorems 2-4 in [12]. It needs just a more careful analysis in applying the martingale convergence theorem. Therefore, we do not give a proof of this theorem.

We now provide a theorem for the process of martingale differences. We define $\mathcal{F}^{(n)}(k) = \sigma\{Z^{(n)}(i), i = 1, 2, ..., k\}$ and denote $M^{(n)}(k) = Z^{(n)}(k) - E[Z^{(n)}(k)|\mathcal{F}^{(n)}(k-1)]$. We consider the following process:

$$W_n(t) = \frac{1}{B(d_n, n)} \sum_{i=1}^{[nt]} M^{(n)}(i).$$

Theorem 3.3. *If conditions C1-C3 and C5 are satisfied, then* $W_n \xrightarrow{D} Y^{(1)}$ *as* $n \to \infty$ *weakly in Skorokhod space* $D(\mathbb{R}_+, \mathbb{R})$, *where* $Y^{(1)}(t) = W(\varphi(d, t)),\ t \in \mathbb{R}_+,\ W(t)$ *is a standard Brownian motion and*

$$\varphi(d, t) = \frac{\gamma_1(d)d}{\nu_\alpha(d, 1)} \int_0^t \mu_\alpha(d, u)du + \frac{\gamma_2(d)t^{1+\beta}}{(1+\beta)\nabla_\beta(d, 1)}.$$

The proof of this theorem is also based on the direct use of the martingale-limit theorem and is similar to the proof of Theorems 2-4 in [12].

We conclude this section with a lemma borrowed from [12] which is required for proofs of main theorems.

Lemma 3.1. *If conditions C1 and C2 are satisfied, then uniformly in* $s \in [0, T]$ *for each fixed* $T > 0$

$$a)\quad \lim_{n\to\infty} \frac{A_n(a_n, [ns])}{n\alpha(n)} = \mu_\alpha(d, s),\ \lim_{n\to\infty} \frac{\sigma_n^2(a_n, [ns])}{n\beta(n)} = \nabla_\beta(d, s),$$

$$b)\quad \lim_{n\to\infty} \frac{\Delta_n^2(a_n, [ns])}{n^2\alpha(n)b_n} = \begin{cases} (1/d)\nu_\alpha(d, s), & \textit{if } d \neq 0 \\ s^{\alpha+2}/(\alpha+1)(\alpha+2), & \textit{if } d = 0 \end{cases}.$$

Lemma 3.1 is also proved in [12].

4 Proofs of main theorems

Proof of Theorem 2.1. Part (a). Since the bootstrap process $\{\mathcal{Z}_n^{\hat\theta_n}(t), t \in \mathbb{R}_+\}$ given $\hat\theta_n$, constitutes an array of branching processes, we show that conditions of Theorem 3.1 are satisfied. It is easy to see that C1 and C4 are trivially satisfied. Condition C2 is also satisfied on the set A with $a = 0$. It follows from A1:

$$B := \{\omega \in \Omega : \hat a_n \to 1\} = \{\omega \in \Omega : f_a^{-1}(\hat a_n) \to f_a^{-1}(1)\} = \{\omega \in \Omega : \hat\theta_n \to \theta_0\},$$

where θ_0 is the true value of θ. Since $b^{\hat\theta_n} := Var(X_{ki}^{\hat\theta_n}|\hat\theta_n) = f_b(\hat\theta_n)$, we immediately obtain from A2 that $b^{\hat\theta_n} \to b$ as $n \to \infty$ for each $\omega \in B$. Taking into account that $A \subset B$, we see that condition C2 is also satisfied on the set A. Hence, the assertion (a) of the theorem follows by Theorem 3.1.

Part (b). Let $P_n^{\hat\theta_n}$ and P_α be probability measures generated by $\{\mathcal{Z}_n^{\hat\theta_n}|\hat\theta_n\}$ and μ_α, respectively. Assume that as $n \to \infty$

$$n(\hat a_n - 1) \xrightarrow{P} 0. \tag{4.1}$$

We prove that any subsequence $\mathbb{N}' \subset \mathbb{N} = \{1, 2, ...\}$ contains another subsequence $\mathbb{N}'' \subset \mathbb{N}'$ such that

$$\int_D g(x) P_n^{\hat\theta_n}(\omega, dx) \xrightarrow{a.s.} \int_D g(x) P_\alpha(dx) \tag{4.2}$$

along $\mathbb{N}''$, for any function $g : D(\mathbb{R}_+\mathbb{R}_+) \mapsto \mathbb{R}$ bounded and continuous in Skorokhod metric. It follows from (4.1) that $n(\hat a_n - 1) \xrightarrow{P} 0$ along any subsequence $\mathbb{N}' \subset \mathbb{N}$. Therefore, there is a subsequence $\mathbb{N}'' \subset \mathbb{N}'$ such that $n(\hat a_n - 1) \to 0$ along $\mathbb{N}''$ for each $\omega \in A$ with $P(A) = 1$. Thus, due to part (a) of Theorem 2.1,

$$\int_D g(x) P_n^{\hat\theta_n}(\omega, dx) \to \int_D g(x) P_\alpha(dx)$$

along $\mathbb{N}''$ for each $\omega \in A$, which implies (4.2). □

Proof of Theorem 2.2. Part (a). We show that conditions of Theorem 3.2 are satisfied. Conditions C1-C3 are satisfied, as in the proof of previous theorem. We just need to check condition C5. For this we denote

$$A^{\hat\theta_n}(k) := E[Z^{\hat\theta_n}(k)|\hat\theta_n],\ B^{2\hat\theta_n}(k) := E[(Z^{\hat\theta_n}(k) - A^{\hat\theta_n}(k))^2|\hat\theta_n]$$

Then we easily obtain that $A^{\hat\theta_n}(n) = A(\hat a_n, n),\ B^{2\hat\theta_n}(n) = B^2(\hat a_n, n)$.

We apply Lemma 3.1 to get on the set $A = \{\omega \in \Omega : n(\hat a_n - 1) \to 0\}$ the following convergence

$$\frac{A(\hat a_n, [ns])}{n\alpha(n)} \to \frac{s^{1+\alpha}}{1+\alpha},\ \frac{\sigma^2(\hat a_n, [ns])}{n\beta(n)} \to \frac{s^{1+\beta}}{1+\beta}$$

$$\frac{\Delta^2(\hat a_n, [ns])}{n^2\alpha(n) f_b(\hat\theta_n)} \to \frac{s^{2+\alpha}}{(1+\alpha)(2+\alpha)}$$

as $n \to \infty$ for each $s \in \mathbb{R}_+$. On the other hand, since $\Delta^2(\hat{a}_n, n)/\Delta^2(n) \to 1$ and $\sigma^2(\hat{a}_n, n)/\sigma^2(n) \to 1$ as $n \to \infty$ on the set A, we have $B^2(\hat{a}_n, n)/B^2(n) \to 1$ on the set A. We conclude from these arguments that

$$\delta^{(2)}(\varepsilon, \hat{\theta}_n) := \frac{1}{B^2(\hat{a}_n, n)} \sum_{k=1}^{n} E[(\bar{\xi}_k)^2 \chi(|\bar{\xi}_k| > \varepsilon B(\hat{a}_n, n))|\hat{\theta}_n]$$

tends to zero as $n \to \infty$ on the set A, due to condition A4.

We now consider

$$\delta_n^{(1)}(\varepsilon, \hat{\theta}_n) = \gamma_1(d) E[(\bar{X}_{ki}^{\hat{\theta}_n})^2 \chi(|\bar{X}_{ki}^{\hat{\theta}_n}| > \varepsilon B(\hat{a}_n, n))|\hat{\theta}_n].$$

If $\gamma_1(0) = 0$, then $\delta_n^{(1)}(\varepsilon, \hat{\theta}_n) \to 0$ as $n \to \infty$ on A. If $\gamma_1(0) \neq 0$, then

$$\frac{\delta_n^{(1)}(\varepsilon, \hat{\theta}_n)}{\gamma_1(0)} \leq \frac{1}{\varepsilon^l B^l(\hat{a}_n, n)} E[|X_{ki}^{\hat{\theta}_n} - \hat{a}_n|^{2+l}|\hat{\theta}_n]. \tag{4.3}$$

It follows from condition A3 that on the set $A \subset B = \{\omega \in \Omega : \hat{\theta}_n \to \theta_0\}$ we have $E[(X_{ki}^{\hat{\theta}_n})^{2+l}|\hat{\theta}_n] \to E[(X_{ki})^{2+l}]$ as $n \to \infty$. If we take this into account, we obtain from (4.3) that $\delta_n^{(1)}(\varepsilon, \hat{\theta}_n) \to 0$ as $n \to \infty$ on the set A.

Thus, condition C5 of Theorem 3.2 is satisfied and we have the assertion (a) of Theorem 2.2.

The proof of part (b) repeats the arguments of the proof of convergence in probability in Theorem 2.1. We just need to consider the sequence of probability measures generated by $\{\mathcal{Y}_n^{\hat{\theta}_n}|\hat{\theta}_n\}$. Therefore, we omit the proof of this part. □

Proof of Theorem 2.3. We use quite standard technique based on Skorokhod's theorem (see [1], Theorem 29.6). We have from (2.5) that $n(\hat{a}_n - 1) \xrightarrow{d} W_0$ as $n \to \infty$. Therefore, due to Skorokhod's theorem there exists a sequence $\{\hat{a}'_n, n \geq 1\}$ of random variables and a random variable W'_0 on a common probability space $(\Omega', \mathcal{F}, Q)$ such that $\hat{a}'_n \stackrel{d}{=} \hat{a}_n$ for all $n \geq 1$, $W'_0 \stackrel{d}{=} W_0$ and $n(\hat{a}'_n(\omega') - 1) \to W'_0$ as $n \to \infty$ for each $\omega' \in \Omega'$.

For any $\omega' \in \Omega'$ we obtain $\hat{\theta}'_n(\omega')$ from equation $a = f_a(\theta)$ as $\hat{\theta}'_n(\omega') = f_a^{-1}(\hat{a}'_n(\omega'))$. Let now $\{X_{ki}'^{(n)}, k, i \geq 1\}$ be a family of i.i.d. random variables such that

$$P\{X_{ki}'^{(n)} = j\} = p_j(\hat{\theta}'_n)$$

for each $\omega' \in \Omega'$ and $n \geq 1$ and $\{\xi_k, k \geq 1\}$ be a sequence of random variables with the probability distributions $\{q_j(k), j \geq 0\}$. We define a new bootstrap process recursively by the relation

$$Z'^{(n)}(k) = \sum_{i=1}^{Z'^{(n)}(k-1)} X_{ki}'^{(n)} + \xi_k, \quad k = 1, 2, ...$$

for each $\omega' \in \Omega'$, $n \geq 1$ with $Z'^{(n)}(0) = 0$. We denote

$$\mathcal{Z}'_n(t) = \frac{Z'^{(n)}([nt])}{A(\hat{a}'_n, n)}, \ \mathcal{Y}'_n(t) = \frac{Z'^{(n)}([nt]) - A(\hat{a}'_n, [nt])}{B(\hat{a}'_n, n)}. \tag{4.4}$$

We introduce for each $E \in \mathcal{B}(D)$ the probability measure $P_n(\theta, E) := P(\mathcal{Z}_n \in E)$, where $\mathcal{Z}_n(t)$ is the normalized original process. Then it is clear that

$$P(\mathcal{Z}_n^{\hat{\theta}_n} \in E|\hat{\theta}_n) = P_n(\hat{\theta}_n, E),\ P(\mathcal{Z}'_n \in E|\hat{\theta}'_n) = P_n(\hat{\theta}'_n, E).$$

We also denote $P(d, E) = P(\pi_\alpha(d, \cdot) \in E)$. Recall that $n(\hat{a}'_n(\omega') - 1) \to W'_0$ as $n \to \infty$ for each $\omega' \in \Omega'$. Therefore, repeating the arguments of the proof of Theorem 2.1, we obtain that

$$\int_D g(x) P_n(\hat{\theta}'_n, dx) \to \int_D g(x) P(W'_0, dx) \tag{4.5}$$

as $n \to \infty$ for any function $g : D(\mathbb{R}_+, \mathbb{R}_+) \mapsto \mathbb{R}$ bounded and continuous in Skorokhod metric.

Since

$$\int_D g(x) P_n(\hat{\theta}'_n, dx) \overset{d}{=} \int_D g(x) P_n(\hat{\theta}_n, dx),\ \int_D g(x) P(W'_0, dx) \overset{d}{=} \int_D g(x) P(W_0, dx),$$

we obtain from (4.5) that

$$\int_D g(x) P_n(\hat{\theta}_n, dx) \xrightarrow{d} \int_D g(x) P(W_0, dx).$$

□

Proof of Theorem 2.4. We consider applicability of Theorem 3.2 to the process $\mathcal{Y}'_n(t)$ defined in (4.4). It follows from the convergence $n(\hat{a}'_n(\omega') - 1) \to W'_0$ as $n \to \infty$, that $\hat{\theta}'_n(\omega') \to \theta_0$ as $n \to \infty$ for each $\omega' \in \Omega'$. Therefore, conditions C1-C3 are trivially satisfied.

We now show that condition C5 is also fulfilled. Since

$$A'_n(k) := E[Z'^{(n)}(k)|\ \hat{a}'_n] = A(\hat{a}'_n, k)$$

and

$$B'^2_n(k) := E[(Z'^{(n)}(k) - A'_n(k))^2|\ \hat{a}'_n] = B^2(\hat{a}'_n, k),$$

applying again Lemma 3.1, we obtain that

$$\frac{A(\hat{a}'_n, [ns])}{n\alpha(n)} \to \mu_\alpha(W'_0, s),\ \frac{\sigma^2(\hat{a}'_n, [ns])}{n\beta(n)} \to \nabla_\beta(W'_0, s),$$

$$\frac{\Delta^2(\hat{a}'_n, [ns])}{n^2\alpha(n) f_b(\hat{\theta}'_n)} \to \frac{1}{W'_0} \nu_\alpha(W'_0, s)$$

as $n \to \infty$ for each $s \in \mathbb{R}_+$ and $\omega' \in \Omega'$. It follows from this and Lemma 3 in [10] that

$$\frac{\Delta^2(\hat{a}'_n, n)}{\Delta^2(1, n)} \to \frac{(\alpha+1)(\alpha+2)}{W'_0} \nu_\alpha(W'_0, 1),\ \frac{\sigma^2(\hat{a}'_n, n)}{\sigma^2(1, n)} \to (\beta+1)\nabla_\beta(W'_0, 1)$$

as $n \to \infty$ for each $\omega' \in \Omega'$. Therefore, for each fixed $\omega' \in \Omega'$ there exists a positive constant $C(\alpha, \beta, W_0'(\omega'))$, such that

$$\frac{B^2(\hat{a}_n', n)}{B^2(1, n)} \to C(\alpha, \beta, W_0') \tag{4.6}$$

as $n \to \infty$. We now consider

$$\delta'^{(2)}(\varepsilon, \omega') := \frac{1}{B^2(\hat{a}_n', n)} \sum_{k=1}^{n} E[(\bar{\xi}_k)^2 \chi(|\bar{\xi}_k| > \varepsilon B(\hat{a}_n', n))].$$

It follows from (4.6) that, if condition C4 is satisfied, then $\delta'^{(2)}(\varepsilon, \omega') \to 0$ as $n \to \infty$ for each fixed $\omega' \in \Omega'$.

To show that

$$\delta_n'^{(1)}(\varepsilon, \omega') = \gamma_1(W_0(\omega')) E[(\bar{X}_{ki}'^{(n)})^2 \chi(|\bar{X}_{ki}'^{(n)}| > \varepsilon B(\hat{a}_n', n))]$$

tends to zero as $n \to \infty$ for each fixed $\omega' \in \Omega'$, we repeat the same arguments as in the proof of part (a) of the Theorem 2.2.

Hence, it follows from Theorem 3.2 that

$$\int_D g(x) Q_n(\hat{\theta}_n', dx) \to \int_D g(x) Q(W_0', dx) \tag{4.7}$$

as $n \to \infty$ for each fixed $\omega' \in \Omega'$, for any function $g : D(\mathbb{R}_+, \mathbb{R}_+) \mapsto \mathbb{R}$ bounded and continuous in Skorokhod metric, where

$$Q_n(\theta, E) := P(\mathcal{Y}_n \in E), \; Q(d, E) := P(W(\psi(d, \cdot) \in E).$$

On the other hand, since $\hat{\theta}_n \stackrel{d}{=} \hat{\theta}_n'$ and $W_0 \stackrel{d}{=} W_0'$, we have

$$\int_D g(x) Q_n(\hat{\theta}_n', dx) \stackrel{d}{=} \int_D g(x) Q_n(\hat{\theta}_n, dx), \; \int_D g(x) Q(W_0', dx) \stackrel{d}{=} \int_D g(x) Q(W_0, dx).$$

Therefore, we obtain from (4.7) that

$$\int_D g(x) Q_n(\hat{\theta}_n, dx) \xrightarrow{d} \int_D g(x) Q(W_0, dx)$$

as $n \to \infty$. □

Proof of Theorem 2.5 is similar to the proof of Theorem 2.2. Here instead of Theorem 3.2 we apply Theorem 3.3. Since conditions of these two theorems are the same, the applicability of the last directly follows from the proof of Theorem 2.2. □

Acknowledgments

I thank the referee and the associate editor for careful reading the first version of the paper and for valuable comments. My sincere thanks to University College of Zayed University, Dubai, UAE, for all the support and facilities I had.

References

[1] P. Billingsley, *Probability and Measure.* Wiley&Sons, New York, 1979.

[2] S. Datta, T.N. Sriram, *A modified bootstrap for branching processes with immigration.* Stoch. Proc. Appl., 56 (1995), 275 – 294.

[3] S. Datta, *Bootstrapping.* In *Encyclopedia of statistical sciences. Second edition, S. Kotz Editor.* Wiley&Sons, New York, 2005.

[4] A.C. Davidson, D.V. Hinkley, *Bootstrap methods and their applications.* Cambridge University Press, Cambridge, 2003.

[5] B. Efron, *Bootstrap methods-another look at the jackknife.* Ann. Statist., 7 (1979), 1 – 26.

[6] B. Efron, R. Tibshirani, *An introduction to the bootstrap.* Chapman and Hall Ltd, New York, 1993.

[7] P. Haccou, P. Jagers, V.A. Vatutin, *Branching processes: Variation, Growth and Extinction of populations.* Cambridge University Press, Cambridge, 2005.

[8] P. Guttorp, *Statistical inference for branching processes.* Wiley, New York, 1991.

[9] S.N. Lahiri, *Bootstrap methods: a review.* In *Frontiers in Statistics. J. Fan and H. Koul Eds.* Imperial College Press, London, 2006.

[10] I. Rahimov, *Functional limit theorems for critical processes with immigration.* Adv. Appl. Probab., 39, 4 (2007), 1054 – 1069.

[11] I. Rahimov, *Limit distributions for weighted estimators of the offspring mean in a branching process.* TEST, 18, 3 (2009), 568 – 583.

[12] I. Rahimov *Approximation of fluctuations in a sequence of nearly critical branching processes.* Stochastic models, 25, 2 (2009), 348 – 373.

[13] I. Rahimov *Approximation of a sum of martingale-differences generated by a bootstrap branching process.* Workshop on Branching Processes and their Applications, April 20-23, 2009, Badajoz, Spain.

[14] J. Shao, D. Tu, *The jackknife and bootstrap.* Springer, New York, 1995.

[15] T.N. Sriram, *Invalidity of bootstrap for critical branching processes with immigration.* Ann. Statist., 22 (1994), 1013 – 1023.

[16] T.J. Sweeting, *On conditional weak convergence.* J. Theoret. Probability, 2, 2 (1989), 461 – 474.

[17] S. Xiong, G. Li *Some results on the convergence of conditional distributions.* Stat. Prob. Letters, 78 (2008), 3249 – 3253.

Ibrahim Rahimov
Department of Mathematics and Statistics
Zayed University
P.O.Box 19282 Dubai, United Arab Emirates
E-mail: Ibrahim.Rahimov@zu.ac.ae

Received: 02.12.2010