Asymptotic Properties of Near Pfeifer Records

ARUP BOSE* and SREELA GANGOPADHYAY†

Indian Statistical Institute, Kolkata

Abstract

Asymptotic properties of the number of near records is known in the literature. We generalize these results to the Pfiefer model which has a wider application. In particular we establish convergence in probability, in the almost sure sense and in distribution for the number of near records under the Pfiefer model.

June 20, 2009

Keywords. Pfeifer records, Near Pfeifer records, Regularly varying functions, Kolmogorov maximal inequality, Limiting distribution.

AMS 2000 Subject Classification. 60F05, 60F15, 60G70.

Address for correspondence:

SREELA GANGOPADHYAY
Statistics and Mathematics Unit
Indian Statistical Institute
203 B.T. Road
Kolkata 700108
India
email address: gangopadhyay7@gmail.com
Alternate email address: res9616@isical.ac.in
Fax: 011 91 33 2577 6035
Alternate fax: 011 91 33 2577 3071.
Phone: 011 91 33 2575 4310
Alternate phone: 011 91 33 2575 3400.

---

*Work done while the author was visiting Department of Economics, University of Cincinnati, USA. Research also supported by J.C.Bose Fellowship, Government of India.

†Research supported by the Department of Science and Technology, Government of India, scheme no. SR/WOS-A/MS-02/2006.

1 Introduction

Insurance companies often change their policy when they receive a claim which exceeds all previous claims, so that under the new policy such claims would be less frequent in probability. The Pfiefer model of records, Pfeifer [12]), has been used to model such situations. It is of interest to the company to study the total value and number of claims which are "very near" to the record claims. Balakrishnan et. al. [3] and Pakes [10] have studied properties of near records in classical setup where observations are i.i.d. with a common continuous distribution. See also Li [6], Li and Pakes [7] and Pakes and Steutel [11]. In this article we study properties of near records for the more general Pfeifer model (Pfeifer [12]) which we refer to as *near precords* for short. The setup is as follows:

Let $\{X_{ij}\}_{i\geq 0, j\geq 1}$ be a double array of independent random variables. For each fixed i, $\{X_{ij}, j \geq 1\}$ are i.i.d. with a common cdf F_i where

$$1 - F_i = (1 - F_0)^{\alpha_i}, \ \forall i = 1, 2, \ldots$$

for some sequence of positive reals $\{\alpha_i\}$.

The first precord R_1^p is by convention X_{01}. Now consider the row $i = 1$. Let

$$\Delta(1) = \inf\{j : X_{1j} > R_1^p\}.$$

Having defined $\Delta(n)$, inductively define

$$\Delta(n+1) = \inf\{j : X_{n+1,j} > X_{n,\Delta(n)}\}, \quad n \geq 1.$$

Then $X_{0,1}, X_{1,\Delta(1)}, \ldots X_{n,\Delta(n)}, \ldots$ are precords denoted by $R_1^p, R_2^p, \ldots R_{n+1}^p, \ldots$.

Let $a > 0$. The number of *near precords* is a sequence of non-negative integers, $\{\xi_n^p(a)\}_{n\geq 1}$, depending on a, defined as

$$\xi_n^p(a) = \#\{j < \Delta(n) : R_n^p - a < X_{nj} < R_n^p\}.$$

For a sequence $\{a_n\}$, the corresponding $\{\xi_n^p(a_n)\}$ will be called the number of *near precords with varying window width*.

The following representation for precords plays a crucial role in our analysis. Suppose $Y_1, Y_2, \ldots$ are independent random variables and

$$Y_i \sim Exp(\alpha_i) \tag{1.1}$$

and for any distribution function F,

$$\psi_F(x) = F^{-1}(1 - e^{-x}). \tag{1.2}$$

Then

$$\left(R_1^p, R_2^p, \ldots R_n^p\right) \stackrel{\mathcal{D}}{=} \left(\psi_{F_0}(Y_1), \psi_{F_0}(Y_1 + Y_2), \ldots \psi_{F_0}(Y_1 + Y_2 + \ldots Y_n)\right) \text{ for all } n,$$

where F_0 is the cdf of X_{01}, the basic underlying cdf of precords. See Arnold et. al. [2]. Unless otherwise stated we will assume that F_0 *is a continuous, strictly increasing cdf with support,* $Supp(F_0) \subset [0, \infty)$ *and* α_n *non-decreasing positive reals, diverging to infinity*. For convenience, we will denote ψ_{F_0} *by* ψ_0 in the sequel.

In Section 2, we derive the distributions of $\xi_n^p(a)$. In Section 3 we study asymptotic properties of $\xi_n^p(a)$ under various conditions on F_0 and α_n. In Section 4, we study the limiting distributions of normalised $\xi_n^p(a_n)$ with varying window width.

2 Distribution of $\xi_n^p(a)$

Balakrishnan et. al. [3] have derived the distribution of $\xi_n(a)$ in the i.i.d. model. Using similar arguments we have the following basic formula for the joint distribution of the number of near records in the Pfeifer model. We need a few notation. For any cdf F,

$$\bar{F}(x) = 1 - F(x).$$

Note that

$$\bar{F_n}(x) = \bar{F_0}^{\alpha_n}(x) \text{ for all } n \geq 0.$$

Let

$$R_n^p(x) = P(R_n^P \leq x) \text{ and } \rho_n(x,a) = \frac{\bar{F_n}(x)}{\bar{F_n}(x-a)}.$$

Let

$$H_i(x_{j+i}, x_j) = P\left(0 < Z < -\log\frac{\bar{F}_{j+i}(x_{j+i})}{\bar{F}_j(x_j)}\right)$$

where $Z \stackrel{\mathcal{D}}{=} \sum_{k=j+1}^{j+i} \beta_k X_k^*$, where $\beta_k = \frac{1}{\alpha_k}$ and X_k^* are i.i.d. Exp(1) r.v.s.

Theorem 1 *(i)* $P(\xi_n^p(a) = k) = \int_R \rho_n(x,a)(1-\rho_n(x,a))^k dR_n^p(x)$.
(ii) The joint distribution of $(\xi_n^p(a_1), \xi_{n+1}^p(a_2), \ldots \xi_{n+k-1}^p(a_k))$ *is given by*

$$P(\xi_n^p(a_1) = r_1, \ldots \xi_{n+k-1}^p(a_k) = r_k)$$
$$= \int_{-\infty}^{\infty}\int_{x_n}^{\infty}\cdots\int_{x_{n+k-2}}^{\infty}\prod_{j=1}^{k}\rho_{n+j-1}(x_{n+j-1},a_j)(1-\rho_{n+j-1}(x_{n+j-1},a_j))^{r_j}\prod_{j=2}^{k}dH_1(x_{n+j-1},x_{n+j-2})dR_n^p(x_n).$$

(iii) The joint distribution of $(\xi_n(a_1), \xi_{n+k}(a_{k+1}))$ *is given by*

$$P(\xi_n^p(a_1) = r_1, \xi_{n+k}^p(a_{k+1}) = r_{k+1})$$
$$= \int_{-\infty}^{\infty}\int_{x_n}^{\infty}\rho_n(x_n,a_1)(1-\rho_n(x_n,a_1))^{r_1}\rho_{n+k}(x_{n+k},a_{k+1})(1-\rho_{n+k}(x_{n+k},a_{k+1}))^{r_{k+1}}dH_k(x_{n+k},x_n)dR_n^p(x_n).$$

Proof (i) Clearly, $P(\xi_n^p(a) = k) = \int_R P(\xi_n^p(a) = k | R_n^p = x) dR_n^p(x)$. Using Nevzorov's [8] deletion argument we compute $P(\xi_n^p(a) = k | R_n^p = x)$ as follows.

On the nth row of the rectangular array, we delete those observations X_{nj} such that $X_{nj} \leq x - a$. The remaining observations on the nth row are all greater than $x - a$ and are conditionally independent given $R_n^p = x$. Denoting these remaining observations by Y_{nj}, we have

$$\begin{aligned} P(Y_{nj} \leq y) =& P(X_{nj} \leq y | X_{nj} > x - a) \\ =& \frac{F_n(y) - F_n(x-a)}{\bar{F_n}(x-a)} = 1 - \frac{\bar{F_n}(y)}{\bar{F_n}(x-a)}. \end{aligned}$$

Therefore,

$$\begin{aligned} P(\xi_n^p(a) = k | R_n^p = x) =& P(Y_{n1} \leq x, Y_{n2} \leq x, \ldots Y_{nk} \leq x, Y_{nk+1} > x) \\ =& [1 - \rho_n(x,a)]^k \rho_n(x,a). \end{aligned}$$

and (i) follows. The proofs of (ii) and (iii) follow from the two simple observations:

(a) For any finite increasing sequence $n_1 < n_2 < ..... < n_i$ and any positive reals $a_1, a_2, \ldots, a_i$, $\xi^p_{n_1}(a_1), \xi^p_{n_2}(a_2), \ldots, \xi^p_{n_i}(a_i)$ are conditionally independent given $R^p_{n_1}, R^p_{n_2}, \ldots, R^p_{n_i}$.

and

(b) $H_i(x_{j+i}, x_j)$ is the transition probability function $P(R^p_{j+i} \leq x_{j+i}|R^p_j = x_j)$. Hence

$$\begin{aligned} P(R^p_{j+i} \leq x_{j+i}|R^p_j = x_j) &= P(\sum_{k=1}^{j+i} \beta_k X^*_k \leq \psi^{-1}_{F_0}(x_{j+i})|\sum_{k=1}^{j} \beta_k X^*_k = \psi^{-1}_{F_0}(x_j)) \\ &= P(\sum_{k=j+1}^{j+i} \beta_k X^*_k \leq \psi^{-1}_{F_0}(x_{j+i}) - \psi^{-1}_{F_0}(x_j)) \\ &= P(\sum_{k=j+1}^{j+i} \beta_k X^*_k \leq -\log \frac{\bar{F}_0(x_{j+i})}{\bar{F}_0(x_j)}). \end{aligned}$$

□

3 Asymptotic behaviour of $\xi^p_n(a)$

Let

$$r_{F_0} = \sup\{support(F_0)\}, \;\; l_{F_0} = \inf\{support(F_0)\} \;\text{ and }\; \beta_n = \frac{1}{\alpha_n}.$$

Note that r_{F_0} may equal ∞. Balakrishnan et. al. [3] have shown that in the i.i.d. model, if $r_{F_0} < \infty$, then $\xi_n(a) \to \infty$ almost surely (a.s.) as $n \to \infty$. The fact that the records R_n have the closed form density $\frac{dR_n}{dF} = \frac{1}{(n-1)!}(-\log \bar{F}(x))^{n-1}$ plays a crucial role in the above work.

Though in the Pfeifer model $\frac{dR^p_n}{dF_0}$ exists, it is not known in a closed form. Hence their arguments cannot be used, unless $\{\alpha_n\}$ are constant for all n. Our goal is to establish some of their results for $\{\xi^p_n(a)\}$ under the Pfiefer model. We need the following Lemma.

Lemma 1 *(i) If* $\sum_{n=1}^{\infty} \beta_n = \infty$ *then* $R^P_n \to r_{F_0}$ *a.s..*

(ii) If $\sum_{n=1}^{\infty} \beta_n < \infty$ *then* $R^p_n \to W$ *a.s. where* W *is a continuous random variable with the same support as* F_0 *and with a strictly increasing cdf on its support.*

Proof (i) Since $\sum_{i=1}^{n} \beta_i \to \infty$, $\sum_{i=1}^{n} \beta_i X^*_i \to \infty$ a.s. where X^*_i are i.i.d. $Exp(1)$, by the Kolmogorov Three-series Theorem. Therefore, $\psi_0(\sum_{i=1}^{n} \beta_i X^*_i) \to r_{F_0}$ a.s. Hence $R^p_n \overset{\mathcal{D}}{=} \psi_0(\sum_{i=1}^{n} \beta_i X^*_i) \to r_{F_0}$ in distribution and hence in probability. Since R^p_n are increasing, this convergence holds a.s..

(ii) If $\sum_{i=1}^{n} \beta_i < \infty$ then $\sum_{i=1}^{n} \beta_i^2 < \infty$. Therefore, $\sum_{i=1}^{n} Var(\beta_i X^*_i - \beta_i) = \sum_{i=1}^{n} \beta_i^2 < \infty$. Khinchine-Kolmogorov's 1-series Theorem implies that $\sum_{i=1}^{n} \beta_i X^*_i \to V$ a.s. where V is a finite random variable.

Note that $P(V > K) \geq P(\beta_1 X^*_1 > K) > 0$ for any $K > 0$, however large. So V is a non-degenerate GGC (generalised gamma convolution) and hence has a strictly positive pdf for $x > l_V = \inf\{support(V)\}$. See Bondesson [4], page 30. Therefore V, an absolutely continuous random variable, has a strictly increasing cdf on its support.

Consequently, $R_n^p \to W = \psi_0(V)$ in distribution and hence a.s., by monotonicity of R_n^p. F_0 is continuous and strictly increasing on (l_{F_0}, r_{F_0}) by our assumption. So W is continuous and has strictly increasing cdf on $(\psi_0(l_V), r_{F_0})$. Now, for any $\epsilon > 0$,

$$P(V < \epsilon) \geq \prod_{i=1}^{n} P(\beta_i X_i^* < \frac{\epsilon}{2^{i+1}}).P(\sum_{i=n+1}^{\infty} \beta_i X_i^* < \frac{\epsilon}{2}). \tag{3.1}$$

By Kolmogorov's maximal inequality,

$$P(\sum_{i=n+1}^{\infty} \beta_i X_i^* < \sum_{i=n+1}^{\infty} \beta_i + \frac{\epsilon}{4}) \geq \lim_{m\to\infty} P(\max_{n+1\leq k\leq m} |\sum_{i=n+1}^{k} (\beta_i X_i^* - \beta_i)| < \frac{\epsilon}{4}) \geq 1 - \frac{\sum_{i=n+1}^{\infty} \beta_i^2}{(\frac{\epsilon}{4})^2}. \tag{3.2}$$

Since $\sum_{i=1}^{n} \beta_i$ converges, we choose n_0 so large that $\sum_{i=n_0+1}^{\infty} \beta_i < \frac{\epsilon}{4}$ and $\frac{\sum_{i=n_0+1}^{\infty} \beta_i^2}{(\frac{\epsilon}{4})^2} < \frac{1}{2}$.
Hence, from (3.2), $P\left(\sum_{i=n_0+1}^{\infty} \beta_i X_i^* < \frac{\epsilon}{2}\right) > \frac{1}{2}$. Further, for $1 \leq i \leq n_0$, $P(\beta_i X_i^* < \frac{\epsilon}{2^{i+1}}) > 0$. It follows from (3.1), $P(V < \epsilon) > 0$. Therefore $l_V = 0$ and hence $support(W) = (l_{F_0}, r_{F_0}) = support(F_0)$. □

From now on, we will assume without loss of generality, $l_{F_0} = 0$.

3.1 Case 1. $r_{F_0} < \infty$

Proposition 1 *(i) $\xi_n^p(a) \to \infty$ in probability.*

(ii) Suppose, $\alpha_n \uparrow \infty$ in such a way that (a) $\sum_{n=1}^{\infty}(1-\lambda)^{\alpha_n} < \infty,\ \forall\ \lambda > 0$ and (b) $\sum_{n=1}^{\infty} \frac{\alpha_n^n}{n!} < \infty$. Then $\xi_n^p(a) \to \infty$ a.s.

Proof. (i) It is enough to show that for any fixed $k > 0$, $P(\xi_n^p(a) \leq k) \to 0$, as $n \to \infty$. Observe that

$$\begin{aligned} P(\xi_n^p(a) \leq k) &= \sum_{j=0}^{k} \int_{-\infty}^{\infty} \rho_n(x,a)(1-\rho_n(x,a))^j dR_n^p(x) \\ &= \int_{-\infty}^{\infty} [1-(1-\rho_n(x,a))^{k+1}] dR_n^p(x) \\ &\leq (k+1) \int_{-\infty}^{\infty} \rho_n(x,a) dR_n^p(x) \\ &= (k+1) E(\rho_n(R_n^p, a)). \end{aligned}$$

Recall that

$$\rho_0(x,a) = \frac{\bar{F}_0(x)}{\bar{F}_0(x-a)} \text{ and } \rho_n(x,a) = \rho_0(x,a)^{\alpha_n}.$$

Fix any $\epsilon \in (0, r_{F_0})$.

$$\begin{aligned} P(\xi_n^p \leq k) &\leq (k+1) E(\rho_n(R_n^p, a)) \\ &= (k+1) E(\rho_n(R_n^p, a) I_{R_n^p < \epsilon}) + (k+1) E(\rho_n(R_n^p, a) I_{R_n^p \geq \epsilon}) \\ &= (k+1)[E_{1,n,\epsilon} + E_{2,n,\epsilon}] \text{ (say).} \end{aligned}$$

Since $0 \leq \rho_n(R_n^p, a) \leq 1$,

$$E_{1,n,\epsilon} \leq P(R_n^p < \epsilon). \tag{3.3}$$

Recall Lemma 1. If $R_n^p \to r_{F_0}$ a.s. then the probability in (3.3) converges to 0.

On the other hand, if $R_n^p \to W$ a.s. (see Lemma 1) then this probability converges to $P(W < \epsilon)$. Since W is continuous and greater than 0 a.s., by choosing ϵ small enough, we can make $P(W < \epsilon)$ and consequently, $E_{1,n,\epsilon}$ as small as we like, for all n large enough. For $E_{2,n,\epsilon}$, observe that for any $\epsilon \in (0, r_{F_0})$, $\rho_0(x,a) \leq (1-\lambda), \ \forall \ x \in [\epsilon, r_{F_0}]$, for some $\lambda > 0$ depending on ϵ. Therefore, $E_{2,n,\epsilon} \leq (1-\lambda)^{\alpha_n} \to 0$, as by our assumption, $\alpha_n \uparrow \infty$.

(ii) Now it is enough to show that $\sum_n (E_{1,n,\epsilon} + E_{2,n,\epsilon}) < \infty$ for small enough ϵ.

Condition (a) implies that for any $\epsilon > 0$, $\sum_n E_{2,n,\epsilon} < \infty$.

Now observe that,

$$\begin{aligned} E_{1,n,\epsilon} &\leq P(R_n^p < \epsilon) \\ &= P(\sum_{i=1}^n \beta_i X_i^* < \epsilon) \\ &\leq P(\sum_{i=1}^n X_i^* < \epsilon\alpha_n) \text{ since } \alpha_i \text{ are non-decreasing.} \end{aligned}$$

As $\sum_{i=1}^n X_i^* \sim \Gamma(n)$, the last probability above has an upper bound $\frac{(\epsilon\alpha_n)^n}{n!}$ which is summable by condition (b), if we choose $\epsilon \in (0,1)$. Now using Borel-Cantelli Lemma we have the desired result. □

Example. Let $\alpha_n = n^\delta, \ \ 0 < \delta < 1$. Then conditions (a) and (b) of Proposition 1 are satisfied.

We now show that with an appropriate random scaling, $\xi_n^p(a)$ has an exponential distribution. The result is analogous to that in Balakrishnan et. al. ([3], Theorem 3.1(ii)), the proof is also similar.

Theorem 2 *$\rho_n(R_n^p, a)\xi_n^p(a) \xrightarrow{\mathcal{D}} \mathcal{E}$ where $\mathcal{E} \sim Exp(1)$.*

Proof. Consider the moment generating function of $\rho_n(R_n^p, a)\xi_n^p(a)$.

$$\begin{aligned} E\left[e^{-\theta\rho_n(R_n^p,a)\xi_n^p(a)}\right] &= E[E(e^{-\theta\rho_n(R_n^p,a)\xi_n^p(a)})|R_n^p] \\ &= \int_{\mathcal{R}} \sum_{k=0}^{\infty} e^{k(-\theta\rho_n(x,a))}\rho_n(x,a)(1-\rho_n(x,a))^k dR_n^p(x) \\ &= E_{R_n^p}\left[\frac{\rho_n(R_n^p,a)}{1-e^{-\theta\rho_n(R_n^p,a)}(1-\rho_n(R_n^p,a))}\right]. \end{aligned} \tag{3.4}$$

Now, $\rho_0(R_n^p, a) \to 0$ if $R_n^p \to r_{F_0}$ a.s. and $\rho_0(R_n^p, a) \to \rho_0(W, a) < 1$ a.s. if $R_n^p \to W$ a.s. Since $\alpha_n \uparrow \infty$, in both the cases $\rho_n(R_n^p, a) \to 0$ a.s. Therefore, $e^{-\theta\rho_n(R_n^p,a)} = 1 - \theta\rho_n(R_n^p, a) + o(\rho_n(R_n^p, a))$. Using this in (3.4) we get

$$E\left[e^{-\theta\rho_n(R_n^p,a)\xi_n^p(a)}\right] = E\left[\frac{1}{1+\theta+\frac{o(\rho_n(R_n^p,a))}{\rho_n(R_n^p,a)}}\right].$$

As $0 \leq \rho_n(R_n^p, a) \leq 1$, the integrand is bounded for any fixed θ with $|\theta| < 1$ and hence by Bounded Convergence Theorem the above expression converges to $\frac{1}{1+\theta}$ which is the moment generating function of $\mathcal{E}(1)$. □

Remark. It follows that R_n^p and $\rho_n(R_n^p, a)\xi_n^p(a)$ are asymptotically independent.

Corollary 1 *(i) If* $\sum_{i=1}^n \beta_i^2 \to \infty$*, then*

$$\frac{\log \xi_n^p(a) - \alpha_n \sum_{i=1}^n \beta_i}{\alpha_n \sqrt{\sum_{i=1}^n \beta_i^2}} \xrightarrow{\mathcal{D}} N(0,1).$$

(ii) If $\sum_{i=1}^\infty \beta_i^2 < \infty$ *and* $\sum_{i=1}^n \beta_i \to \infty$*, then*

$$\frac{\log \xi_n^p(a) - \alpha_n \sum_{i=1}^n \beta_i}{\alpha_n} \xrightarrow{\mathcal{D}} V_1.$$

(iii) If $\sum_{i=1}^\infty \beta_i < \infty$*, then*

$$\frac{\log \xi_n^p(a)}{\alpha_n} \xrightarrow{\mathcal{D}} V_2.$$

Here V_1 *and* V_2 *are continuous random variables with densities.*

Proof. (i) By Theorem 2, $\rho_n(R_n^p, a)\xi_n^p(a) \xrightarrow{\mathcal{D}} \mathcal{E}$. Taking log on both sides we have

$$\alpha_n \log(\bar{F}_0(R_n^p)) - \alpha_n \log(\bar{F}_0(R_n^p - a)) + \log \xi_n^p(a) \xrightarrow{\mathcal{D}} \log \mathcal{E}. \tag{3.5}$$

Now, $-\log(\bar{F}_0(R_n^p)) \stackrel{\mathcal{D}}{=} \sum_{i=1}^n \beta_i X_i^*$, where $X_i^* \sim Exp(1)$. Since $\sum_{i=1}^n \beta_i^2 \to \infty$, by CLT

$$\frac{\sum_{i=1}^n \beta_i X_i^* - \sum_{i=1}^n \beta_i}{\sqrt{\sum_{i=1}^n \beta_i^2}} \xrightarrow{\mathcal{D}} N(0,1).$$

Note that $\bar{F}_0(R_n^p - a) \to \bar{F}_0(r_{F_0} - a)$, a constant (>0) a.s. Therefore, adding and subtracting $\alpha_n \sum_{i=1}^n \beta_i$ on left side of (3.5) and dividing both sides by $\alpha_n \sqrt{\sum_{i=1}^n \beta_i^2}$ we get (i).

For (ii) and (iii), we use similar arguments as in (i) and use the fact that $\sum_{i=1}^n \beta_i(X_i^* - 1)$ converges a.s. to $S = \sum_{i=1}^\infty \beta_i(X_i^* - 1)$, if $\sum_{i=1}^\infty \beta_i^2 < \infty$. Since S is a convolution of $\beta_1(X_1^* - 1)$ and $\sum_{i=2}^\infty \beta_i(X_i^* - 1)$ and the former has a density, we infer S has a density and consequently V_1 and V_2 have densities. □

3.2 Case 2. $r_{F_0} = \infty$

In this case, the asymptotic behaviour of $\xi_n^p(a)$ depends on $\rho(a) = \lim_{x\to\infty} \rho_0(x,a)$.

If $\rho(a) < 1$ or $R_n^p \to W$, then $\rho_n(R_n^p, a) \to 0$ a.s. and arguments of Proposition 1 and Theorem 2 go through and we obtain the following Theorem. We omit the details of the arguments.

Theorem 3 *If* $\rho(a) < 1$ *or,* $\rho(a) = 1$ *but* $R_n^p \to W$*, then*

(i) $\xi_n^p(a) \to \infty$ *in probability and*

(ii)$\rho_n(R_n^p, a)\xi_n^p(a) \xrightarrow{\mathcal{D}} \mathcal{E}$.

Unfortunately, when $\rho(a) = 1$ and $R_n^p \to \infty$, i.e. $\sum_{i=1}^\infty \beta_i = \infty$, the situation becomes quite complicated. We deal with a special case below. Let $\bar{F}_0(x) = \frac{1}{x^\gamma}$, as $x \to \infty$ (where $\gamma > 0$ is a constant), then

$\rho(a) = 1$. Further,

$$\rho_n(x,a) = \rho_0(x,a)^{\alpha_n} = \left(1 - \frac{a}{x}\right)^{\gamma\alpha_n}. \tag{3.6}$$

Theorem 4 *Assume $\bar{F}_0(x) = \frac{1}{x^\gamma}$, as $x \to \infty$ and $\sum_{i=1}^{\infty} \beta_i = \infty$. If*

$$\lim_{n\to\infty} \frac{q\gamma \log \alpha_n - \sum_{i=1}^{n} \beta_i}{\sqrt{\sum_{i=1}^{n} \beta_i^2}} = -\infty, \text{ for some constant } q > 1$$

then $P(\xi_n^p(a) = 0) \to 1$.

Proof. Observe that, for all $x > \frac{a}{1-\left(1-\beta_n^q\right)^{\frac{1}{\gamma}}}$ $(= x_n)$, $\left(1-\frac{a}{x}\right)^\gamma > 1-\beta_n^q$. By the Mean Value Theorem,

$$x_n = \frac{a}{\beta_n^q \frac{1}{\gamma} \gamma_n^{\frac{1}{\gamma}-1}}\ , \tag{3.7}$$

where $1-\beta_n^q < \gamma_n < 1$, so that $\gamma_n \to 1$. Now from (3.6) it follows that, for all $x > x_n$, $\rho_n(x,a) > \left(1-\beta_n^q\right)^{\alpha_n} \to 1$ as $\alpha_n \uparrow \infty$. Therefore, given $\epsilon > 0$, there exists n_0, depending on q and ϵ, such that for all $n > n_0$, $\rho_n(x,a) > 1-\epsilon$, $\forall\ x > x_n$.

Fix any $\epsilon > 0$ and a $q > 1$ for which the condition of the theorem holds. Then for all $n > n_0$,

$$P(\xi_n^p(a) = 0) = \int_0^\infty \rho_n(x,a) dR_n^p(x) > \int_{x_n}^\infty \rho_n(x,a) dR_n^p(x) > (1-\epsilon) P(R_n^p > x_n). \tag{3.8}$$

Also,

$$P(R_n^p > x_n) = P\left[\sum_{i=1}^n \beta_i X_i^* > \psi_0^{-1}(x_n)\right] = P\left[\sum_{i=1}^n \beta_i X_i^* > -\log \bar{F}_0(x_n)\right]. \tag{3.9}$$

Using the expression (3.7) for x_n, we have

$$\begin{aligned} -\log \bar{F}_0(x_n) &= -q\gamma \log \beta_n + \gamma \log \gamma - \gamma(\frac{1}{\gamma} - 1) \log \gamma_n + \gamma \log a \\ &= q\gamma \log \alpha_n + k + o(1), \end{aligned}$$

where $k(= \gamma \log(\gamma a))$ is a constant. Therefore, the probability on the right side of (3.9) can be written as

$$P\left[\frac{\sum_{i=1}^n \beta_i X_i^* - \sum_{i=1}^n \beta_i}{\sqrt{\sum_{i=1}^n \beta_i^2}} > \frac{q\gamma \log \alpha_n - \sum_{i=1}^n \beta_i}{\sqrt{\sum_{i=1}^n \beta_i^2}} + \frac{k + o(1)}{\sqrt{\sum_{i=1}^n \beta_i^2}}\right]. \tag{3.10}$$

Now two cases can occur.

Case 1. $\sum_{i=1}^n \beta_i^2 \to \infty$. Then by the Central Limit Theorem, the left side of the inequality in (3.10) converges in distribution to standard normal variate.

Case 2. $\sum_{i=1}^\infty \beta_i^2 < \infty$. Then by the Khinchine-Kolmogorov Theorem, the left side of the inequality in (3.10) converges in distribution to a random variable Z. That $support(Z) = (-\infty, \infty)$ is easy to prove using Kolmogorov's maximal inequality and the fact that $\sum_{i=1}^n \beta_i \to \infty$.

It is clear from (3.8), (3.9) and (3.10) that both in Case 1 and Case 2, the claim of the theorem follows immediately under the given condition. □

Example. If $\alpha_n = n^\delta$ and $\delta < 1$ then the condition of Theorem 4 is satisfied for all $\gamma > 0$. If $\delta = 1$ then for $\gamma < 1$ we can find a $q > 1$ such that $q\gamma < 1$ and hence the condition of Theorem 4 is satisfied.

Remark. If the condition of Theorem 4 is not satisfied then it appears to be hard to conclude anything about the behaviour of $P(\xi_n^p(a) = 0)$. However it is easy to see that

if

$$\limsup_{n\to\infty} \frac{q\gamma\alpha_n - \sum_{i=1}^n \beta_i}{\sqrt{\sum_{i=1}^n \beta_i^2}} < \infty \text{ for some } q > 1$$

then $\liminf_{n\to\infty} P(\xi_n^p(a) = 0) > 0$.

4 Limiting distribution of number of near precords with varying window width

Pakes [10] obtained several interesting limit theorems for number of near records with varying window width in classical setup. In this section we investigate whether analogous results can be derived for the Pfeifer model under some suitable conditions on F_0 and α_n. *Throughout we will assume* $r_{F_0} = \infty$. First we will prove a lemma which we will use in the sequel.

Lemma 2 *Suppose ψ_0 is regularly varying with index $\nu > 0$ such that the derivative ψ_0' exists and is monotone.*

(i) If $\sum_{i=1}^n \beta_i^2 \to \infty$

$$\frac{R_n^p - \psi_0(\sum_{i=1}^n \beta_i)}{(\sqrt{\sum_{i=1}^n \beta_i^2})\psi_0'(\sum_{i=1}^n \beta_i)} \stackrel{\mathcal{D}}{\to} N(0,1).$$

(ii) If $\sum_{i=1}^\infty \beta_i^2 < \infty$ *and* $\sum_{i=1}^\infty \beta_i = \infty$*, then*

$$\frac{R_n^p - \psi_0(\sum_{i=1}^n \beta_i)}{\psi_0'(\sum_{i=1}^n \beta_i)} \stackrel{\mathcal{D}}{\to} Z,$$

where Z is a finite random variable.

Proof. Observe that by the Mean Value Theorem

$$R_n^p - \psi_0(\sum_{i=1}^n \beta_i) \stackrel{\mathcal{D}}{=} (\sum_{i=1}^n \beta_i X_i^* - \sum_{i=1}^n \beta_i)\psi_0'(V_n^*) \tag{4.1}$$

where V_n^* is a random variable lying between $\sum_{i=1}^n \beta_i X_i^*$ and $\sum_{i=1}^n \beta_i$. By our assumption on ψ_0 and ψ_0' we have ψ_0' is regularly varying with index $\nu - 1$. Therefore, for any $u_n \to \infty$, and $y > 0$, $\lim_{u_n\to\infty} \frac{\psi_0'(u_n y)}{\psi_0'(u_n)} \to y^{\nu-1}$. So the limit is continuous and by our assumption $\psi_0'(u_n y)$ is monotone in y. Hence it follows that the convergence of $\frac{\psi_0'(u_n y)}{\psi_0'(u_n)}$ is locally uniform in y, (see Resnick [13], page 1). Also since $\sum_{i=1}^n \beta_i \to \infty$, by SLLN, $\frac{\sum_{i=1}^n \beta_i X_i^*}{\sum_{i=1}^n \beta_i} \to 1$ a.s. Therefore by the above discussion and setting $\sum_{i=1}^n \beta_i = u_n$ and $\frac{\sum_{i=1}^n \beta_i X_i^*}{\sum_{i=1}^n \beta_i} = y_n$, we have,

$$\frac{\psi_0'(\sum_{i=1}^n \beta_i X_i^*)}{\psi_0'(\sum_{i=1}^n \beta_i)} = \frac{\psi_0'((\sum_{i=1}^n \beta_i)\frac{\sum_{i=1}^n \beta_i X_i^*}{\sum_{i=1}^n \beta_i})}{\psi_0'(\sum_{i=1}^n \beta_i)} = \frac{\psi_0'(u_n y_n)}{\psi_0'(u_n)} \to 1, \ \ a.s.$$

Then by monotonicity of ψ_0' and the fact that V_n^* lies between $\sum_{i=1}^n \beta_i X_i^*$ and $\sum_{i=1}^n \beta_i$ it follows,

$$\frac{\psi_0'(V_n^*)}{\psi_0'(\sum_{i=1}^n \beta_i)} \to 1 \ \ a.s. \tag{4.2}$$

Now, (i) follows by (4.1), (4.2) and CLT for $\frac{\sum_{i=1}^n \beta_i(X_i^*-1)}{\sqrt{\sum_{i=1}^n \beta_i^2}}$. (ii) follows from (4.1), (4.2) and Khinchine-Kolmogorov's 1-series Theorem for $\sum_{i=1}^n \beta_i(X_i^*-1)$. $\square$

Using the above lemma we can find the limiting distribution of $\xi_n^p(a_n)$ for a suitable choice of $\{a_n\}$. In the following, we will denote ψ_0^{-1} by Λ_0. Observe that if ψ_0' exists and is monotone then Λ_0' exists and is monotone, as $\Lambda_0'(x) = \frac{1}{\psi_0'(\psi_0^{-1}(x))}$, (recall that ψ_0 is continuous and strictly increasing by our assumption and hence $\psi_0' > 0$ on $(0,\infty)$, if it exists). The following theorem gives limiting distribution for $\xi_n^p(a_n)$.

Theorem 5 *Suppose ψ_0 is regularly-varying with index $\nu > 0$ and the derivative ψ_0' exists and is monotone. Let $a > 0$ be given. Then there exists a sequence of positive reals $\{a_n\}$ such that $\xi_n^p(a_n) \to \xi$ in distribution, where*

(i) ξ has geometric distribution with parameter e^{-a}, if $\sum_{i=1}^n \beta_i \to \infty$.

(ii) ξ has mixed geometric distribution, with probability generating function $E(s^\xi) = E\big[\frac{e^{-a\Lambda_0'(W)}}{1-s(1-e^{-a\Lambda_0'(W)})}\big]$ where W is the almost sure limit of R_n^p, if $\sum_{i=1}^\infty \beta_i < \infty$.

Proof. Let $\{a_n\}$ be a sequence of positive reals to be chosen later suitably. Easy computation yields

$$E(s^{\xi_n^p(a_n)}) = E\left[\frac{\rho_n(R_n^p, a_n)}{1 - s(1-\rho_n(R_n^p, a_n)))}\right].$$

Observe that for any fixed s with $|s| < 1$ the integrand is bounded since $0 \le \rho_n(R_n^p, a_n) \le 1$. Hence $\{\frac{\rho_n(R_n^p, a_n)}{1-s(1-\rho_n(R_n^p(a_n)))}\}$ is uniformly integrable. So it is enough to show that $\rho_n(R_n^p, a_n))$ converges in distribution to the appropriate limit. First we will prove (i). Here two cases can arise.

Case 1. $\sum_{i=1}^n \beta_i^2 \to \infty$. For this case, define $A_n = \psi_0(\sum_{i=1}^n \beta_i)$, $B_n = (\sqrt{\sum_{i=1}^n \beta_i^2})\psi_0'(\sum_{i=1}^n \beta_i)$ and $Z_n = \frac{R_n^p - A_n}{B_n}$. Let $g_n(x) = B_n x + A_n$ so that $g_n(Z_n) = R_n^p$. Recall $\Lambda_0(x) = -\log \bar{F}_0(x)$ and $\rho_n(R_n^p, a_n) = \big(\frac{\bar{F}_0(R_n^p)}{\bar{F}_0(R_n^p - a_n)}\big)^{\alpha_n}$. Now using the Mean Value Theorem we have,

$$\begin{aligned}
-\log \rho_n(R_n^p, a_n) &= \alpha_n \Lambda_0(g_n(Z_n)) - \alpha_n \Lambda_0(g_n(Z_n - \frac{a_n}{B_n})) \\
&= \alpha_n \frac{a_n}{B_n} \Lambda_0'(g_n(Z_n^*)) g_n'(Z_n^*) \\
&= \alpha_n \frac{a_n}{B_n} \Lambda_0'(B_n Z_n^* + A_n) B_n \\
&= \alpha_n a_n \Lambda_0'\big(A_n(1 + \frac{B_n}{A_n} Z_n^*)\big) \frac{\psi_0'(\sum_{i=1}^n \beta_i)}{\psi_0'(\sum_{i=1}^n \beta_i)} \qquad (4.3)
\end{aligned}$$

where Z_n^* is a random variable lying between Z_n and $Z_n - \frac{a_n}{B_n}$. By our assumption on ψ_0 (see Resnick [13], page21),

$$\lim_{x\to\infty} \frac{x\psi_0'(x)}{\psi_0(x)} = \nu. \qquad (4.4)$$

Therefore,

$$\begin{aligned}
\lim_{n\to\infty} \frac{B_n}{A_n} &= \lim_{n\to\infty} \frac{(\sqrt{\sum_{i=1}^n \beta_i^2})\psi_0'(\sum_{i=1}^n \beta_i)}{\psi_0(\sum_{i=1}^n \beta_i)} \\
&= \lim_{n\to\infty} \frac{\sqrt{\sum_{i=1}^n \beta_i^2}}{\sum_{i=1}^n \beta_i} \lim_{n\to\infty} \frac{(\sum_{i=1}^n \beta_i)\psi_0'(\sum_{i=1}^n \beta_i)}{\psi_0(\sum_{i=1}^n \beta_i)} = 0 \qquad (4.5)
\end{aligned}$$

since the first limit in (4.5) is 0 and second limit is ν by (4.4). Also, if $\frac{a_n}{B_n}$ converges to 0, then since $Z_n \xrightarrow{\mathcal{D}} N(0,1)$ and since Z_n^* lie between Z_n and $Z_n - \frac{a_n}{B_n}$ a.s., $Z_n^* \xrightarrow{\mathcal{D}} N(0,1)$. This together with (4.5) gives $1 + \frac{B_n}{A_n} Z_n^* = 1 + o_p(1)$. Since by our assumption on ψ_0, Λ_0 is a regularly varying function with index $\frac{1}{\nu}$ (see Resnick [13], page 23) and Λ_0' is monotone, we can use similar argument as in the proof of Lemma 2. Here we set $A_n = u_n$ and $1 + \frac{B_n}{A_n} Z_n^* = y_n$ and replace ψ_0' by Λ_0'. Consequently,

$$\frac{\Lambda_0'\big(A_n(1 + \frac{B_n}{A_n} Z_n^*)\big)}{\Lambda_0'(A_n)} = \frac{\Lambda_0'(u_n y_n)}{\Lambda_0'(u_n)} = 1 + o_p(1). \tag{4.6}$$

Now choose

$$a_n = \frac{a\psi_0'(\sum_{i=1}^n \beta_i)}{\alpha_n}.$$

Observe that for this choice of a_n, $\frac{a_n}{B_n} \to 0$. Putting this value of a_n in (4.3) and using (4.6) we have

$$-\log \rho_n(R_n^p, a_n) = a(1 + o_p(1))\Lambda_0'(A_n)\psi_0'(\sum_{i=1}^n \beta_i) = a(1 + o_p(1)), \tag{4.7}$$

since $\Lambda_0'(A_n)\psi_0'(\sum_{i=1}^n \beta_i) = \Lambda_0'(\psi_0(\sum_{i=1}^n \beta_i))\psi_0'(\sum_{i=1}^n \beta_i) = 1$. Therefore,

$$\lim_{n\to\infty} -\log \rho_n(R_n^p, a_n) = a \text{ in probability.}$$

This proves (i) for **Case 1**.

Case 2. $\sum_{i=1}^\infty \beta_i^2 < \infty$.

We define A_n, B_n, g_n and Z_n as Case 1. The only difference is, in this case $Z_n^* \xrightarrow{\mathcal{D}} Z$, where Z is a random variable which is not standard normal. But clearly the same proof goes through. This completes the proof of (i).

Now to prove (ii). Since $\sum_{i=1}^\infty \beta_i < \infty$ now, $R_n^p \to W$ a.s., where W is a continuous random variable (see Lemma 1). Using the Mean Value Theorem, we have

$$-\log \rho_n(R_n^p, a_n) = \alpha_n \Lambda_0(R_n^p) - \alpha_n \Lambda_0(R_n^p - a_n) = \alpha_n a_n \Lambda_0'(R_n^{p*}), \tag{4.8}$$

where R_n^{p*} is a random variable lying between $R_n^p - a_n$ and R_n^p. Choose $a_n = \frac{a}{\alpha_n}$. Since $a_n = \frac{a}{\alpha_n} \to 0$, $\lim_{n\to\infty} R_n^p - a_n = W$ a.s. which implies $R_n^{p*} \to W$ a.s. Since by our assumption Λ_0' is monotone, D, the set of discontinuities of Λ_0' is at most countable. So $P(W \in D) = 0$. Therefore $\Lambda_0'(R_n^{p*}) \xrightarrow{\mathcal{D}} \Lambda_0'(W)$. Thus we get from (4.8), $\rho_n(R_n^p, a_n) \xrightarrow{\mathcal{D}} e^{-a\Lambda_0'(W)}$ which completes the proof of (ii). □

Example. (i) If $\alpha_n = n^\delta$, with $0 < \delta < 1$ and $F_0(x) = 1 - e^{-x^{\frac{1}{\nu}}}$ with $\nu > 0$ then $\psi_0(x) = x^\nu$ and all the conditions of Theorem 5 (i) are satisfied.

(ii) If $\alpha_n = n^\delta$ with $\delta > 1$ and $\psi_0(x) = x^\nu$ with $\nu > 0$, then all conditions of (ii) of the theorem is satisfied.

Remark. For Theorem 5 (ii) to hold it is not necessary for ψ_0 to be regularly varying, as is obvious from the proof of (ii). All we need is, $\sum_{i=1}^\infty \beta_i < \infty$ and Λ_0' exists and has at most countably many discontinuities. Therefore if $\alpha_n = n^\delta$, with $\delta > 1$ and $\psi_0(x) = e^x$ then the p.g.f. of $\xi_n^p(a_n)$ converges to the same limit, with the same choice of a_n as in the proof of (ii) of the theorem.

Acknowledgment. The research of Arup Bose has been supported by the J.C.Bose Fellowship, Department of Science and Technology, Govt. of India. His work was done while he was visiting Department of Economics, University of Cincinnati, USA. Sreela Gangopadhyay thanks Department of Science and Technology, Govt. of India for funding this research (scheme no. SR/WOS-A/MS-02/2006) and Indian Statistical Institute, Kolkata for providing research facilities.

References

[1] Arnold, B. C. and Villaseñor, J. A. (2007). On limit laws for sums of Pfeifer records. *Extremes*, **10(1)**: 235-248

[2] Arnold, B. C.; Balakrishnan, N. and Nagaraja, H.N. (1998). *Records*. Wiley, New York.

[3] Balakrishnan, N.; Pakes, A. G. and Stepanov, A. (2005). On the number and sum of near-record observations. *Adv. Appl. Prob.*, **37**: 765-780

[4] Bondesson, L. (1992). *Generalized Gamma Convolution and Related Classes of Distributions and Dencities*. (Lecture Notes Statist. **76**). Springer, New York.

[5] Bose, Arup and Gangopadhyay, Sreela. (To appear). Convergence of linear functions of Pfeifer records. *Extremes*.

[6] Li, Yun. (1999). A note on the number of records near the maximum. *Stat. Prob. Letters*, **43**: 153-158

[7] Li, Yun and Pakes, A. G. (2001). On the number of near-maximum insurance claims. *Insurance Math. Econom.* **28**: 309-318

[8] Nevzorov, V. B. (1986). On the kth record moments and generalizations. *Zapiski Nauchn Sem. LOMI* **153**: 115-121 (in Russian).

[9] Nevzorov, V. B. (2000). *Records: Mathematical Theory*(Transl. Math. Monogr. **194**). American Mathematical Society, Providence. Rhode Island.

[10] Pakes, A. G. (2007). Limit theorems for numbers of near records. *Extremes*, **10**: 207-224.

[11] Pakes, A. G. and Steutel, F. W. (1997). On the number of records near the maximum. *Austral. J. Stat.* **39**: 179-193

[12] Pfeifer, D. (1984). Limit laws for inter-record times from nonhomogeneous record values. *J. Organ. Behav. Stat.* **1**: 69-74.

[13] Resnick, S. I. (1987). *Extreme values, regular variation, and point processes*. Springer-Verlag, New York.

ARUP BOSE
Statistics and Mathematics Unit
Indian Statistical Institute
203 B. T. Road, Kolkata 700108
INDIA.
email: bosearu@gmail.com

SREELA GANGOPADHYAY
Statistics and Mathematics Unit
Indian Statistical Institute
203 B. T. Road, Kolkata 700108
INDIA.
email: gangopadhyay7@gmail.com res9616@isical.ac.in