Convex concentration inequalities for nondecreasing processes

Thierry KLEIN
Université de versailles
45 avenue des Etats-Unis
78035 Versailles cedex
klein@math.uvsq.fr

September the 15th 2003

Abstract

In this paper, we prove convex concentration inequalities for discrete and continuous time counting processes. Then we apply these inequalities to prove that the supremum of independent binomial random variables and the supremun of independent Poisson random variables satisfy convex concentration inequalities.

AMS Classification: 60E15, 60F10.
Keywords: Martingales, counting processes, records, convex concentration inequality, Negatively associated variables, 3−ary search trees.

1 Introduction.

In this paper, we will introduce the concept of binomial (resp. Poissonian) convex concentration inequality for discrete (resp. continuous) time process see Definition 1 (resp. Definition 2). We will then give examples of processes who satisfy these inequalities.

This concept was first introduced by Hoeffding in [6]. In this paper, Hoeffding compares $\mathbb{E}\left(\phi(S_n)\right)$ with $\mathbb{E}\left(\phi(S_n^\star)\right)$, when $S_n = \sum_{i=1}^n X_i$ is the sum of independent Bernoulli distributed random variables with parameters p_i and $S_n^\star$ is $B(n,\bar{p})-$distributed with $\bar{p}$ the arithmetic mean of the p_i's.

Proposition 1 *(Hoeffding [6], Shorack-Wellner [12]). Let $b_1, \ldots, b_n$ be independent random variables Bernoulli distributed with parameters p_i and $S_n = b_1 + \cdots + b_n$. Let $\bar{p} = \dfrac{p_1 + \ldots + p_n}{n}$ then for any convex function ϕ we have*

$$\mathbb{E}\Big(\phi(S_n)\Big) \leq \mathbb{E}\Big(\phi\big(B(n, \bar{p})\big)\Big). \tag{1}$$

These inequalities are very useful to derive tail inequalities as pointed by Hoeffding [6], Bretagnolle [3] who gave a functional version of this result, Pinelis in [7] and [8] studies a more general case where the function ϕ is in a general class of functions. Shao in [11] treats the case of Negatively Associated (N.A) random variables and shows how convex concentrations inequalities lead to classical inequalities like Rosenthal maximal inequality or Kolmogorov inequality. In particular, he is able to extend Hoeffding's inequality on the probability bounds for the sum of a random sample without replacement from a finite population. Bentkus in [1] uses convex concentration inequalities to give bounds for probabilities tails of discrete martingales with bounded jumps.

In this paper, we introduce a class of discrete processes which satisfy convex concentration inequalities. Our approach is similar to Shao's approach [11]. Indeed our first result (Theorem 1) states that under some appropriate hypothesis on the discrete process $(Z_n)_{n\in\mathbb{N}}$ (Assumption 1), for any convex function ϕ,

$$\mathbb{E}\left(\phi(Z_n)\right) \leq \mathbb{E}\left(\phi(S_n)\right), \tag{2}$$

where S_n is $B(n, \mathbb{E}(Z_n)/n)-$distributed[1]. The key argument in the proof of this result is that $(Z_{n+1} - Z_n, Z_n)$ is N.A. for any $n \in \mathbb{N}$.

Next, we give an analogue of Theorem 1 for continuous time counting processes $(A_t)_{t\geq 0}$ (see Dellacherie-Meyer [5] or Brémaud for complete study of these processes and in particular for properties of their compensator). Our result for continuous time process states that under some appropriate hypothesis (Assumption 2 and Assumption 3), for any convex function ϕ

$$\mathbb{E}\left(\phi(A_t)\right) \leq \mathbb{E}\left(\phi(Y_t)\right), \tag{3}$$

[1] $B(n,p)$ is the binomial distribution with parameters n and p, $\mathcal{E}(\lambda)$ is the exponential distribution of parameter λ, $b(p)$ is the Bernoulli distribution of parameter p and $\mathcal{P}(\mu)$ is the Poisson distribution of parameter μ.

where Y_t is $\mathcal{P}(\mathbb{E}(Z_t))-$distributed (Assumption 2, concerning the absolutely continuity of the compensator of A_t is due to Reynaud-Bourret [10]). The proof of the continuous time theorem (Theorem 2) relies on differential equations.

Section 4 is devoted to applications of Theorems 1 and 2. First, we will prove that suprema of binomial (resp. Poissonian) independent random variables are more concentrated in the sence of convex concentration inequality than a single binomial (resp, Poissonian) variable. In other words, let $(Y_i)_{1\leq i\leq p}$ be independent random variables with distribution $B(n, p_i)$ and $Z_n = \sup(Y_1, \ldots, Y_p)$. For any convex function ϕ

$$\mathbb{E}\left(\phi(Z_n)\right) \leq \mathbb{E}\left(\phi(S_n)\right), \tag{4}$$

where S_n is $B(n, \mathbb{E}(Z_n)/n)-$distributed. In the same way, if $\left(N^{(i)}\right)_{1\leq i\leq p}$ are independent random variables with Poisson distribution with parameter μ_i and $A_t = \sup(N^{(1)}, \ldots, N^{(p)})$ for any convex function ϕ

$$\mathbb{E}\left(\phi(A_t)\right) \leq \mathbb{E}\left(\phi(P_t)\right), \tag{5}$$

where P_t is $\mathcal{P}(\mathbb{E}(A_t))-$distributed. The key argument here, is that we are able to compute the compensator of $(A_t)_{t\geq 0}$. The result is, in fact, a concentration inequality for the supremum of a set indexed Poisson process, when the class of sets is a class of disjoint sets. So it is quite natural to formulate the question below.

(Q) Does the process $(\sup(\Pi_t(A),\ A \in \mathcal{A}))_{t\geq 0}$ satisfies a convex
Poissonnian concentration inequality when $(\Pi_t)_{t\geq 0}$ is Poisson process .

Reynaud-Bourret (see [9]) proved that the answer is positive if we restrict the functions ϕ to be of the form $\phi_\lambda(x) = \exp(\lambda x)$. We can then conjecture that the question (Q) has a positive answer.

In the last application we study the example of $3-$ary search trees, and we show that they are an example for which theorem 1 is valid.

2 Definitions and statement of results

Let $(Z_n)_{n\in\mathbb{N}^\star}$ be a nondecreasing discrete time process, with $Z_0 = 0$ and with jumps equal to 1. In this paper we are interested in concentration inequalities for the process Z.

Definition 1 . *A process $(X_n)_n$ is said to satisfy a binomial convex concentration inequality if for any $n \in \mathbb{N}$, any convex function ϕ we have*

$$\mathbb{E}\left(\phi(X_n)\right) \leq \mathbb{E}\left(\phi(Y_n)\right), \tag{6}$$

where Y_n is $B(n, \mathbb{E}(X_n)/n)-$distributed.

We will also consider continuous time counting processes. We recall that $(A_t)_{t\geq 0}$ is a counting process if it is a random increasing piecewise constant function with $A_0 = 0$ and with jumps equal to 1 (for a complete description of these processes see Brémaud [2]). Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration and assume that $(A_t)_{t\geq 0}$ is $(\mathcal{F}_t)-$measurable. Let $(\Lambda_t)_{t\geq 0}$ be the compensator of the counting process A, i.e. the nondecreasing function such that $(M_t = A_t - \Lambda_t)_{t\geq 0}$ is a martingale (see Brémaud [2] or Dellacherie and Meyer [5] for a complete description of compensators). In the sequel we are interested in concentration inequalities for the process A.

Definition 2 . *A process $(X_t)_{t\geq 0}$ is said to satisfy a Poissonian convex concentration inequality, if for any $t \geq 0$ and any convex function ϕ we have*

$$\mathbb{E}\left(\phi(X_t)\right) \leq \mathbb{E}\left(\phi(Y_t)\right), \tag{7}$$

where Y_t is $\mathcal{P}(\mathbb{E}(X_t))-$distributed.

2.1 Theorem for discrete time processes

Let $(Z_n)_{n\in\mathbb{N}^\star}$ be a nondecreasing discrete time process, with $Z_0 = 0$ and jumps equal to 1 i.e. $Z_{n+1} - Z_n = 0$ or $Z_{n+1} - Z_n = 1$. We suppose that $(Z_n)_{n\in\mathbb{N}^\star}$ satisfies the following assumption

Assumption 1 . *For any fixed n, the sequence*

$$\left(\mathbb{P}(Z_{n+1} = k+1 \mid Z_n = k)\right)_{k\geq 0}$$

is nonincreasing.

Theorem 1 *(Discrete time). Under Assumption 1, the process* $(Z_n)_{n\in\mathbb{N}}$ *satisfies a binomial convex concentration inequality. In other words, for any convex function* ϕ *we have*

$$\mathbb{E}\left(\phi(Z_n)\right) \leq \mathbb{E}\left(\phi(Y_n)\right), \tag{8}$$

where Y_n *is* $B(n, \mathbb{E}(X_n)/n)-$*distributed.*

2.2 Theorem for continuous time processes

Let $(A_t)_{t\geq 0}$ be a counting process, whose compensator $(\Lambda_t)_{t\geq 0}$ satisfies the following two assumptions.

Assumption 2 . *The compensator* $(\Lambda_t)_{t\geq 0}$ *is absolutely continuous and a.s. finite on* $[0, T]$.

Note that Assumption 2 implies that A has a.s. a finite number of jumps (recall the jumps are equal to 1). In the sequel we will denote by λ_s the derivative of $d\Lambda_s$ with respect to ds (see Reynaud-Bourret [10] who introduces this assumption and gives other applications for counting processes).

Assumption 3 . $\mathbb{E}(\lambda_t \mid A_{t-})$ *is a nonincreasing function of* A_{t-}.

Theorem 2 *(Continuous time). Under assumptions 2 and 3, the process* $(A_t)_{t\geq 0}$ *satisfies a Poissonian convex concentration inequality. In other words , for any convex function* ϕ *we have*

$$\mathbb{E}\left(\phi(A_t)\right) \leq \mathbb{E}\left(\phi(Y_t)\right), \tag{9}$$

where Y_t *is* $\mathcal{P}(\mathbb{E}(X_t))-$*distributed.*

3 Proofs

3.1 Proof of theorem 1

Theorem 1 will be a consequence of Theorem 3, which is Theorem 1 in Shao [11]. We briefly recall Shao's setting. A finite family of random variables

$\{X_i,\ 1 \le i \le n\}$ is said to be negatively associated $(N.A.)$ if for every pair of disjoint subsets A_1 and A_2 of $\{1, 2, \ldots, n\}$,

$$\mathrm{Cov}\{f_1(X_i, i \in A_1), f_2(X_j, j \in A_2)\} \le 0, \tag{10}$$

whenever f_1 and f_2 are coordinatewise increasing and the covariance exists. An infinite family is N.A. if every finite subfamily is N.A.

Theorem 3 *(shao [11]). Let $\{X_i, 1 \le i \le n\}$ be a N.A sequence and let $\{X_i^\star, 1 \le i \le n\}$ be a sequence of independent random variables such that X_i and $X_i^\star$ have the same distribution for each $i = 1, 2, \ldots, n$. Then*

$$\mathbb{E}\Big((f \sum_{i=1}^{n} X_i)\Big) \le \mathbb{E}\Big(f\big(\sum_{i=1}^{n} X_i^\star\big)\Big) \tag{11}$$

for any convex function f on $\mathbb{R}$, whenever the expectation on the right hand side of (11) exists.

Remark 1 . *The proof of Theorem 3 requires only that (S_n, X_{n+1}) is N.A. for any $n \in \mathbb{N}$.*

Theorem 3 implies the following lemma.

Lemma 1 . *Let ϕ be a convex function. Under the assumptions of Theorem 1, we have*

$$\mathbb{E}(\phi(Z_n)) \le \mathbb{E}(\phi(S_n)), \tag{12}$$

where S_n is the sum of independent Bernoulli variables, $a_1 + \cdots + a_n$, such that $\mathbb{E}(a_i) = \mathbb{E}\,(Z_i - Z_{i-1})$.

Proof: Let $b_{n+1} = Z_{n+1} - Z_n$ and let us prove that (Z_n, b_{n+1}) is N.A for any $n \in \mathbb{N}$. Let $t \ge 0$, using Assumption 1 we have

$$\mathbb{P}(b_{n+1} = 1 \mid Z_n \ge t) \le \mathbb{P}(b_{n+1} = 1).$$

This equation can be written as

$$\mathbb{P}(Z_n \ge t,\ b_{n+1} = 1) \le \mathbb{P}(Z_n \ge t)\mathbb{P}(b_{n+1} = 1).$$

As $b_{n+1} = 0$ or $b_{n+1} = 1$ we get, for any $(s,t) \in \mathbb{R}^2$,

$$\mathbb{P}(Z_n \geq t,\ b_{n+1} \geq s) \leq \mathbb{P}(Z_n \geq t)\mathbb{P}(b_{n+1} \geq s).$$

In other words

$$\mathrm{Cov}\{I_{Z_n \geq t}\ I_{b_{n+1} \geq s}\} \leq 0. \tag{13}$$

From this inequality we get that, for any nondecreasing functions f and g,

$$\mathrm{Cov}\{f(Z_n), g(b_{n+1})\} \leq 0. \tag{14}$$

From (14) and Theorem 3 (cf. Remark 1), we then get Lemma 1. □

Theorem 1 is an easy consequence of both Lemma 1 and Proposition 1.

3.2 Proof of theorem 2

We will use differential equation technics to prove Theorem 2. The key point is the lemma below which gives a concrete description of the compensator of $\Big(\phi(A_t) - \phi(A_0)\Big)_{t\geq 0}$.

Lemma 2 . *Let ϕ be a nondecreasing convex function. Then the predictable compensator $\Big(\Lambda_t(A_t,\phi)\Big)_{t\geq 0}$ of $\Big(\phi(A_t) - \phi(A_0)\Big)_{t\geq 0}$ is defined by*

$$\Lambda_t(A_t,\phi) = \int_0^t \Big(\phi(1 + A_{s-}) - \phi(A_{s-})\Big)\lambda_s ds. \tag{15}$$

Proof: Using the fact that the process $(A_t)_{t\geq 0}$ is piecewise constant with jumps equal to 1 we have

$$\phi(A_t) - \phi(A_0) = \int_0^t \Big(\phi(1 + A_{s-}) - \phi(A_{s-})\Big)dA_s.$$

As the process $(A_t)_{t\geq 0}$ is càdlàg, the process $(A_{t-})_{t\geq 0}$ is left continuous and so is the process $(\phi(1 + A_{s-}) - \phi(A_{s-}))_{s\geq 0}$. Using Theorem T8, p. 27 in Brémaud [2] we get that

$$Y_t = \int_0^t \Big(\phi(1 + A_{s-}) - \phi(A_{s-})\Big)(dA_s - \lambda_s ds)$$

is a $(\mathcal{F}_t)-$martingale. This ends the proof of Lemma 2. $\square$

In order to prove Theorem 2, we will exhibit differential equations satisfied by $\mathbb{E}\left(\phi(A_t)\right)$ and $\mathbb{E}\left(\phi(N_t)\right)$. Denote by $\mathcal{C}$ the set of all convex functions and by $\mathcal{C}_2$ the set of all convex functions of the class C^2. Let $(N_t)_{t\geq 0}$ be a Poisson point process on $\mathbb{R}^+$, with $\mathbb{E}(N_t) = \mathbb{E}(A_t)$. Let

$$h(\phi, t) = \mathbb{E}\left(\phi(A_t)\right), \ g(\phi, t) = \mathbb{E}\left(\phi(N_t)\right).$$

For $a \in \mathbb{R}$, set $A_t^a = A_t + a, \ N_t^a = N_t + a$. Let

$$h_a(\phi, t) = \mathbb{E}\left(\phi(A_t^a)\right), \ g_a(\phi, t) = \mathbb{E}\left(\phi(N_t^a)\right).$$

Note that $h_0 = h$ and $g_0 = g$. Then

$$g(\phi, t) = \sum_{k=0}^{\infty} \phi(k) e^{\mathbb{E}(A_t)} \mathbb{E}(A_t)^k / k!.$$

Using the definition of λ_t and Fubini's theorem, we get

$$\frac{d}{dt}\mathbb{E}(A_t) = \mathbb{E}(\lambda_t).$$

Consequently

$$\frac{dg}{dt}(\phi, t) = \mathbb{E}(\lambda_t)\Big(- \sum_{k=0}^{\infty} \phi(k) e^{\mathbb{E}(A_t)} \mathbb{E}(A_t)^k / k! + \sum_{k=0}^{\infty} \phi(k+1) e^{\mathbb{E}(A_t)} \mathbb{E}(A_t)^k / k! \Big).$$

This equation can be written in the following way

$$\frac{dg}{dt}(\phi, t) = \mathbb{E}(\lambda_t)\Big(\mathbb{E}(\phi(N_t + 1) - \phi(N_t) \Big),$$

whence

$$\frac{dg_0}{dt}(\phi, t) + \mathbb{E}(\lambda_t) g_0(\phi, t) - \mathbb{E}(\lambda_t) g_1(\phi, t) = 0. \tag{16}$$

Let us now deal with h. From Lemma 2 we have

$$h(\phi, t) = \mathbb{E}(\phi(A_t)) = \mathbb{E}(\phi(A_0)) + \mathbb{E}\left(\int_0^t \Big(\phi(1 + A_{s-}) - \phi(A_{s-}) \Big) \lambda_s ds \right). \tag{17}$$

From Fubini's theorem $t \mapsto h(\phi, t)$ is absolutely continuous with respect to Lebesgue's measure and, denoting by $\frac{dh}{dt}$ its derivative,

$$\frac{dh}{dt}(\phi, t) = \mathbb{E}\Big(\phi(1 + A_{t-}) - \phi(A_{t-})\Big)\lambda_t\Big). \tag{18}$$

Let $\mathbb{E}^{A_{t-}}$ denote the expectation conditionally to A_{t-}. Then

$$\begin{aligned}\frac{dh}{dt}(\phi, t) &= \mathbb{E}\left(\mathbb{E}^{A_{t-}}\Big(\big(\phi(1 + A_{t-}) - \phi(A_{t-})\big)\lambda_t\Big)\right) \\ &= \mathbb{E}\Big(\big(\phi(1 + A_{t-}) - \phi(A_{t-})\big)\mathbb{E}^{A_{t-}}(\lambda_t)\Big).\end{aligned}$$

Now, using the convexity of ϕ, on one hand $\big(\phi(1+A_{t-}) - \phi(A_{t-})\big)$ is a nondecreasing function of A_{t-}, and on the other hand, from Assumption 3, $\mathbb{E}^{A_{t-}}(\lambda_t)$ is a nonincreasing function of A_{t-}. Hence $\big(\phi(1 + A_{t-}) - \phi(A_{t-}), \mathbb{E}^{A_{t-}}(\lambda_t)\big)$ is negatively associated, which ensures that

$$\frac{dh}{dt}(\phi, t) \leq \mathbb{E}\big(\phi(1 + A_{t-}) - \phi(A_{t-})\big)\mathbb{E}(\lambda_t).$$

From the convexity of ϕ

$$\phi(1 + A_{t-}) - \phi(A_{t-}) \leq \phi(1 + A_t) - \phi(1 + A_t),$$

because $A_{t-} \leq A_t$. Whence

$$\frac{dh}{dt}(\phi, t) \leq \mathbb{E}\big(\phi(1 + A_t) - \phi(A_t)\big)\mathbb{E}(\lambda_t). \tag{19}$$

In other words

$$\frac{dh_0}{dt}(\phi, t) + \big(h_0(\phi, t) - h_1(\phi, t)\big)\mathbb{E}(\lambda_t) \leq 0. \tag{20}$$

Replacing ϕ by $\phi_a : x \mapsto \phi(x + a)$ in (16) and (20). We get, for any $a \in \mathbb{R}$,

$$\frac{dg_a}{dt}(\phi, t) + \big(g_a(\phi, t) - g_{a+1}(\phi, t)\big)\mathbb{E}(\lambda_t) = 0, \tag{21}$$

$$\frac{dh_a}{dt}(\phi, t) + \big(h_a(\phi, t) - h_{a+1}(\phi, t)\big)\mathbb{E}(\lambda_t) \leq 0. \tag{22}$$

Now, define, for $u \in \mathbb{R}$ and $x \in \mathbb{R}^+$, the function ϕ_u by

$$\phi_u(x) = (u - x)_+ = \sup(u - x, 0),$$

and consider $\mathcal{E} = \{\phi_u, u \in \mathbb{R}\}$.
It is easy to see that $\phi_u(x + y) = 0$ as soon as $y \geq u$. Hence

$$h_y(\phi_u, t) = g_y(\phi_u, t) = 0 \quad \text{for any } y \geq u. \tag{23}$$

Let y be the first integer greater than u. From equations (23),

$$h_y(\phi_u, t) \leq g_y(\phi_u, t). \tag{24}$$

Now, let us prove, by backward induction on k, that $h_k(\phi_u, t) \leq g_k(\phi_u, t)$ for any k in $[0, y]$. If $h_k(\phi_u, t) \leq g_k(\phi_u, t)$ at rank k then

$$\frac{dg_{k-1}}{dt}(\phi_u, t) + g_{k-1}(\phi_u, t)\mathbb{E}(\lambda_t) = g_k(\phi_u, t)\mathbb{E}(\lambda_t), \tag{25}$$

$$\frac{dh_{k-1}}{dt}(\phi_u, t) + h_{k-1}(\phi_u, t)\mathbb{E}(\lambda_t) \leq g_k(\phi_u, t)\mathbb{E}(\lambda_t). \tag{26}$$

Both the initial condition

$$h_{k-1}(\phi_u, 0) = g_{k-1}(\phi_u, 0) = (u - k + 1)_+$$

and equations (25) and (26) imply that $h_{k-1}(\phi_u, t) \leq g_{k-1}(\phi_u, t)$. Hence by induction we get that $h_0(\phi_u, t) \leq g_0(\phi_u, t)$. Whence Theorem 2 is proved for any $\phi \in \mathcal{E}$.
Now, if $\phi \in \mathcal{C}_2$, thanks to Taylor formula, we can write

$$\phi(x) = \phi(0) + x\phi'(0) + \int_0^{+\infty} (x - u)_+ \phi''(u)du. \tag{27}$$

Now $(x - u)_+ = (x - u) + (u - x)_+$. Hence equation (27) becomes

$$\phi(x) = \phi(0) + x\phi'(0) + \int_0^{\infty} \left((x - u) + (u - x)_+\right) \phi''(u)du. \tag{28}$$

Then

$$\mathbb{E}\left(\phi(A_t)\right) = \phi(0) + \mathbb{E}(A_t)\phi'(0) + \mathbb{E}\Big(\int_0^{\infty} \left((A_t - u) + (u - A_t)_+\right) \phi''(u)du\Big).$$

As the functional inside the integral is nonnegative Fubini's theorem applies and consequently

$$\mathbb{E}\left(\phi(A_t)\right) = \phi(0) + \mathbb{E}(A_t)\phi'(0) + \int_0^\infty \mathbb{E}\Big(\left((A_t - u) + (u - A_t)_+\right)\Big)\phi''(u)du.$$

Now, from the validity of Theorem 2 for the elements of $\mathcal{E}$, we get

$$\mathbb{E}\left(\phi(A_t)\right) \leq \phi(0) + \mathbb{E}(N_t)\phi'(0) + \int_0^\infty \mathbb{E}\Big(\left((N_t - u) + (u - N_t)_+\right)\Big)\phi''(u)du.$$

Using again Fubini's theorem we have

$$\mathbb{E}\left(\phi(A_t)\right) \leq \mathbb{E}\left(\phi(N_t)\right). \tag{29}$$

We complete the proof using a density argument since $\mathcal{C}_2$ is dense in $\mathcal{C}$. □

4 Applications.

In this section, we will give applications of Theorems 1 and 2 of Section 2. The two first applications show that suprema of independent binomial random variables (resp. Poisson variables) satisfy a binomial (resp Poissonian) convex concentration inequality. The third deals with 3−ary search trees. We will show that the process of saturated nodes in an 3−ary search tree is an easy example of discrete time models which satisfied Assumption 1.

4.1 Supremum of binomial random variables

Let $p_1 \geq p_2 \geq \ldots \geq p_l$ be a nonincreasing sequence of reals. The aim of this section is to give a concentration inequality for $Z = \sup(B_1, \ldots, B_l)$ where the B_i's are independent $B(n, p_i)-$distributed random variables.

Theorem 4 . *If $Z = \sup(B_1, \ldots, B_l)$, then, for any convex function ϕ, we have $\mathbb{E}(\phi(Z)) \leq \mathbb{E}(\phi(Y))$, where the B_i's are independent $B(n, p_i)-$distributed random variables and $Y \sim B(n, \mathbb{E}(Z)/n)$.*

4.1.1 A discrete time representation

Here, we introduce a discrete time counting process $(Z_u)_{u\in\mathbb{N}}$ such that

$$Z = Z_{nl}. \tag{30}$$

Next, we apply Theorem 1 to Z_u with $u = nl$.
Let $X_{ij},\ i = 1..n,\ j = 1..l$ be independent variables such that X_{ij} is Bernoulli distributed with parameter p_j. Assume that the p_j's are nonincreasing. If $u = an + b$ with $0 \le b < n$, we define Z_u by

$$Z_u = \max\Big(S_n(p_1), S_n(p_2), \ldots, S_n(p_a), S_b(p_{a+1})\Big), \tag{31}$$

where $S_m(p_j) = \sum_{i=1}^{m} X_{ij}$ is $B(m, p_j)-$distributed.

Lemma 3 . *$(Z_u)_{u\in\mathbb{N}}$ satisfies the hypothesis of Theorem 1.*

The proof of Lemma 3 requires the two technical lemmas below whose proofs are postponed to the end of the section.

Lemma 4 . *Let Y be a $B(n,p)-$distributed random variable and let us denote by G its distribution function. Then for any $k \ge 1$ we have*

$$G^2(k) - G(k-1)G(k+1) \ge 0. \tag{32}$$

Lemma 5 . *Assume that $p_1 \ge p_2$. Set*

$$I_k(p_1, p_2) = \frac{\mathbb{P}\Big(S_j(p_2) < k\Big)\mathbb{P}\Big(S_n(p_1) = k\Big)}{\mathbb{P}\Big(S_j(p_2) = k\Big)\mathbb{P}\Big(S_n(p_1) \le k\Big)}. \tag{33}$$

Then the sequence $(I_k(p_1, p_2))_{k=1,\ldots,j}$ is nondecreasing with respect to k for any $j \in \{0, \ldots, n-1\}$.

4.1.2 Proof of lemma 3.

Define $F(m,j,l) = \mathbb{P}(S_m(p_j) \leq k)$,

$$N_{a+1}(k) = P(S_b(p_{a+1}) = k) \prod_{i=1}^{a} F(n,i,k),$$

and, for any $i \in \{1,\ldots,a\}$,

$$N_i(k) = F(b,a+1,k-1) \prod_{m=1}^{i-1} F(n,m,k)\mathbb{P}(S_n(p_i) = k) \prod_{m=i+1}^{a} F(n,m,k-1).$$

Set $u_k = \mathbb{P}(Z_{j+1} = k+1 \mid Z_j = k)$. Then

$$u_k = \frac{N_{a+1}(k)}{\sum_{i=1}^{a+1} N_i(k)} \mathbb{P}\big(X_b(p_j = 1)\big). \tag{34}$$

Let

$$c_k = 1 + \sum_{i=1}^{a} \frac{N_i(k)}{N_{a+1}(k)}. \tag{35}$$

From (34), we get that $(u_k)_{k\in\mathbb{N}}$ is nonincreasing if and only if $(c_k)_{k\in\mathbb{N}}$ is nondecreasing. Let

$$v_i(k) = \frac{N_i(k)}{N_{a+1}(k)}.$$

It is enough to prove that each sequence $(v_i(k))_{k\in\mathbb{N}}$ is nondecreasing . From the definition of the numbers $N_1(k),\ldots,N_{a+1}(k)$,

$$\begin{aligned} v_i(k) &= \frac{\mathbb{P}(S_j(p_a) < k)}{\mathbb{P}(S_j(p_a) = k)} \frac{\mathbb{P}(S_n(p_i) = k)}{\mathbb{P}(S_n(p_i) \leq k)} \frac{\prod_{m=i+1}^{a} \mathbb{P}(S_n(p_m) < k)}{\prod_{m=i+1}^{a} \mathbb{P}(S_n(p_m) \leq k)} \\ &= I_k(p_i, p_a) \frac{\prod_{m=i+1}^{a} \mathbb{P}(S_n(p_m) < k)}{\prod_{m=i+1}^{a} \mathbb{P}(S_n(p_m) \leq k)}. \end{aligned}$$

Using Lemmas 4 and 5 we get that $(v_i(k))_{k\in\mathbb{N}}$ is a product of two nondecreasing sequences. Consequently $(c_k)_{k\in\mathbb{N}}$ is nondecreasing, which ends up the proof. □

4.1.3 Proofs of the technical lemmas

Proof of lemma 4: It is a well known log-concavity result (see Pinelis [8] for instance). Anyway it can easily be proven by induction that

$$\log G(k) \geq \frac{1}{2}\Big(\log G(k-1) + \log G(k+1)\Big).$$

□

Proof of Lemma 5: We will first prove this lemma when $p_1 = p_2 = p$. Set $j = n-m$, then $I_k(p_1,p_2)$ becomes

$$I_k(p_1,p_2) = \frac{\mathbb{P}\Big(S_{n-m} < k\Big)\mathbb{P}\Big(S_n = k\Big)}{\mathbb{P}\Big(S_{n-m} = k\Big)\mathbb{P}\Big(S_n \leq k\Big)}.$$

Set

$$\tilde{I}_k = \frac{(n-m-k)!}{(n-k)!}\frac{\mathbb{P}\Big(S_{n-m} < k\Big)}{\mathbb{P}\Big(S_n \leq k\Big)},$$

then

$$I_k(p_1,p_2) = \frac{n!}{(n-m)!}(1-p)^m \tilde{I}_k.$$

As the factor in front of $\tilde{I}_k$ is independent of k, $(I_k(p_1,p_2))_{k\in\mathbb{N}}$ is nondecreasing if and only if $(\tilde{I}_k)_{k\in\mathbb{N}}$ is nondecreasing. Now let J_k be the inverse of $\tilde{I}_k$. Then $J_k = J_k^{(1)} + J_k^{(2)}$ with

$$\begin{cases} J_k^{(1)} = \dfrac{(n-k)!}{(n-m-k)!}\Big(\dfrac{\mathbb{P}(S_{n-m} < k, S_m \leq k - S_{n-m})}{\mathbb{P}(S_{n-m} < k)}\Big), \\ \\ J_k^{(2)} = \dfrac{(n-k)!}{(n-m-k)!}\Big(\dfrac{\mathbb{P}(S_{n-m} = k)\mathbb{P}(S_m = 0)}{\mathbb{P}(S_{n-m} < k)}\Big). \end{cases}$$

Consequently, it is enough to show that $\Big(J_k^{(1)}\Big)_{k\in\mathbb{N}}$ and $\Big(J_k^{(2)}\Big)_{k\in\mathbb{N}}$ are nonincreasing. Set $r_i = \mathbb{P}(S_{n-m} = i)$ and $q_i = \mathbb{P}(S_l \leq i)$. Then, setting

$$\gamma_1(k) = \frac{(n-k-1)!}{(n-m-k-1)!}\frac{1}{\sum_{j=0}^{k-1} r_j \sum_{j=0}^{k} r_j},$$

we get

$$\frac{J_k^{(1)} - J_{k+1}^{(1)}}{\gamma_1(k)} = (n-k)\sum_{i=0}^{k-1} r_i q_{k-i} \sum_{j=0}^{k} r_j - (n-m-k)\sum_{i=0}^{k} r_i q_{k-i} \sum_{j=0}^{k-1} r_j.$$

Hence $J_k^{(1)} - J_{k+1}^{(1)}$ has the same sign as

$$\delta_1(k) = \left((n-k)\sum_{i=0}^{k-1} r_i q_{k-i} \sum_{j=0}^{k} r_j - (n-m-k)\sum_{i=0}^{k} r_i q_{k-i} \sum_{j=0}^{k-1} r_j \right).$$

Now

$$\delta_1(k) \geq (n-k) r_k \Big(\sum_{i=0}^{k-1} r_i q_{k-i} - q_0 \sum_{i=0}^{k-1} r_i \Big).$$

The right hand side of this inequality is positive since the sequence $(q_i)_{i\in\mathbb{N}}$ is nondecreasing. Hence $(J_k^{(1)})_{k\in\mathbb{N}}$ is nonincreasing.
Let us deal now with $J_k^{(2)}$. Denoting by F the distribution function of S_{n-m} and setting

$$\gamma_2(k) = \frac{(n-k-1)!}{(n-m-k-1)!} \frac{(1-p)^m}{F(k-1)F(F)},$$

we have

$$\begin{aligned} J_k^{(2)} - J_{k+1}^{(2)} = \gamma_2(k) \Big(&(n-k)\big(F(k) - F(k-1)\big)F(k) \\ &- (n-m-k)\big(F(k+1) - F(k)\big)F(k-1). \Big). \end{aligned} \tag{36}$$

Using Lemma 4, we see that the right hand side of (36) is nonnegative. Then $(J_k^{(2)})_{k\in\mathbb{N}}$ is nonincreasing, whence $(J_k)_{k\in\mathbb{N}}$ is nonincreasing. Which implies that $(I_k)_{k\in\mathbb{N}}$ is nondecreasing.
Consider now the case where $p_1 > p_2$ and $j = n - m$. Write

$$I_k(p_1, p_2) = I_k(p_1, p_1) \frac{I_k(p_1, p_2)}{I_k(p_1, p_1)}$$

and set

$$L_k := \frac{I_k(p_1, p_2)}{I_k(p_1, p_1)} = \frac{\mathbb{P}(S_j(p_2) < k)\mathbb{P}(S_j(p_1) = k)}{\mathbb{P}(S_j(p_2) = k)\mathbb{P}(S_j(p_1) < k)}.$$

We now prove that $(L_k)_{k\in\mathbb{N}}$ is nondecreasing, which will be enough to conclude.
For $i = 1, 2$ set $r_i = p_i/(1-p_i)$ (note that as $p_1 > p_2$ we have $r_1 < r_2$). Setting

$$\gamma_3 = \left(\frac{1-p_1}{1-p_2}\right)^j,$$

we get

$$L_k = \gamma_3 \frac{r_1^k}{r_2^k} \frac{\mathbb{P}(S_j(p_2) < k)}{\mathbb{P}(S_j(p_1) < k)}.$$

Expanding L_k, we see that $L_{k+1} - L_k$ as the same sign as Δ_k, with

$$\Delta_k := \frac{\sum_{i=0}^{k-1} \binom{j}{i} r_1^{i+1}}{\sum_{i=0}^{k} \binom{j}{i} r_1^i} - \frac{\sum_{i=0}^{k-1} \binom{j}{i} r_2^{i+1}}{\sum_{i=0}^{k} \binom{j}{i} r_2^i}.$$

Let

$$C_k(r) = \frac{r\sum_{i=0}^{k-1} \binom{j}{i} r^i}{\sum_{i=0}^{k} \binom{j}{i} r^i} = \frac{rA_{k-1}(r)}{A_{k-1}(r) + \binom{j}{k} r^k}, \tag{37}$$

with $A_{k-1}(r) = \sum_{i=0}^{k-1} \binom{j}{i} r^i$. Then it is obvious that $\Delta_k = C_k(r_1) - C_k(r_2)$. Hence Lemma 5 will be proved if we prove that $C_k(r)$ is nondecreasing. Taking the derivative with respect to r in (37) we see that the sign of $C_k'(r)$ is the same as the one of

$$d_k(r) = A_{k-1}^2 + \binom{j}{j} r^k \left(rA_{k-1}'(r) - (k-1)A_{k-1}(r)\right).$$

Now

$$d_k(r) = \Big(\sum_{i=0}^{k-1} \binom{j}{i} r^i\Big)^2 - \binom{j}{k} \Big((k-1) + \sum_{i=1}^{k-1} \binom{j}{i} (k-1-i) r^i\Big) r^k \tag{38}$$

is a polynomial function in r for which the coefficient of r^{k+i} is

$$\sum_{u=i+1}^{k-1} \binom{j}{u} \binom{j}{k+i-u} - \binom{j}{k} \binom{j}{i} (k-1-i).$$

For $0 \le i \le 2k-2$ and $i+1 \le u \le k-1$, it is easy to check that

$$\binom{j}{i+1} \binom{j}{k-1} \le \binom{j}{u} \binom{j}{k+i-u},$$

whence

$$\binom{j}{k}\binom{j}{i} \leq \binom{j}{u}\binom{j}{k+i-u}. \tag{39}$$

This last inequality implies that $d_k(r)$ is nonnegative. Which concludes the proof of Lemma 5. □

4.2 Supremum of Poisson random variables

Let $\mu_1 \geq \ldots \geq \mu_p$ be a finite nonincreasing sequence of real numbers. The aim of this section is to give a concentration inequality for

$$W = \sup(Y_1, \ldots, Y_p)$$

where the Y_i's are independent and $\mathcal{P}(\mu_i)-$distributed.

Theorem 5 . *For any convex function ϕ we have if $Y \sim \mathcal{P}(\mathbb{E}(W))$ the following inequality*

$$\mathbb{E}(\phi(W)) \leq \mathbb{E}(\phi(Y)). \tag{40}$$

4.2.1 A continuous time model representation

Here, we introduce a continuous time counting process $(A_t)_{t\geq 0}$, such that

$$W = A_1. \tag{41}$$

Next, we will apply Theorem 2 to A_t with $t = 1$.
Let μ be equal to $\mu_1 + \ldots + \mu_p$. Let $(T_i)_{i\in\mathbb{N}^*}$ be i.i.d random variables $\mathcal{E}(\mu)-$distributed. Define the process S_n by $S_0 = 0$ and for any $n > 0$, $S_n = \sum_{j=1}^{n} T_j$. It is well known that $N_t = \sum_{k=1}^{+\infty} 1\!\!1_{\{S_k \leq t\}}$ is a Poisson point process. Consider now the nonincreasing sequence of reals $(t_i)_{1\leq i\leq p}$ with sum 1 define by $t_i\mu = \mu_i$. Let $a_i = \sum_{j=1}^{i} t_i$. Then $N_{a_i} - N_{a_{i-1}}$ is $\mathcal{P}(\mu_i)-$distributed. By homogeneity we can assume that $\mu = 1$ in the sequel. We define for $t \leq 1$,

$$N^{(i)} = N_{a_i} - N_{a_{i-1}} \tag{42}$$

and

$$k(t) = \sup(i,\ a_i \leq t).$$

We consider

$$A_t = \sup(N^{(1)}, \dots, N^{(k(t))}, N_t - N_{a_{k(t)}}).$$

Lemma 6 *(Compensator of A_t). If $a_i \le t < a_{i+1}$ define λ_t by setting $\lambda_t = 1$ if $A_t = N_t - N_{a_{k(t)}}$ and $\lambda_t = 0$ otherwise. Then $\Lambda_t = \int_0^t \lambda_u du$ is the compensator of A_t.*

Proof of Lemma 6: Let t be in $[a_i, a_{i+1}[$ and $s < t$ we will show that

$$\mathbb{E}^{\mathcal{F}_s}\left(A_t - A_s - \int_s^t \lambda_u du\right) = 0. \tag{43}$$

Suppose first that $s > a_i$. Then

$$A_t = \sup\{A_{a_i}, N_t - N_{a_i}\}, \tag{44}$$

$$A_s = \sup\{A_{a_i}, N_s - N_{a_i}\}. \tag{45}$$

If $A_s = N_s - N_{a_i}$, we have $A_t = N_t - N_{a_i}$ and $\lambda_u = 1$ for any $u \in [s, t]$. As the event

$$B := \{A_s = N_s - N_{a_i}\} \tag{46}$$

is $\mathcal{F}_s$-measurable and $N_t - N_s$ is independent of $\mathcal{F}_s$, we have

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_s}\left(\left(A_t - A_s - \int_s^t \lambda_u du\right)1\!\!1_{\{B\}}\right) &= \mathbb{E}^{\mathcal{F}_s}\left(\left(N_t - N_s - (t-s)\right)1\!\!1_{\{B\}}\right) \\ &= 1\!\!1_{\{B\}}\mathbb{E}^{\mathcal{F}_s}\left(N_t - N_s - (t-s)\right) \\ &= 1\!\!1_{\{B\}}\mathbb{E}\left(N_t - N_s - (t-s)\right) = 0. \end{aligned} \tag{47}$$

Now on B^c we have $N_s - N_{a_i} < A_s$, whence $A_s = A_{a_i}$. Then

$$\delta := A_{a_i} - (N_s - N_{a_i})$$

is a positive number and is $\mathcal{F}_s$-measurable. Now if $N_t - N_s < \delta$ we have $A_t = A_{a_i}$ and $\lambda_u = 0$ for all $u \in [s, t]$. This implies

$$\mathbb{E}^{\mathcal{F}_s}\left(\left(A_t - A_s - \int_s^t \lambda_u du\right)1\!\!1_{\{B^c\}}1\!\!1_{\{N_t - N_s < \delta\}}\right) = 0. \tag{48}$$

If $N_t - N_s \geq \delta$ let τ be the first time in $]s,t]$ such that $N_\tau - N_{a_i} = A_{a_i}$. Note that τ is a $(\mathcal{F}_t)_{t\geq 0}$-stopping time. Then $A_t = N_t - N_{a_i}$, $A_s = A_{a_i}$, $\lambda_u = 0$ if $u \in [s,\tau[$ and $\lambda_u = 1$ if $u \in [\tau, t]$. Hence

$$\mathbb{E}^{\mathcal{F}_s}\Big(\big(A_t - A_s - \int_s^t \lambda_u du\big)\mathbb{1}_{\{B^c\}}\mathbb{1}_{\{N_t - N_s \geq \delta\}}\Big) = \\ \mathbb{E}^{\mathcal{F}_s}\Big(\Big(N_t - N_\tau - (t-\tau)\Big)\mathbb{1}_{\{B^c\}}\mathbb{1}_{\{N_t - N_s \geq \delta\}}\Big). \tag{49}$$

Clearly $\{N_t - N_s > \delta\} = \{N_t - N_\tau > 0\} = \{\tau \leq t\}$. Therefore

$$\mathbb{E}^{\mathcal{F}_s}\Big(\big(A_t - A_s - \int_s^t \lambda_u du\big)\mathbb{1}_{\{B^c\}}\mathbb{1}_{\{N_t - N_s \geq \delta\}}\Big) = \\ \mathbb{1}_{\{B^c\}}\mathbb{E}^{\mathcal{F}_s}\Big(\mathbb{E}^{\mathcal{F}_\tau}\Big(\big(N_t - N_\tau - (t-\tau)\big)\mathbb{1}_{\{N_t - N_\tau \geq 0\}}\Big)\Big) = 0. \tag{50}$$

Putting together equations (47) (48) (50), we get

$$\mathbb{E}^{\mathcal{F}_s}\left(A_t - A_s - \int_s^t \lambda_u du\right) = 0, \tag{51}$$

for any $a_i < s < t < a_{i+1}$. Using similar arguments we see that equation (51) is valid for any $0 < s < t < 1$. □

Thanks to the following lemma, Theorem 5 is an easy consequence of Theorem 2, with $t = 1$.

Lemma 7 . *$(A_t, \lambda_t)_t$ satisfies Assumption 3*

Proof of Lemma 7: First, for a fixed t, $A_t = A_{t-}$ almost surely. Hence $\mathbb{E}(\lambda_t \mid A_{t-}) = \mathbb{E}(\lambda_t \mid A_t)$ almost surely. If $t \in [a_i,\ a_{i+1}[$, then $\lambda_t = 1$ iff $A_t = N_t - N_{a_i}$. Consequently

$$\mathbb{E}(\lambda_t \mid A_t) = \mathbb{P}(A_t = N_t - N_{a_i} \mid A_t). \tag{52}$$

We now prove that the sequence $(u_k(t))_k$ defined below is nonincreasing. Let

$$u_k(t) := \mathbb{P}(A_t = N_t - N_{a_i} \mid A_t = k). \tag{53}$$

If $V_i = N_t - N_{a_i}$ then

$$u_k(t) = \frac{\mathbb{P}(V_i = k, N^{(j)} \leq k,\ j = 1..i)}{\mathbb{P}(A_t = k)}. \tag{54}$$

The end of the proof needs the following lemma whose proof is postponed to the end of the section.

Lemma 8 . (i) *For any* $1 \leq j \leq i$,

$$v_k(j) = \frac{\mathbb{P}(V_i = k,\ N^{(j)} \leq k)}{\mathbb{P}(V_i = k,\ N^{(j)} \leq k) + \mathbb{P}(V_i < k,\ N^{(j)} = k)}. \tag{55}$$

is nonincreasing with respect to k.
(ii) *For any* $i > 0$,

$$M_i(k) = \frac{\mathbb{P}(N^{(i)} < k)}{\mathbb{P}(N^{(i)} \leq k)} \tag{56}$$

is nondecreasing with respect to k.

Recall we have to show that $(u_k(t))_{k\in\mathbb{N}}$ is nonincreasing. From the independence of the random variables $N^{(j)}$ and $N_t - N_{a_i}$, setting

$$\begin{aligned} D(k) =& \mathbb{P}(V = k) \prod_{j=1}^{i} \mathbb{P}(N^{(j)} \leq k) \\ &+ \mathbb{P}(V < k) \sum_{u=0}^{i-1} \mathbb{P}(N^{(i-u)} = k) \prod_{l > i-u} \mathbb{P}(N^{(l)} < k) \prod_{l < i-u} \mathbb{P}(N^{(l)} \leq k), \end{aligned}$$

we get

$$u_k(t) = \frac{\mathbb{P}(V = k) \prod_{j=1}^{i} \mathbb{P}(N^{(j)} \leq k)}{D(k)}. \tag{57}$$

Set $W_k(t) = (u_k(t))^{-1}$. Rewriting (57) we have

$$\begin{aligned} W_k =& 1 + \frac{\mathbb{P}(V < k)\mathbb{P}(N^{(i)} = k)}{\mathbb{P}(V = k)\mathbb{P}(N^{(i)} \leq k)} + \frac{\mathbb{P}(V < k)\mathbb{P}(N^{(i)} < k)\mathbb{P}(N^{(i-1)} = k)}{\mathbb{P}(V = k)\mathbb{P}(N^{(i)} \leq k)\mathbb{P}(N^{(i-1)} \leq k)} \\ &+ \cdots + \frac{\mathbb{P}(V < k)\mathbb{P}(N^{(i)} < k) \ldots \mathbb{P}(N^{(2)} < k)\mathbb{P}(N^{(1)} = k)}{\mathbb{P}(V = k)\mathbb{P}(N^{(i)} \leq k) \ldots \mathbb{P}(N^{(1)} \leq k)}. \end{aligned}$$

Any component of this sum is of the form $\alpha_k\beta_k$, with

$$\alpha_k = \frac{\mathbb{P}(V < k)\mathbb{P}(N^{(j)} = k)}{\mathbb{P}(V = k)\mathbb{P}(N^{(j)} \leq k)} \text{ and } \beta_k = \prod_l \frac{\mathbb{P}(N^{(l)} < k)}{\mathbb{P}(N^{(l)} \leq k)}.$$

Using Lemma 8, we see that both α_k and β_k are nondecreasing sequences of k, which ends the proof of Lemma 7. □

4.2.2 Proof of lemma 8

(i) Let $t \in [a_i,\ a_{i+1}[$. Let us prove that $(v_k(j))_k$ is nonincreasing for all $j \leq i$. Recall that

$$v_k(j) = \frac{\mathbb{P}(V_i = k,\ N^{(j)} \leq k)}{\mathbb{P}(V_i = k,\ N^{(j)} \leq k) + \mathbb{P}(V_i < k,\ N^{(j)} = k)}. \tag{58}$$

Fix j and set $v_k = v_k(j)$. It is enough to show that

$$U_k = \frac{\mathbb{P}(V < k)\mathbb{P}(N^{(j)} = k)}{\mathbb{P}(V = k)\mathbb{P}(N^{(j)} \leq k)} \tag{59}$$

is nondecreasing in k.

As $\theta_{i+1} < t_j$, and $t - a_i < \theta_j$, there exists θ in $]0, 1[$ such that $t - a_i = \theta t_j$. Equation (59) becomes now

$$U_k = \frac{\sum_{n=0}^{k-1} \theta^n t_j^n / n!}{\theta^k \sum_{n=0}^{k} t_j^n / n!}. \tag{60}$$

Now, we set, $S_j(u) = \sum_{n=0}^{u} t_j^n / n!$ and

$$K(\theta, k) = \theta^{k+1} S_j(k) S_j(k+1).$$

Then

$$(U_{k+1} - U_k)K(\theta, k) = S_j(k) \left(\sum_{n=0}^{k} \frac{\theta^n t_j^n}{n!} \right) - \theta S_j(k+1) \left(\sum_{n=0}^{k-1} \frac{t_j^n \theta^n}{n!} \right).$$

We will show that for $\theta \in]0, 1[$ and $t_j > 0$,

$$P(\theta) = S_j(k) \left(\sum_{n=0}^{k} \frac{\theta^n t_j^n}{n!} \right) - \theta S_j(k+1) \left(\sum_{n=0}^{k-1} \frac{t_j^n \theta^n}{n!} \right)$$

is nonnegative. We have

$$P(\theta) = S_j(k) + \sum_{i=1}^{k} \left(\frac{t_j^i}{i!} S_j(k) - \frac{t_j^{i-1}}{(i-1)!} S_j(k+1) \right) \theta^i. \tag{61}$$

Let us first simplify the coefficient in front of θ

$$\begin{aligned} P(\theta) = S_j(k) &+ \theta \left(S_j(k) t_j - S_j(k+1) \right) \\ &+ \sum_{i=2}^{k} \left(\frac{t_j^i}{i!} S_j(k) - \frac{t_j^{i-1}}{(i-1)!} S_j(k+1) \right) \theta^i. \end{aligned} \tag{62}$$

Now using the fact that $\theta < 1$ and $\theta > \theta^2$ we get

$$P(\theta) \geq \theta \left(t_j S_j(k-1) + \frac{k t_j^{k+1}}{(k+1)!} \right) + \sum_{i=2}^{k} \left(\frac{t_j^i}{i!} S_j(k) - \frac{t_j^{i-1}}{(i-1)!} S_j(k+1) \right) \theta^i. \tag{63}$$

We can do the same thing for each power of θ, and we get that $P(\theta) \geq 0$ for any $\theta \in [0,1]$.
(ii) Let $P_k = \sum_{n=1}^{k} t_i^n / n! = \sum_{n=1}^{k} c_n$, then

$$\begin{aligned} \frac{M_i(k+1)}{M_i(k)} = \frac{P_k^2}{P_{k+1} P_{k-1}} &= \frac{P_{k-1}^2 + c_k(2P_{k-1} + c_k)}{P_{k-1}^2 + c_k(P_{k-1} + c_{k+1} P_{k-1}/c_k)} \\ &= \frac{P_{k-1}^2 + c_k P_{k-1} + c_k P_k}{P_{k-1}^2 + c_k P_{k-1} + c_{k+1} P_{k-1}}. \end{aligned}$$

Expanding the polynomial function

$$\delta_k = c_k P_k - c_{k+1} P_{k-1} = c_k \left(P_k - \frac{t_i}{k+1} P_{k-1} \right),$$

we see that $\delta_k > 0$, which implies that $M_i(k)$is nondecreasing. □

4.3 3-ary search trees

An 3-ary search tree is a data structure that grows by the progressive insertion of keys into a tree with branch factor 3. Each node contains $0, 1, 2$ keys

and gives rise to 3 branches as soon as it contains 2 keys. We call saturated a node containing two keys.
For each $i \in \{1,2,3\}$ let $X_n(i)$ denote the number of nodes containing $i-1$ keys after having introduced $n-1$ keys in the tree. The purpose is to give a binomial convex concentration inequality for $X_n^{(3)}$. In other words we have the following theorem.

Theorem 6 . *The number of saturated nodes in an $3-$ ary search tree satisfies a binomial convex concentration inequality, i.e. for any convex function ϕ,*

$$\mathbb{E}\left(\phi(X_n^{(3)})\right) \leq \mathbb{E}\left(\phi(Y)\right), \tag{64}$$

where Y is $B(n, \mathbb{E}(X_n^{(3)})/n)-$distributed.

4.3.1 Construction of an $3-$ary tree

Let us first recall Chauvin and Pouyanne description of $3-$ary search trees (see [4] for a general description of $m-$ary search trees). One throws a sequence of numbers in $[0,1]$ named keys, uniformly in $[0,1]^{\mathbb{N}^\star}$. The keys are placed one after another in an 3-ary tree (one node root, from each node grow 3 branches). The following recursive rule describes the way a key named k is inserted in the tree.
i) If the root contains strictly less than $m-1$ keys, then k is inserted in the root. One draws usually keys in a root from left to right in increasing order.
ii) If the root is already saturated, i.e. if it contains $m-1$ keys named k_1, k_2, ordered such that $k_1 < k_2$, then corresponds to each interval $I_1 =]-\infty, k_1[,\ I_2 =]k_1, k_2[,\ I_3 =]k_2, \infty[$ a subtree being itself an 3-ary search tree. one draws usually the branches corresponding to I_1, I_2, I_3 from left to right. In this situation, k is inserted in the subtree that corresponds to the interval I_j such that $k \in I_j$. Let $\mathcal{F}_n$, the $\sigma-$field generated up to time n. For each $i \in \{1,2,3\}$ and $n \geq 1$, we define $X_n^{(i)}$ as the number of node which contains $i-1$ keys after the insertion of the $n-1-$th key; such nodes are named nodes of type i. Nodes of type m are called saturated.

4.3.2 Proof of theorem 6

We are interested in the saturated nodes. We recall the two following equations that can be found in [4].

$$n - 1 = 2X_n^{(3)} + X_n^{(2)}, \tag{65}$$

$$n = X_n^{(1)} + 2X_n^{(2)}. \tag{66}$$

Hence if $X_n^{(3)}$ is known then $X_n^{(1)}$ and $X_n^{(2)}$ are also known. It is clear that $X_{n+1}^{(3)} = X_n^{(3)}$ or $X_{n+1}^{(3)} = X_n^{(3)} + 1$ (the number of saturated nodes is an nondecreasing function). $X_n^{(3)}$ increases only if the n-th keys is added in a node of type 2, this is done with probability $\frac{2}{n}X_n^{(2)}$. Hence

$$\begin{aligned}\mathbb{P}(X_{n+1}^{(3)} = X_n^{(3)} + 1 \mid X_n^{(3)}) &= \mathbb{P}(X_{n+1}^{(3)} = X_n^{(3)} + 1 \mid X_n^{(1)}, X_n^{(2)}) \\ &= \frac{2}{n}X_n^{(2)} = \frac{2}{n}(n - 1 - 2X_n^{(3)}).\end{aligned}$$

The last equation implies that $(X_n^{(3)})_{n\in\mathbb{N}}$ satisfies Assumption 1. Theorem 6 follows. □

Remark 2 . *We proved that for $m = 3$, the process of saturated nodes $\left(X_n^{(m)}\right)_{n\in\mathbb{N}}$ satisfies a binomial convex concentration inequality. The problem to know if this concentration inequality holds for $m > 3$ is open.*

References

[1] V. Bentkus. *on Hoeffding's inequality* To appear in Ann. of Probab. 2003

[2] P. Brémaud. Point processes and queues. Springer-Verlag. 1981.

[3] J. Bretagnolle. Statistique de Kolmogorov-Smirnov pour un échantillon non équiréparti dans l' aspects statistics et aspects physiques des processus gaussiens. *Paris: Editions du centre nationnal de la recherche scientifique.* 1981.

[4] B. Chauvin, N. Pouyanne. $m-$ary search trees when $m > 26$: a strong asymptotics for the space requirements. Submitted to *Random Structures and Algorithms* 2002.

[5] C. Dellacherie, P.A. Meyer. Probabilités et potentiel. Chap. 5 to 8 Hermann, Paris. 1980.

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

[7] I. Pinelis. Optimal Bounds for the Distributions of Martingales in Banach Spaces. *Ann. Prob.* Vol. 22. No. 4, 1679-1706 1994

[8] I. Pinelis. Optimal Tail Comparison Based on Comparison Moments *Eberlein, Ernst (ed.) et al., High dimensional probability. Proceedings of the conference, Oberwolfach, Germany, August 1996. Basel: Birkhäuser. Prog. Probab.* Vol. 43, 297-314 1998.

[9] P. Reynaud-Bourret. Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities. *Probab. Theory. relat. Fields.* 126, 103-153 2003.

[10] P. Reynaud-Bourret. Exponential inequalities for counting processes. *preprint of the Georgia Institute of Technology.* 2002.

[11] Q.M. Shao. A Comparison Theorem on Moment Inequalies Between Negatively Associated and Independent Random Variables. *Journ. of Theor. Prob.*, 2000. Vol. 13, 2. 343-356.

[12] G. Shorack, J. Wellner. Empirical processes with applications to statistics. John Willey and Sons, Inc., New York. 1986.