EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 15, No. 2, 2022, 511-527
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Extensions of two classical Poisson limit laws to non-stationary independent sequences

Aladji Babacar Niang[1], Harouna Sangaré[2], Tchilabalo Abozou Kpanzou[3,*], Gane Samb Lo[4], Nafy Ngom[1]

[1] *LERSTAD, Gaston Berger University, Saint-Louis, Sénégal.*
Imhotep Mathematical Center (IMC),
[2] *DER MI, FST, Université des Sciences, des Techniques et des Technologies de Bamako (USTT-B), Mali. Affiliated to LERSTAD, Gaston Berger University, Saint-Louis, Sénégal*
[3] *University of Kara, Kara, Togo. Affiliated to LERSTAD, Gaston Berger University, Saint-Louis, Sénégal*
[4] *LERSTAD, Gaston Berger University, Saint-Louis, Sénégal (main affiliation)*
Department of Pure and Applied Mathematics, African university of Science and Technology, Abuja, Nigeria
LSTA, Pierre and Marie Curie University, Paris VI, France (associated researcher)

Abstract. In earlier stages in the introduction to asymptotic methods in probability theory, the weak convergence of sequences $(X_n)_{n\geq 1}$ of binomial random variables (*rv*'s) to a Poisson law is classical and easy to prove. A version of such a result concerning sequences $(Y_n)_{n\geq 1}$ of negative binomial *rv*'s also exists. In both cases, X_n and $Y_n - n$ are by-row sums $S_n[X]$ and $S_n[Y]$ of arrays of Bernoulli *rv*'s and corrected geometric *rv*'s respectively. When considered in the general frame of asymptotic theorems of by-row sums of *rv*'s of arrays, these two simple results in the independent and identically distributed scheme can be generalized to non-stationary data and beyond to non-stationary and dependent data. Further generalizations give interesting results that would not be found by direct methods. In this paper, we focus on generalizations to the non-stationary independent data in the frame of the central limit theorem for independent random variables.

2020 Mathematics Subject Classifications: 60B10, 60F05, 60G70, 62G30

Key Words and Phrases: Summands of independent and square integrable random variables; weak convergence of arrays; Poisson limits, binomial and negative binomial laws; Bernoulli and corrected Geometric laws; non-stationary.

---

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v15i2.4323

Email addresses: niang.aladji-babacar@ugb.edu.sn (AB Niang), harounasangare@fst-usttb-edu.ml (H Sangaré), t.kpanzou@univkara.net (TA Kpanzou), gane-samb.lo@ugb.edu.sn (GS Lo), fany.ngom@ugb.edu.sn (N Ngom)

1. Introduction

1.1. Preliminaries

The approximation of a sequence of binomial probability laws $(\mathcal{B}(n,p_n))_{n\geq 1}$ associated to a sequence of *r.v*'s $(Z_n)_{n\geq 1}$ [such that the sequence of probabilities $(p_n)_{n\geq 1}$ converges to zero and $np_n \to \lambda > 0$ as $n \to +\infty$] to a Poisson law $\mathcal{P}(\lambda)$ is a classical and easy-to-prove result in probability theory. This approximation has very important applications in real-life problems, especially in lack of powerful computers. In this simple case, each Z_n is a sum of n independent and identically distributed (*iid*) Bernoulli $\mathcal{B}(p_n)$-random variables $\{(X_{j,n})_{1\leq j\leq n},\ n\geq 1\}$, *i.e.*, $Z_n = X_{1,n} + X_{2,n} + \cdots + X_{n,n}$. When we depart from the identical distributivity assumption, the problem may get more and rapidly complex, even if the independence assumption is still required. The situation becomes more interesting if the random variables $X_{j,n}$ are non-stationary and independent. For the definition and more details on the Binomial and Poisson laws, see [4].

In this paper, we aim at giving non trivial generalizations of such results in the frame of the central limit theorem for independent random variables.

Also, there is a *negative* version of the described result. Indeed, if we call a binomial law as a positive binomial law $\mathcal{PB}(n,p)$, $n \geq 1$, $0 < p < 1$ in opposition to a negative binomial law $\mathcal{NB}(n,p)$, we have the following two results concerning positive binomial and negative binomial laws respectively.

Let us make this precision for once: throughout this paper, all limits are meant as $n \to +\infty$ unless the contrary is specified.

Proposition 1. *Let $(X_n)_{n\geq 1}$ be a sequence of random variables in some probability space $(\Omega,\mathcal{A},\mathbb{P})$ such that:*

1) $\forall n \geq 1,\ X_n \sim \mathcal{B}(n,p_n),\ n\geq 1$;

2) $p_n \to 0$ and $np_n \to \lambda \in \mathbb{R}_+ \setminus \{0\}$ as $n \to +\infty$.

Then

$$X_n \rightsquigarrow \mathcal{P}(\lambda).$$

Next, we have:

Proposition 2. *Let $(X_n)_{n\geq 1}$ be a sequence of random variables in some probability space $(\Omega,\mathcal{A},\mathbb{P})$ such that:*

1) $\forall n \geq 1,\ X_n \sim \mathcal{NB}(n, p_n),\ n \geq 1$*;*

2) $(1 - p_n) \to 0$ *and* $n(1 - p_n) \to \lambda \in \mathbb{R}_+ \setminus \{0\}$ *as* $n \to +\infty$*.*

Then

$$X_n - n \rightsquigarrow \mathcal{P}(\lambda).$$

Remark. These two results are proved in [5]. Although the proofs are direct, we do not encounter the second in some classical books as [1, 2], [3], [6]. For that reason, we give it below.

Proof of Proposition 2.

Let us use the convergence of characteristic functions. Let X_n be a sequence of $\mathcal{NB}(n, p_n)$-random variables and X be a $\mathcal{P}(\lambda)$ random variable. We have

$$S_n = X_n - n = Z_1 + \cdots + Z_n,$$

where $Z,\ Z_1, \cdots,\ Z_n$ are independent random variables such that each $Z_i + 1$ follows a geometric law of parameter p_n, that is

$$\Phi_Z(t) = \frac{p_n}{1 - q_n e^{it}}, \quad t \in \mathbb{R}.$$

So, for $t \in \mathbb{R}$ fixed,

$$\Phi_{S_n}(t) = \exp\Bigg(n \left(\log p_n - \log\left(1 - q_n e^{it}\right)\right)\Bigg).$$

We have, as $n \to +\infty$,

$$n \log p_n = n \log(1 - q_n) = -nq_n + o(nq_n)$$

and

$$-n \log\left(1 - q_n e^{it}\right) = nq_n e^{it} + o(nq_n).$$

Hence we get for any $t \in \mathbb{R}$,

$$\Phi_{S_n}(t) = \exp\Bigg(nq_n(e^{it} - 1) + o(nq_n)\Bigg) \to e^{\lambda(e^{it}-1)} = \Phi_{\mathcal{P}(\lambda)}(t). \ \Box$$

Our aim here is to provide non-trivial generalizations of such simple results to non-stationary and independent data. The used methods will later allow further generalizations even with dependent data.

Let us prepare generalizations by transforming both results as sums of random variables.

1.2. The Central limit theorem frame

It is known that a binomial random variable $X_n \sim \mathcal{B}(n, p_n)$ has the same law as a sum of n *iid* Bernoulli distributed random variables:

$$X_n =^d X_{1,n} + \cdots + X_{n,n},$$

where $X_{1,n}, \cdots, X_{n,n}$ are independent and follow all the $\mathcal{B}(p_n)$-law. Also $X_n \sim \mathcal{NB}(n, p_n)$ has the same law as a sum of n *iid r.v*'s:

$$X_n =^d X^*_{1,n} + \cdots + X^*_{n,n},$$

where $X^*_{1,n}, \cdots, X^*_{n,n}$ are independent and each $X^*_{j,n}$ follows the geometric law $\mathcal{G}(p_n)$. In the second, we rather use

$$X_n - n = \sum_{i=1}^{n} \left(X^*_{i,n} - 1\right) =: \sum_{i=1}^{n} X_{i,n},$$

where the $X_{i,n}$'s are independent and each $X_{i,n}$ follows the law $\mathcal{G}^*(p_n) = \mathcal{G}(p_n) - 1$, and such a law is called a corrected geometric law, for convenience. In both cases, we have to study an array

$$X \equiv \biggl\{ \{X_{k,n},\ 1 \leq k \leq k(n)\},\ n \geq 1 \biggr\},$$

of random variables defined in the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that here:

1) $\forall n \geq 1,\ k(n) = n$;

2) $\forall n \geq 1$, the variables $X_{1,n}, \cdots, X_{k(n),n}$ are independent;

3) The sequence $X_{1,n}, \cdots, X_{k(n),n}$ is stationary for $n \geq 1$;

4) $\forall k \in [1, k(n)]$, $X_{k,n} \sim \mathcal{B}(p_n)$ or $X_{k,n} \sim \mathcal{G}(p_n) - 1$.

We see that we are in the *CLT* frame and each of points (2) and (3) can be changed to lead to generalizations. Since, we want to generalize the limiting binomial laws and negative binomial laws, we keep the same hypotheses on the marginal laws

$$\forall k \in [1, k(n)],\ X_{k,n} \sim \mathcal{B}(p_{k,n}) \quad or \quad X_{k,n} \sim \mathcal{G}(p_{k,n}) - 1$$

and try to answer to the questions (Q1) and (Q2) below:

(Q1) Given an array X of random variables with $k(n) \to +\infty$ such that the elements of each row are independent and $\mathcal{B}(p_{k,n})$-$r.v$'s, do we still have

$$S_n[X] = \sum_{k=1}^{k(n)} X_{k,n} \rightsquigarrow \mathcal{P}(\lambda), \tag{1}$$

when some of the assumptions (1), (2) and (3) [but mainly (2) and (3)] are violated, and under what sufficient conditions this should hold?

(Q2) Given an array X of random variables with $k(n) \to +\infty$ such that the elements of each row are independent and $\mathcal{G}^*(p_{k,n})$-$r.v$'s, do we still have (1) when some of the assumptions (1), (2) and (3) [but mainly (2) and (3)] are violated, and under what sufficient conditions this should hold?

Although direct handlings of these questions might be possible, we think that a general and extensible solution resides in the *CLT* frame, since it will prepare further generalizations for dependent data.

Therefore, we organize the paper as follows. In Section 2, we recall the frame of the *CLT* problem as stated in [6]. In Section 3, we state and prove the results. We conclude the paper by conclusive remarks in Section 4.

2. Notation and *G-CLT* for summands of independent random variables

Let us consider the array

$$X \equiv \biggl\{ \{X_{k,n},\ 1 \leq k \leq k_n = k(n)\},\ n \geq 1 \biggr\},$$

of square integrable random variables defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We denote $F_{k,n}$ as the cumulative distribution function (*cdf*) of $X_{k,n}$. We also denote by $a_{k,n} = \mathbb{E}(X_{k,n})$ and $\sigma^2_{k,n} = \mathbb{V}ar(X_{k,n})$, $1 \leq k \leq k(n)$, if these expectations or variances exist. We also suppose that

$$k(n) \to +\infty \ as \ n \to +\infty.$$

The central limit theorem problem consists in finding, whenever possible, the weak limit law (in type) of the by-row sums of the array X, *i.e.* the summands:

$$S_n[X] = \sum_{k=1}^{k(n)} X_{k,n}, \ n \geq 1.$$

Historically, the *CLT* was discovered with the convergence of a binomial law (which has the same law as a sum of *iid* Bernoulli random variables) to the standard Gaussian law (due to Laplace, De Moivre, etc., around 1731, see [6] for a review). For a long period, the Gaussian limit was automatically meant in the *CLT* problem. Many authors, among them Lévy, Gnedenko, Kolmogorov, etc., characterized the class of possible limit laws under the *Uniform Asymptotic Negligibility (UAN)* condition, exactly as the class of infinitely decomposable distributions. The longtime association of *CLT*'s with Gaussian limits explains that some authors reserve the vocable *CLT* for Gaussian limits and for other possible limits, they use different vocables. Here we use the vocable of ***G-CLT*** to cover all possible limit laws G beyond the Gaussian law.

Here we suppose that the $X_{k,n}$'s are integrable with finite variances. For an array X, we define some important hypotheses used in the formulation of the *CLT* problem.

(1) The *UAN* condition: for any $\varepsilon > 0$,

$$U(n, \varepsilon, X) = \sup_{1 \leq k \leq k_n} \mathbb{P}(|X_{k,n} - a_{k,n}| \geq \varepsilon) \to 0. \tag{2}$$

(2) The Bounded Variance Hypothesis (*BVH*): there exists a constant $c > 0$, such that

$$\sup_{n \geq 1} MV(n, X) \leq c,$$

where

$$MV(n, X) = \mathbb{V}ar(S_n[X]), \ n \geq 1.$$

(3) The Variance Convergence Hypothesis (*VCH*):

$$MV(n, X) \to c \in]0, +\infty[.$$

According to the state of the art in *CLT*'s theory for centered, square integrable and independent by-row arrays of random variables, the summands weakly converge to a probability law associated to the *cdf* G and to the characteristic function (*cha.f*) ψ_G under the UAN condition and the BVH if and only if the sequence of distribution functions (*df*)

$$K_n(x) = \sum_{k=1}^{k(n)} \int_{-\infty}^{x} y^2 dF_{k,n}(y), \; x \in \mathbb{R}, \; n \geq 1,$$

pre-weakly converges to a *df* K, denoted $K_n \rightsquigarrow_{pre} K$, that is for any continuity point x of K denoted as $[x \in C(K)]$, we have

$$K_n(x) \to K(x),$$

and the *cha.f* $\psi_G(\circ)$ of G is given by $\exp(\psi[K](\circ))$ with

$$\forall u \in \mathbb{R}, \; \psi[K](u) = \int \frac{e^{iux} - 1 - iux}{x^2} \, dK(x).$$

If we have the *VCH*, the convergence criterion is replaced by the weak convergence $K_n \rightsquigarrow K$. Moreover, the limit law G is necessarily an infinitely decomposable law.

In the non centered case, with the same hypotheses above on the random variables of the array, the summands weakly converge to a probability law associated to the *cdf* G^* and to the *cha.f* ψ_{G^*} under the UAN condition and the BVH if and only if

$$\sum_{k=1}^{k(n)} a_{k,n} \to a, \; a \in \mathbb{R}$$

and the sequence of distribution functions (*df*)

$$K_n^*(x) = \sum_{k=1}^{k(n)} \int_{-\infty}^{x} y^2 dF_{k,n}(y + a_{k,n}), \; x \in \mathbb{R}, \; n \geq 1,$$

pre-weakly converges to a *df* K^* and the *cha.f* $\psi_{G^*}(\circ)$ of G^* is given by $\exp(\psi[K^*](\circ))$ with

$$\forall u \in \mathbb{R}, \; \psi[K^*](u) = \int \frac{e^{iux} - 1 - iux}{x^2} \, dK^*(x).$$

If we have the *VCH*, the convergence criterion is replaced by the weak convergence $K_n^* \rightsquigarrow K^*$. Moreover, the limit law G^* is of the form $G^* = G + a$, with G is necessarily a centered and infinitely decomposable law.

By specializing the limit law as a Gaussian law or a Poisson law, which clearly are infinitely decomposable laws, we have the following characterizations.

C1. Under the conditions

$$\Big(\forall n \geq 1,\ \forall 1 \leq k \leq k(n),\quad a_{k,n} = 0\Big)\ \ \textit{and}\ \ \sum_{k=1}^{k(n)} \sigma_{k,n}^2 = 1,$$

the summands $S_n[X]$ of the array X converges to standard Gaussian law and $\max_{1 \leq k \leq k(n)} \sigma_{k,n}^2 \to 0$ if and only if the following Lynderberg-Gaussian condition holds:

$$\forall \varepsilon > 0, \quad L_{n,G}(\varepsilon) = \sum_{k=1}^{k(n)} \int_{(|x| \geq \varepsilon)} x^2 \ dF_{k,n}(x) \to 0. \tag{3}$$

C2. Under the conditions

$$\max_{1 \leq k \leq k(n)} \sigma_{k,n}^2 \to 0 \ \ \textit{and}\ \ \sum_{k=1}^{k(n)} \sigma_{k,n}^2 \to \lambda,\ \lambda > 0,$$

the summands $S_n[X]$ of the array X converges to a translated Poisson law $\mathcal{P}(a,\lambda) \equiv a + \mathcal{P}(\lambda),\ \ a \in \mathbb{R}$, if and only if

$$\sum_{k=1}^{k(n)} a_{k,n} \to a + \lambda$$

and the following Lynderberg Poisson-type condition holds:

$$\forall \varepsilon > 0, \quad L_{n,P}(\varepsilon) = \sum_{k=1}^{k(n)} \int_{(|x-1| \geq \varepsilon)} x^2 \ dF_{k,n}(x + a_{k,n}) \to 0. \tag{4}$$

3. Statements of the results

As announced, we focus here on the non-stationary independent scheme.

First, we consider uniform conditions of the convergence of the probabilities $p_{k,n}$ (in the Bernoulli case) and $q_{k,n}$ (in the corrected geometric case) to zero to unveil refined versions of the extensions. Later, we will provide more general conditions.

Theorem 1. *Let*

$$X = \left\{ \{X_{k,n},\ 1 \leq k \leq k_n = k(n)\},\ n \geq 1 \right\},$$

be an array of by-row independent Bernoulli random variables, that is:

(1) $\forall n \geq 1$, $\forall 1 \leq k \leq k(n)$, $X_{k,n} \sim \mathcal{B}(p_{k,n})$, *with* $0 < p_{k,n} < 1$ *and:*

(2) $\sup_{1 \leq k \leq k(n)} p_{k,n} \to 0$;

(3) $\sum_{1 \leq k \leq k(n)} p_{k,n} \to \lambda \in]0,\ +\infty[$.

Then we have

$$S_n[X] \rightsquigarrow \mathcal{P}(\lambda).$$

Proof of Theorem 1. Throughout this proof, the notation $\ell_{k,n} = \overline{o}_n(1)$, for k ranging over some set I_n means that the sequence $\ell_{k,n}$ goes to zero as $n \to +\infty$ uniformly in $k \in I_n$. So Assumption (2) means that

$$p_{k,n} = \overline{o}_n(1) \quad and \quad q_{k,n} = 1 - p_{k,n} = 1 + \overline{o}_n(1).$$

We have to check the ***UAN*** condition. By using Chebychev's inequality, we have, for any $\varepsilon > 0$,

$$\begin{aligned} U(n, \varepsilon, X) &= \sup_{1 \leq k \leq k_n} \mathbb{P}(|X_{k,n} - a_{k,n}| \geq \varepsilon) \\ &\leq \varepsilon^{-2} \sup_{1 \leq k \leq k_n} \mathbb{V}ar(X_{k,n}) \\ &= \varepsilon^{-2}\ p_{k,n}\ q_{k,n} \\ &= \varepsilon^{-2}\ \overline{o}_n(1)(1 + \overline{o}_n(1)) \to 0. \end{aligned}$$

The ***VCH*** also holds since

$$MV(n, X) = \sum_{1 \leq k \leq k(n)} \mathbb{V}ar(X_{k,n})$$

$$\begin{aligned} &= \sum_{1\leq k\leq k(n)} p_{k,n}\ q_{k,n} \\ &= (1+\overline{o}_n(1)) \sum_{1\leq k\leq k(n)} p_{k,n} \rightarrow \lambda. \end{aligned}$$

Besides

$$\sum_{1\leq k\leq k(n)} \mathbb{E}(X_{k,n}) = \sum_{1\leq k\leq k(n)} p_{k,n} \rightarrow \lambda.$$

So, we are in the position of applying the conditions of weak convergence to a Poisson law by checking the Poisson Lynderbeg condition (4). We have for any $\varepsilon > 0$

$$L_{n,P}(\varepsilon) =: \sum_{k=1}^{k(n)} L_{n,k,P}(\varepsilon),$$

with

$$\begin{aligned} L_{n,k,P}(\varepsilon) &= \int_{(|x-1|\geq\varepsilon)} x^2\ dF_{k,n}(x+p_{k,n}) \\ &= \int_{(|X_{k,n}-p_{k,n}-1|\geq\varepsilon)} |X_{k,n}-p_{k,n}|^2\ d\mathbb{P}_{k,n} \\ &= p_{k,n}\biggl(1_{(|X_{k,n}-p_{k,n}-1|\geq\varepsilon)}|X_{k,n}-p_{k,n}|^2\biggr)_{(X_{k,n}=1)} \\ &+ (1-p_{k,n})\biggl(1_{(|X_{k,n}-p_{k,n}-1|\geq\varepsilon)}|X_{k,n}-p_{k,n}|^2\biggr)_{(X_{k,n}=0)} \\ &= p_{k,n}1_{(|p_{k,n}|\geq\varepsilon)}\ (1-p_{k,n})^2 + (1-p_{k,n})1_{(|p_{k,n}+1|\geq\varepsilon)}\ p_{k,n}^2 \\ &= p_{k,n}1_{(|\overline{o}_n(1)|\geq\varepsilon)}(1+\overline{o}_n(1))^2 + \overline{o}_n(1)(1+\overline{o}_n(1))1_{(|\overline{o}_n(1)+1|\geq\varepsilon)}\ p_{k,n}. \end{aligned}$$

We only need to get (4) for $0 < \varepsilon < \varepsilon_0$, for a fixed $\varepsilon_0 > 0$. Let us fix $\varepsilon_0 = 1/2$. So, for n large enough,

$$1_{(|\overline{o}_n(1)|\geq\varepsilon)} = 0 \ \ and \ \ 1_{(|\overline{o}_n(1)+1|\geq\varepsilon)} = 1$$

and hence

$$\sum_{k=1}^{k(n)} L_{n,k,P}(\varepsilon) = \overline{o}_n(1)(1+\overline{o}_n(1)) \sum_{k=1}^{k(n)} p_{k,n}$$

$$= \overline{o}_n(1)(1+\overline{o}_n(1))(\lambda+o(1)) \to 0.$$

The proof is complete. ■

Theorem 2. *Let*

$$X = \biggl\{ \{X_{k,n},\ 1 \leq k \leq k_n = k(n)\},\ n \geq 1 \biggr\},$$

be an array of by-row-independent corrected geometric random variables, that is:

(1) $\forall n \geq 1,\ \forall 1 \leq k \leq k(n),\ X_{k,n} \sim \mathcal{G}^*(p_{k,n}),$ *with* $0 < p_{k,n} = 1 - q_{k,n} < 1$ *and:*

(2) $\sup_{1 \leq k \leq k(n)} q_{k,n} \to 0$*;*

(3) $\sum_{1 \leq k \leq k(n)} q_{k,n} \to \lambda \in]0,\ +\infty[$.

Then we have

$$S_n[X] \rightsquigarrow \mathcal{P}(\lambda).$$

Proof of Theorem 2. Assumption (2) of the theorem means that

$$q_{k,n} = \overline{o}_n(1),\ p_{k,n} = 1 + \overline{o}_n(1) \ \ and \ \ (1/p_{k,n})^i = 1 + \overline{o}_n(1),\ i = 1, 2, 3.$$

We have to check the ***UAN*** condition. By using Chebychev's inequality, we have, for any $\varepsilon > 0$,

$$\begin{aligned} U(n,\varepsilon,X) &= \sup_{1 \leq k \leq k_n} \mathbb{P}(|X_{k,n} - a_{k,n}| \geq \varepsilon) \\ &\leq \varepsilon^{-2} \sup_{1 \leq k \leq k_n} \mathbb{V}ar(X_{k,n}) \\ &= \varepsilon^{-2} \sup_{1 \leq k \leq k_n} \frac{q_{k,n}}{p_{k,n}^2} \\ &= \varepsilon^{-2} \overline{o}_n(1)(1+\overline{o}_n(1)) \to 0. \end{aligned}$$

The ***VCH*** also holds since

$$MV(n,X) = \sum_{1 \leq k \leq k(n)} \mathbb{V}ar(X_{k,n})$$

$$
\begin{aligned}
&= \sum_{1\leq k\leq k(n)} \frac{q_{k,n}}{p_{k,n}^2}\\
&= (1+\overline{o}_n(1)) \sum_{1\leq k\leq k(n)} q_{k,n} \to \lambda.
\end{aligned}
$$

Besides

$$
\sum_{1\leq k\leq k(n)} \mathbb{E}(X_{k,n}) = \sum_{1\leq k\leq k(n)} \frac{q_{k,n}}{p_{k,n}} = (1+\overline{o}_n(1)) \sum_{1\leq k\leq k(n)} q_{k,n} \to \lambda.
$$

Here again, we are in the position of applying the conditions of weak convergence to a Poisson law by checking the Poisson Lynderbeg condition (4). We have for any $0 < \varepsilon < 1/2$

$$
L_{n,P}(\varepsilon) =: \sum_{k=1}^{k(n)} L_{n,k,P}(\varepsilon),
$$

with

$$
\begin{aligned}
L_{n,k,P}(\varepsilon) &= \int_{(|x-1|\geq\varepsilon)} x^2 \ dF_{k,n}(x+a_{k,n})\\
&= \int_{\left(\left|X_{k,n}-\frac{q_{k,n}}{p_{k,n}}-1\right|\geq\varepsilon\right)} \left|X_{k,n}-\frac{q_{k,n}}{p_{k,n}}\right|^2 \ d\mathbb{P}_{X_{k,n}}\\
&= \sum_{j=0}^{+\infty} p_{k,n}q_{k,n}^j \left(1_{\left(\left|X_{k,n}-\frac{q_{k,n}}{p_{k,n}}-1\right|\geq\varepsilon\right)} \left|X_{k,n}-\frac{q_{k,n}}{p_{k,n}}\right|^2\right)_{(X_{k,n}=j)}\\
&= p_{k,n}1_{\left(\left|\frac{q_{k,n}}{p_{k,n}}+1\right|\geq\varepsilon\right)} \left(\frac{q_{k,n}}{p_{k,n}}\right)^2 \quad (for\ j=0)\\
&+ p_{k,n}q_{k,n}1_{\left(\left|\frac{q_{k,n}}{p_{k,n}}\right|\geq\varepsilon\right)} \left(1-\frac{q_{k,n}}{p_{k,n}}\right)^2 \quad (for\ j=1)\\
&+ \sum_{j=2}^{+\infty} p_{k,n}q_{k,n}^j 1_{\left(\left|j-1-\frac{q_{k,n}}{p_{k,n}}\right|\geq\varepsilon\right)} \left(j-\frac{q_{k,n}}{p_{k,n}}\right)^2 \quad (for\ j\geq 2).
\end{aligned}
$$

By the same remarks used in the precedent proof, we have for n large enough,

$$
1_{\left(\left|\frac{q_{k,n}}{p_{k,n}}+1\right|\geq\varepsilon\right)} = 1_{(|\overline{o}_n(1)(1+\overline{o}_n(1))+1|\geq\varepsilon)} = 1
$$

and next

$$p_{k,n} 1_{\left(\left|\frac{q_{k,n}}{p_{k,n}}+1\right| \geq \varepsilon\right)} \left|\frac{q_{k,n}}{p_{k,n}}\right|^2 = (1+\overline{o}_n(1))q_{k,n}^2 = \overline{o}_n(1)(1+\overline{o}_n(1))\ q_{k,n}. \quad (L1)$$

Also, we have for n large enough,

$$1_{\left(\left|\frac{q_{k,n}}{p_{k,n}}\right| \geq \varepsilon\right)} = 1_{(|\overline{o}_n(1)(1+\overline{o}_n(1))| \geq \varepsilon)} = 0$$

and next

$$p_{k,n} q_{k,n} 1_{\left(\left|\frac{q_{k,n}}{p_{k,n}}\right| \geq \varepsilon\right)} \left(1 - \frac{q_{k,n}}{p_{k,n}}\right)^2 = 0. \quad (L2)$$

Now, for n large enough and for any $j \geq 2$,

$$1_{(|j-1+\overline{o}_n(1)(1+\overline{o}_n(1))| \geq \varepsilon)} = 1$$

and thus,

$$\begin{aligned}
A_{j,k,n} &:= p_{k,n} q_{k,n}^j 1_{\left(\left|j-1-\frac{q_{k,n}}{p_{k,n}}\right| \geq \varepsilon\right)} \left(j - \frac{q_{k,n}}{p_{k,n}}\right)^2 \\
&= \Big((1+\overline{o}_n(1))q_{k,n}\Big)\ q_{k,n}^{j-1} 1_{\left(\left|j-1+\overline{o}_n(1)(1+\overline{o}_n(1))\right| \geq \varepsilon\right)} \Big(j + \overline{o}_n(1)(1+\overline{o}_n(1))\Big)^2 \\
&= \Big((1+\overline{o}_n(1))q_{k,n}\Big)\ q_{k,n}^{j-1} (j + \overline{o}_n(1)(1+\overline{o}_n(1)))^2 \\
&\leq \Big((1+\overline{o}_n(1))q_{k,n}\Big)\ q_{k,n}^{j-1} (j+1)^2 \\
&\leq 2\Big((1+\overline{o}_n(1))q_{k,n}\Big)\ q_{k,n}^{j-1} (j^2+1),
\end{aligned}$$

where we apply the C_2-inequality in the last line. So we have

$$\sum_{j \geq 2} A_{j,k,n} \ \leq\ 2(1+\overline{o}_n(1))\ q_{k,n}\ B(n,k), \quad (L3)$$

with

$$B(n,k) = \sum_{j\geq 2} q_{k,n}^{j-1} + \sum_{j\geq 2} j^2 q_{k,n}^{j-1} =: B(n,k,1) + B(n,k,2).$$

We have

$$B(n,k,1) = \left(\sum_{j\geq 0} q_{k,n}^{j}\right) - 1 = \frac{q_{k,n}}{p_{k,n}} = \overline{o}_n(1)(1+\overline{o}_n(1)). \quad (L4a)$$

Next

$$\begin{aligned}
B(n,k,2) &= \sum_{j\geq 2} j q_{k,n}^{j-1} + \sum_{j\geq 2} j(j-1) q_{k,n}^{j-1} \\
&= \left(\sum_{j=1}^{+\infty} j q_{k,n}^{j-1} - 1\right) + \left(q_{k,n} \sum_{j=2}^{+\infty} j(j-1) q_{k,n}^{j-2}\right) \\
&= \left(\left\{\sum_{j=0}^{+\infty} q_{k,n}^{j}\right\}' - 1\right) + q_{k,n} \left\{\sum_{j=0}^{+\infty} q_{k,n}^{j}\right\}'' \\
&= \left(\frac{1}{p_{k,n}^2} - 1\right) + \left(q_{k,n} \frac{2}{p_{k,n}^3}\right) \\
&= \frac{1-p_{k,n}^2}{p_{k,n}^2} + \frac{2q_{k,n}}{p_{k,n}^3} \\
&= \frac{p_{k,n}(1-p_{k,n}^2) + 2q_{k,n}}{p_{k,n}^3} \\
&= \frac{p_{k,n} q_{k,n}(1+p_{k,n}) + 2q_{k,n}}{p_{k,n}^3} \\
&= \frac{q_{k,n}\left(p_{k,n}(1+p_{k,n}) + 2\right)}{p_{k,n}^3} \\
&\leq \frac{4q_{k,n}}{p_{k,n}^3} \\
&= 4\overline{o}_n(1)(1+\overline{o}_n(1)). \quad (L4b)
\end{aligned}$$

Hence

$$B(n,k) \leq C\overline{o}_n(1)(1+\overline{o}_n(1)), \quad (L4c)$$

for some $C > 0$ by (L4a) and (L4b).

Finally, by putting together (L1), (L2), (L3) and (L4c), we get

$$L_{n,P}(\varepsilon) \leq \overline{o}_n(1)(1+\overline{o}_n(1))(\lambda+o(1))\left\{1+2C(1+\overline{o}_n(1))\right\} \to 0.$$

This completes the proof. ■

A simple application. Let us give a simple application to a classical example. We suppose that we observe occurrences of landing crashes at some airport (A) over a period $T > 0$. We know that those crashes are usually of very low probabilities. Over n landings, we denote X_n the number of crashes at times $k(n)$. Usually, we suppose that the data (of landing crashes) are observations of *iid* Bernoulli $\mathcal{B}(p_n)$ and then, the approximation $X_n \approx Z \sim \mathcal{P}(\lambda)$, with $\lambda = np_n$ can be used. That formula was systematically used with limited performance of computers. However, with powerful computers, we no-longer need that approximation to compute the related p-values $\mathbb{P}(X_n > t)$ of the statistical tests since we know the explicit form of $S_n[X]$. In the software **R**, the code $1 - pbiniom(t, p_n)$ gives the desired values.

Now suppose we can use the independence hypothesis only and not the stationary distribution. Hence the distribution of X_n is the convolution product of Bernoulli $\mathcal{B}(p_{k,n})$ distributions and its law is not simple. So, the simplest way to compute $\mathbb{P}(X_n > t)$ should be using the approximation $\mathbb{P}(\mathcal{P}(\lambda) > t)$ with $\lambda = p_{1,n} + \cdots + p_{k(n),n}$. So it is better to use the non-stationary scheme since the stationary hypothesis is usually a *working hypothesis*, not confirmed, and p_n is computed as the average number of crashes.

These two theorems actually are still particular cases of two more general results.

Theorem 3. *Let*

$$X = \left\{\{X_{k,n},\ 1 \leq k \leq k_n = k(n)\},\ n \geq 1\right\},$$

be an array of by-row-independent Bernoulli random variables, that is:

(GP1) $\forall n \geq 1$, $\forall 1 \leq k \leq k(n)$, $X_{k,n} \sim \mathcal{B}(p_{k,n})$, *with* $0 < p_{k,n} < 1$ *and:*

(GP2) $\sup_{1\leq k\leq k(n)} p_{k,n}(1-p_{k,n}) \to 0$*;*

(GP3) $\sum_{1\leq k\leq k(n)} p_{k,n}(1-p_{k,n}) \to \lambda \in]0,\ +\infty[$ *and* $\sum_{1\leq k\leq k(n)} p_{k,n} \to \lambda$*;*

(GP4) for $0 < \varepsilon < 1$, $n \geq 1$, $1 \leq k \leq k(n)$, *and for*

$$B(\varepsilon, k, n) = \sum_{j=0}^{1} 1_{\left(\left|j - q_{k,n} - 1\right| \geq \varepsilon\right)} \left(j - p_{k,n}\right)^2 p_{k,n}^{j} q_{k,n}^{1-j},$$

we have

$$B(\varepsilon, n) = \sum_{k=1}^{k(n)} B(\varepsilon, k, n) \to 0.$$

Then we have

$$S_n[X] \rightsquigarrow \mathcal{P}(\lambda).$$

Theorem 4. *Let*

$$X = \left\{ \{X_{k,n},\ 1 \leq k \leq k_n = k(n)\},\ n \geq 1 \right\},$$

be an array of by-row-independent corrected geometric random variables, that is:

(GN1) $\forall n \geq 1$, $\forall 1 \leq k \leq k(n)$, $X_{k,n} \sim \mathcal{G}^*(p_{k,n})$, *with* $0 < p_{k,n} = 1 - q_{k,n} < 1$ *and:*

(GN2) $\sup_{1 \leq k \leq k(n)} (q_{k,n}/p_{k,n}^2) \to 0$;

(GN3) for $h \in \{1, 2\}$, $\sum_{1 \leq k \leq k(n)} (q_{k,n}/p_{k,n}^h) \to \lambda \in]0, +\infty[$.

(GN4) for $0 < \varepsilon < 1$, $n \geq 1$, $0 \leq k \leq k(n)$, *and for*

$$B(\varepsilon, k, n) = \sum_{j=0}^{+\infty} 1_{\left(\left|j - \frac{q_{k,n}}{p_{k,n}} - 1\right| \geq \varepsilon\right)} \left(j - \frac{q_{k,n}}{p_{k,n}}\right)^2 p_{k,n} q_{k,n}^{j},$$

we have

$$B(\varepsilon, n) = \sum_{k=1}^{k(n)} B(\varepsilon, k, n) \to 0.$$

Then we have

$$S_n[X] \rightsquigarrow \mathcal{P}(\lambda).$$

Proofs of Theorems 3 and 4.

In the proofs of Theorems 1 and 2, the ***UAN***, the ***CVH***, the Poisson Lynderberg condition and the convergence of $\mathbb{E}S_n[X]$ are the general conditions for $S_n[X] \rightsquigarrow \mathcal{P}(\lambda)$. In Theorems 3 and 4, uniform convergence simplifies these conditions, making the proofs lighter. □

4. Concluding remarks

The extensions we provide are the first general results. The central limit theorem frame seems to be the appropriate way to get more general extensions. Theorems 1 and 2 can be done by direct methods. However, general forms in Theorems 3 and 4 could hardly be obtained in direct methods. They are products of the *CLT* frame.

References

[1] W Feller. *An introduction to Probability Theory and its Applications. Volume I.* Third Editions. John Wiley & Sons Inc, New-York, 1968.

[2] W Feller. *An introduction to Probability Theory and its Applications. Volume II.* Third Editions. John Wiley & Sons Inc, New-York, 1968.

[3] A Gut. *Probability : A Graduate Course.* Springer Science+Business Media, Inc, Singapore, 2005.

[4] GS Lo. *Mathematical Foundations of Probability Theory.* SPAS Books Series., Calgary, 2018.

[5] GS Lo, TA Kpanzou, M Ngom, and AB Niang. *Weak Convergence (IA): Sequences of random vectors.* SPAS Books Series., Calgary, 2021.

[6] M Loève. *Probability Theory I.* Springer-Verlag., New-York, 1977.