On Independent Copies of Negative Binomial Distribution

N. Vadiraja\textsuperscript{1,*}, S. Nagesh\textsuperscript{2}

Assistant Professor – Statistics, Department of Community Medicine, Mysore Medical College and Research Institute, Mysuru, Karnataka, India \textsuperscript{1}, Assistant Professor, DOS in Statistics, Karnataka university, Dharvada, Karnataka, India \textsuperscript{2}

*correspondence author

- Short title: Limit Points of vector sequences
- AMS classification: Primary 60F15; secondary 60G17

Abstract

This paper finds a limiting region for the number of subjects required and hence number of failed in screening test in multi-centric clinical trials. This situation follows a properly normalized independent vector sequences comprising of moving maxima \(Y_{m,n}(k(n))\) for \(m \geq 1\) multi centric set up in clinical trials, where \(1 \leq k(n) \leq n\). Results are given for bi-centric and multi-centric situations, under certain conditions on \(k(n)\) for \(p=\infty\) case.

Keywords: Almost sure limit set; Clinical trial; Moving maxima; Independent copies; Vector sequence

1. Introduction

The number of failure subjects to make it to the clinical trial till the fixed number of inclusions is reached on the \(j^{th}\) day follows Negative Binomial Distribution (NBD). Consider \(k(n)\) number of days, and hence the maximum failures moves and hence screening subjects, as and when \(k(n)\) changes. This is exactly the moving maxima, which is due to Rothmann and Russo (1991). The scheme of finding number of failures for fixed number of inclusion of subjects on each day is adopted at each center. The moving maxima of number of screening subjects, that include failures and test passed subjects, on \(j^{th}\) day at each centre constitute vector sequence of independent components of \(i^{th}\) centre moving maxima. Thus to provide optimum resources at the centre to minimize the cost involved, doctors/company might be interested to know the strong limiting regions in which the moving maxima of number of screening subjects of multi-centre lie.

Let \(r\) be the number of subjects passes the screening tests i.e. the sample size required for the multi-centric trial. Let \(\{X_n, n \geq 1\}\) be a sequence of number of screening subjects required to meet \(r\) and is independent identically distributed random variables (i.i.d.r.v) with common probability mass function

\[P(X=k)=p(k)\equiv (k-1)c_0(1-a+r)^{k-r}, k=r, r+1,\ldots, 0<a<1.\]

Define, moving maxima \(Y_{i(n)}=\max(\{X_{n+1}, X_{n+2}, \ldots, X_{n+k(n)}\})\) where \(k(n)\) is a sequence of positive integers, \(2 \leq k(n) \leq n\), for \(i^{th}\) multi-centre, \(i=1,2,3,\ldots\)

Condition on \(k(n)\) in Hebbar and Vadiraja(1997) is used in this paper.

\(k(n)\) is non-decreasing \hspace{1cm} (1.1)

\[\text{Sup} [ k(n+1) – k(n) ] \leq \mu \text{ (finite)} \hspace{1cm} (2.1)\]

and

\[K(n) = [n/(\log n)^{\alpha(n)}] \text{ where } t(n) \rightarrow p, 0 \leq p \leq \infty \text{ as } n \rightarrow \infty \hspace{1cm} (3.1)\]
Let \( b_n = -\log n/\log(1-a) \) be a real sequence and that \( \log(n)/\log n \to \Delta \) as \( n \to \infty \) where \( \Delta \in [0,1) \).

In view of this, it is planned to get the strong limiting regions for vector sequences of independent copies of moving maxima for Negative Binomial Distribution (NBD). For \( p \in [0,\infty) \), Vadiraja and Nagesha (2016) showed the limit sets. Here it is proceeded with \( p = \infty \).

Let \( Y_{k(n)} \) be the forward moving maxima. Then the above results hold good.

2. Proofs.

The proof of Theorem 1 is built up through the following lemmas.

Let for every \( a_0 < a < a_1 \), there exists a constant \( c > 0 \) such that
\[
ca_i < p_i < ca_i^1 \quad \text{for all } i. \quad (1.2)
\]

**Lemma 1.** Let \( S_1 = \{(x,y): \Delta \leq x, y \leq 1+\Delta , x+y \leq 1+\Delta \} \) where \( \Delta \in [0,1) \).

**Theorem 2.** The almost sure limit set of the vector sequence
\[
\{ Y_{k(1)} / b_n, Y_{k(2)} / b_n, \ldots, Y_{k(n)} / b_n \} \quad n \geq 1, m > 0
\]
coincides with the region \( S_1 = \{(x,y,z): \Delta \leq x,y,z \leq 1+\Delta , x+y+z \leq 1+\Delta \} \) where \( \Delta \in [0,1) \).

**Remark:** Let \( Y_{k(n)} = \max(X_{n+1}, X_{n+2}, \ldots, X_{n+k(n)}) \) be the forward moving maxima. Then the above results hold good.

**Proof:**
Note that \( \sum_{i=x}^{b(n)} p_i \to 0 \) for \( i \) large
\[
k(l_i) \cdot \sum_{i=x}^{b(n)} p_i \to 0 \quad \text{for } i \text{ large and } x>\Delta
\]
\[
k(l_i) \cdot \sum_{i=y}^{b(n)} p_i \to 0 \quad \text{for } i \text{ large and } y>\Delta
\]
\[
(2.2) \text{ is achieved as follows. For all } i \text{ large},
\]
\[
P(Y_{k(l_i)} > (x+\epsilon) b_n, Y_{k(l_i)}^{2} > y b_n) = \left\{ \sum_{i=(x+\epsilon) b(n)}^{b(n)} p_i^1 \right\}^{k(l_i)} \left\{ \sum_{i=y b(n)}^{b(n)} p_i^2 \right\}^{k(l_i)}
\]
\[
= \text{const.} \left\{ 1-\left\{ 1-(1-a_0)^{(x+\epsilon) b(n)} \right\}^{k(l_i)} \right\} \left\{ 1-\left\{ 1-(1-a_0)^{y b(n)} \right\} \right\}
\]
in view of (1.2),
\[
= \text{const.} \cdot k^2(l_i) \left( 1-a_0 \right)^{(x+\epsilon) b(n)} \left( 1-a_0 \right)^{y b(n)} \]
\[=\text{const. exp} \left\{ \log \left( k^2(l_1) \left( 1-a_0 \right)^{x+y+c} h \right) \right\} \]

\[=\text{const. exp} \left\{ 2\log k(l_1) + (x+y+c)b_h \ast \log(1-a_0) \right\} \]

\[=\text{const. exp} \left\{ 2\log k(l_1) - (x+y+c) \log l_i \log(1-a_0) / \log(1-a_0) \right\} \]

\[=\text{const. exp} \left\{ -\log l_i \left( \log(1-a_0) / \log(1-a_0) \right) \right\} \]

\[=\text{const. exp} \left\{ -\log l_i \left( \log(1-a_0) / \log(1-a_0) \right) \right\} \]

\[=\text{const. exp} \left\{ -\log l_i \left( \log(1-a_0) / \log(1-a_0) \right) \right\} \]

\[=\text{const. exp} \left\{ -\log l_i \left( \log(1-a_0) / \log(1-a_0) \right) \right\} \]


For every \( \epsilon > 0 \) and \( i \) large,

\[\theta(x+y-2\Delta + \epsilon) = 1 + \delta_1 \quad \text{(8.2)}\]

where \( \delta_1 = [\theta \epsilon/2] > 0 \).

In view of (8.2) and (7.2),

\[\Sigma P(Y_{1(ki)}^1 > b_h(x+c), Y_{2(ki)}^2 > yb_h) < \infty\]

Through Borel-Cantelli (B-C) lemma, (2.2) follows. The proof of (3.2) is similar. The proof of (4.2) is established through B-C lemma.

Note that

\[P(Y_{1(ki)}^1 > b_h(x), Y_{2(ki)}^2 > b_h(y)) = \text{const. } k^2(l_1) \left( 1-a_1 \right)^{x+y+c} h \ast (1-a_1)^{y} h \quad \text{(9.2)}\]

in view of (1.2). For every \( \epsilon > 0 \) and \( i \) large, we have

\[\text{RHS(9.2)} \geq \text{const. } l_i^{-1(x+y-2\Delta + \epsilon)} \]

\[\text{= const. } i^{-(1-\delta_2)}\]

where \( \delta_2 = 30\epsilon/2 > 0 \) for every \( \epsilon > 0 \) and \( i \) large. To prove (4.2), it is sufficient to show \( Y_{1(ki)}^1 \)'s are independent for all \( i \) large, \( i = 1,2,... \). Observe that,

\[l_i - k(l_i) + 1 - l_{i+1} = l_i \left[ 1-k(l_i)/l_i \right. + 1/l_i - l_{i+1}/l_i \right\} \quad \text{(10.2)}\]

\[\text{RHS (10.2)} \rightarrow \infty \text{ for large } i \text{ and for } \theta > 0 \text{ i.e } x+y \leq 2\Delta \]

as \( l_i/l_i \rightarrow 1 \left( k(l_i) \ast l_i/l_i, l_i \right) \rightarrow 0 \), \( 1 - l_{i+1}/l_i \sim \text{hi}^{(0,1)} \)

Hence, whenever \( \theta > 1 \), i.e. \( (x+y-2\Delta) < 1 \)

\[\text{R.H.S(10.2)} \sim l_i \text{ as } i \rightarrow \infty \text{. Further for } (1-\Delta)^{-1} < \theta < 1, \text{ the expression inside the square bracket of (10.2) is } \sim \text{hi}^{(0c)} \]

\[\text{1) as } i \rightarrow \infty, \text{ since } i^{(1-\theta)} \ast k(l_i)/l_i \rightarrow 0 \]

Thus, for \( \theta > (1-\Delta)^{-1} \), i.e. for \( x+y < 1 + \Delta \),
R.H.S (10.2) tends to \( \infty \) as \( i \to \infty \).

Thus, the events under consideration are independent, for all \( i \) large.

**Lemma 2.2.** For all \( x \geq \Delta \), \( y \geq \Delta \) with \( x+y > 1+ \Delta \) and for every \( \varepsilon > 0 \),

\[
P(Y_{1}^{k(n)} > b_{d}(x+c), Y_{2}^{k(n)} > b_{d}(y+c) \ i.o.) = 0
\]

(11.2)

**Proof:**

\[
P(Y_{1}^{k(n)} > b_{d}(x+c), Y_{2}^{k(n)} > b_{d}(y+c)) = \{\sum_{i=(x+c) b_{d} \to \infty}^{k(n)} p_{i}^{1}\} \{\sum_{i=(y+c) b_{d} \to \infty}^{k(n)} p_{i}^{2}\} \]

\[
= \text{const.} \{1-(1-(1-a_{0})^{(x+c) b_{d}})^{k(n)}\} \{1-(1-(1-a_{0})^{(y+c) b_{d}})^{k(n)}\}
\]

in view of (1.2).

\[
= \text{const.} k^{2}(\delta) (1-a_{0})^{(x+c) b_{d}} (1-a_{0})^{(y+c) b_{d}}
\]

\[
= \text{const.} \exp^{-k(n) \sum_{i=(x+c) b_{d} \to \infty}^{k(n)} p_{i}^{1}} \]

in view of (10.2) as \( n \to \infty \). From (1.2) for all large \( i \)

For every \( \varepsilon > 0 \), \( x+y > 1+\Delta \) and for \( n \) large,

\[
\theta(x+y-2\Delta+2\varepsilon) > 1+\delta, \ \delta_{3}=30\varepsilon/2 > 0.
\]

(13.2)

An appeal to (13.2), (12.2) and B.C lemma, the lemma is proved.

**Lemma 3.2.** For every \( \varepsilon > 0 \) and \( x_{0} = \Delta \)

\[
P(Y_{1}^{k(n)} < (x_{0}-\varepsilon) b_{d} \ i.o.) = 0 \]

(14.2)

\[
P(Y_{2}^{k(n)} < (x_{0}-\varepsilon) b_{d} \ i.o.) = 0 \]

(15.2)

**Proof:**

(14.2) is established by showing the following and (16.2) follows on similar lines.

\[
P(Y_{1}^{k(n)} \leq (x_{0}-\varepsilon) b_{d} \ i.o.) = 0 \]

(16.2)

and

\[
P(Y_{1}^{k(n)} \leq (x_{0}+\varepsilon) b_{d} \ i.o.) = 1 \]

(17.2)

Note that by the independence,

\[
P(Y_{1}^{k(n)} \leq (x_{0}-\varepsilon) b_{d}) = \{\sum_{i=0 \to (x_{0}-\varepsilon) b_{d}}^{k(n)} p_{i}^{1}\}^{k(n)}
\]

\[
= \exp^{-k(n) \sum_{i=(x_{0}-\varepsilon) b_{d} \to \infty}^{k(n)} p_{i}^{1}} \]

(18.2)

in view of (5.2) as \( n \to \infty \). From (1.2) for all large \( i \)

\[
\text{R.H.S}(18.2) \leq \exp^{-\text{const.} k(n) \sum_{i=(x_{0}-\varepsilon) b_{d} \to \infty}^{k(n)} (1-a_{0})^{i}}
\]

\[
\leq \exp^{-\text{const.} k(n) (1-a_{0})^{(x_{0}+\varepsilon) b_{d} \to \infty}}
\]

\[
\leq \text{const.} k(n) (1-a_{0})^{(x_{0}-\varepsilon) b_{d} \to \infty} \]

(19.2)

M being a positive integer. Fix M large so that
\( \text{RHS}(19.2) \leq n^{-\frac{M_e}{2}} \leq n^{(1+\delta_4)}, \delta_4 > 0. \)  \hspace{1cm} (20.2)

Hence an appeal to B-C lemma, (16.2) is shown.

Next, we show (17.2). Consider,

\[ P(Y_{k(0)}^{1} \leq (x_0+\varepsilon)b_n) \geq \lim_{N \to \infty} \left\{ \sum_{i=0}^{(x_0+\varepsilon)bN} p_1 \right\}^{k(N)} \]

\[ = \lim_{N \to \infty} \exp \left\{ -k(N) \sum_{i=(x_0+\varepsilon)bN}^{\infty} p_1 \right\}^{k(N)} \]

\[ \geq \lim_{N \to \infty} \exp \left\{ -\text{const.} k(N) (1-a_1)^{(x_0+\varepsilon)bN} \right\} \]

\[ \geq \lim_{N \to \infty} \exp \left\{ -\text{const.} N^{\varepsilon_2} \right\} \]

similar to that at (20.2).

\[ = 1 \]

Hence, (17.2). Thus, the proof of (14.2) is complete. Similarly (15.2) can be shown. Hence the proof of lemma.

**Proof of Theorem 1:** S is a required limit set by lemmas 2.2 and 3.2. It is concluded with the fact that the limit set is necessarily closed from the lemma 1.2. This completes the proof of theorem 1.

**References**