

Corrigé ENS 2011

Corrigé proposé par Martine Ginestet et Francis Denise (UPA)

Remerciements à Sébastien Kerner pour les tableaux de variations

I Caractérisation algorithmique de la taille d'un cluster

1. Le nombre maximal de voisin de i est égal à $n - 1$

Notons Z_i le nombre de voisins de i

$Z_i = \sum_{j \neq i} Z_{i,j}$, somme de Bernoulli indépendantes de même paramètre p d'où

$$Z_i \hookrightarrow \mathcal{B}(n-1, p), E(Z_i) = (n-1)p, V(Z_i) = (n-1)pq$$

2. Exemple

	1	2	3	4	5	6
Etape 0	A	N	N	N	N	N
Etape 1	I	A	A	N	N	N
Etape 2	I	I	A	A	N	N
Etape 3	I	I	I	A	N	N
Etape 4	I	I	I	I	N	N

$|C_1| = 4$, l'algorithme s'arrête à l'étape 4

A chaque étape, on rend inactif un voisin, donc forcément, les inactifs en fin d'algorithme sont des sommets reliés à 1.

D'autre part, à la fin de l'algorithme, il ne reste plus de sommets activés.

Si on suppose qu'il existe un sommet du cluster neutre, il existe un chemin le reliant au sommet 1, et tous les sommets de ce chemin ont été activés à un moment, puis rendus inactifs.

L'ensemble des sommets inactifs est égal au cardinal du cluster.

A chaque étape, un sommet est inactivé, donc l'algorithme s'arrête à l'étape $|C_1|$

3. a) A_0 est la variable certaine égale à 1, A_{K_0} est la variable certaine égale à 0
 $P(A_1 = 0) = P(\text{'aucun sommet n'est relié à 1'}) = P(Z_1 = 0) = q^{n-1}$

- b) A chaque étape 1 et 1 seul sommet est inactivé, notons I_k le nombre de sommets inactifs

$$\begin{cases} \text{si } k \leq K_0, & I_k = k \\ \text{si } k > K_0, & I_k = K_0 \end{cases}$$

$$|C_1| = I_{K_0} = K_0$$

K_0 est le premier entier k tel que $A_k = 0$

- c) Si $k \leq K_0$, à l'étape $k - 1$, il y a :

- A_{k-1} actifs ; il en restera parmi cela $A_{k-1} - 1$ actifs à l'étape k

- il y a $k - 1$ inactifs

- B_k est le nombre de sommets voisins de i_k qui sont neutres et qui deviennent alors actifs à l'étape k . Il reste $n - (k - 1) - A_{k-1}$ sommets neutres qui ont chacun la probabilité p d'être voisin de i_k , de manière indépendante. D'où

$$B_k \hookrightarrow \mathcal{B}(n - k + 1 - A_{k-1}, p)$$

Le nombre d'actifs à l'étape k est alors : $\boxed{\text{Si } k \leq K_0, \quad A_k = A_{k-1} - 1 + B_k}$

II. Etude d'une approximation déterministe de la taille d'un cluster

4. $P(X = k) = \sum_{m=1}^{\infty} P(X = k/N = m)P(N = m)$
 $E(X) = \sum_{k=0}^{\infty} kP(X = k) = \sum_{m=1}^{\infty} P(N = m)E(X/N = m)$ (**intersion des $\sum$, non au programme des BCPST**)
 $E(X) = \sum_{m=1}^{\infty} P(N = m)mq = qE(N)$ car $X/(N = m) \hookrightarrow \mathcal{B}(m, q)$

$$\boxed{E(X) = qE(N)}$$

5. A_k étant majorée par n , elle admet une espérance

$$\begin{aligned} E(A_k) &= E(1_{k \leq K_0} A_k) + E(1_{k > K_0} A_k) = E(1_{k \leq K_0} A_k) \text{ car si } k \geq K_0 \text{ alors } A_k = 0 \\ E(1_{k \leq K_0} A_k) &= E(1_{k \leq K_0} (A_{k-1} - 1 + B_k)) = E(1_{k \leq K_0} A_{k-1}) - E(1_{k \leq K_0}) + E(1_{k \leq K_0} B_k) \\ E(A_{k-1}) &= E(1_{k \leq K_0} A_{k-1}) \text{ car si } k > K_0 \text{ alors } A_{k-1} = 0 \\ E(1_{k \leq K_0}) &= P(k \leq K_0) \\ E(1_{k \leq K_0} B_k) &= E(B_k) P(k \leq K_0) \\ E(B_k) &= pE(n - k + 1 - A_{k-1}) = p(n - k + 1 - E(A_{k-1})) \text{ d'après le 4.} \\ \text{D'où } E(A_k) &= E(A_{k-1}) + P(k \leq K_0) [-1 + p(n - k + 1 - E(A_{k-1}))] \\ \text{Pour } k \text{ fixé et } n \text{ grand, } P(k \leq K_0) &\text{ est proche de 1 (on peut l'admettre?)} \\ \text{On peut donc approximer par :} \\ E(A_k) &\simeq E(A_{k-1}) + [-1 + p(n - k + 1 - E(A_{k-1}))] = (1 - p)E(A_{k-1}) - 1 + p(n - k + 1 - E(A_{k-1})) \\ \text{En remplaçant par :} \end{aligned}$$

$$\boxed{E(A_k) \simeq (1 - \frac{\lambda}{n})E(A_{k-1}) + \lambda - 1 - \lambda \frac{k-1}{n}}$$

$$\boxed{E(A_0) = 1}$$

6. $u_k = a_k + \alpha + \beta k$

$$u_{k-1} = a_{k-1} + \alpha + \beta(k-1)$$

$$\text{En remplaçant : } u_k = \left(1 - \frac{\lambda}{n}\right) u_{k-1} + k \frac{\lambda}{n} (\beta - 1) + \frac{\lambda}{n} (\alpha - \beta + 1) + \beta + \lambda - 1$$

$$\text{Choisissons : } \begin{cases} \beta - 1 = 0 \\ \frac{\lambda}{n} (\alpha - \beta + 1) + \beta + \lambda - 1 = 0 \end{cases} \iff \begin{cases} \beta = 1 \\ \alpha = -n \end{cases}$$

$$\text{Alors : } u_k = \left(1 - \frac{\lambda}{n}\right) u_{k-1} \text{ et } u_0 = 1 - n \text{ d'où } u_k = \left(1 - \frac{\lambda}{n}\right)^k (1 - n)$$

$$\boxed{\forall k, a_k = (1 - n) \left(1 - \frac{\lambda}{n}\right)^k + n - k}$$

7. a) $f_n(y) = n - y - (n-1)e^{-b_n y}$ $b_n = -\ln\left(1 - \frac{\lambda}{n}\right) > 0$

$$f'_n(y) = -1 + b_n(n-1)e^{-b_n y}$$

$$f'_n(y) > 0 \iff e^{-b_n y} > \frac{1}{b_n(n-1)} \iff y < \frac{\ln(b_n(n-1))}{b_n} = \alpha_n$$

f_n est croissante puis décroissante, la valeur maximale est atteinte en α_n

$$b_n(n-1) \sim \frac{\lambda}{n} = \lambda$$

Si $\lambda \neq 1$, $\lim \ln(b_n(n-1)) = \ln \lambda$ qui est du signe de $\lambda - 1$

Donc pour n assez grand, α_n est du signe de $\lambda - 1$

b) $-b_n \ln n = \ln\left(1 - \frac{\lambda}{n}\right) \ln n \sim -\frac{\lambda}{n} \ln n$ d'où $-b_n \ln n = -\lambda \frac{\ln n}{n} + o\left(\frac{\ln n}{n}\right)$

$$e^u = 1 + u + o(u) \text{ et } \lim \frac{\ln n}{n} = 0 \text{ d'où } e^{-b_n \ln n} = 1 - \lambda \frac{\ln n}{n} + o\left(\frac{\ln n}{n}\right)$$

$$f_n(\ln n) = n - \ln n - (n-1) \left[1 - \lambda \frac{\ln n}{n} + o\left(\frac{\ln n}{n}\right)\right]$$

$$f_n(\ln n) = -\ln n + 1 + \lambda \frac{\ln n}{n} (n-1) + o(\ln n) = -\ln n + \lambda \ln n + o(\ln n)$$

$$\boxed{\text{Si } \lambda \neq 1, f_n(\ln n) \sim (\lambda - 1) \ln n}$$

Si $\lambda = 1$, $f_n(\ln n) = n - \ln n - (n-1) \exp(\ln(1 - \frac{1}{n}) \ln n)$

$$\ln\left(1 - \frac{1}{n}\right) \ln n = -\frac{1}{n} \ln n - \frac{1}{2n^2} \ln n + o\left(\frac{1}{n^2} \ln n\right)$$

$$\exp(\ln(1 - \frac{1}{n}) \ln n) = 1 - \frac{1}{n} \ln n + \frac{1}{2n^2} (\ln n)^2 + o\left(\frac{1}{n^2} (\ln n)^2\right)$$

$$f_n(\ln n) = n - \ln n - n + 1 + \ln n - \frac{1}{2n} (\ln n)^2 + o\left(\frac{1}{n} (\ln n)^2\right) = 1 - \frac{1}{2n} (\ln n)^2 + o\left(\frac{1}{n} (\ln n)^2\right)$$

$$\boxed{\text{Si } \lambda = 1, f_n(\ln n) \rightarrow 1}$$

c) Soit $\alpha \in [0, 1]$

$$\frac{f_n(\alpha n)}{n} = 1 - \alpha - \left(1 - \frac{1}{n}\right) e^{-\alpha n b_n} \quad \text{et} \quad \alpha n \ln\left(1 - \frac{\lambda}{n}\right) \sim -\alpha \lambda$$

$$\boxed{\lim \frac{f_n(\alpha n)}{n} = g(\alpha) = 1 - \alpha - e^{-\alpha \lambda}}$$

$$g'(\alpha) = -1 + \lambda e^{-\alpha \lambda}$$

$$g''(\alpha) = -\lambda^2 e^{-\alpha \lambda} < 0$$

$$g'(0) = -1 + \lambda > 0$$

$$g'(1) = -1 + \lambda e^{-\lambda} < 0 \quad (\text{Etudier } h : \lambda \mapsto -1 + \lambda e^{-\lambda})$$

$$g(0) = 0 \text{ et } g(1) = 1 - \alpha - e^{-\alpha} < 0$$

x	0	α^*	1
$g'(x)$	+	0	-
$g(x)$	0	0	-

Thm de la bijection : $\boxed{\text{il existe un unique } \alpha^* \in]0, 1[, \text{ tel que } g(\alpha^*) = 0}$

Remarque : il n'est pas utile de montrer que g' s'annule, car comme $g(1) < 0$, $g(0) = 0$ et $g'(0) > 0$, g admet forcément un maximum sur $]0, 1[$

8. On admet que $|C_1|$ est le premier entier k tel que $a_k \leq 0$

On cherche donc le premier entier k tel que $f_n(k) \leq 0$

Si $\lambda > 1$

Si $\alpha > \alpha^*$, $g(\alpha) < 0$, donc pour n assez grand, $f_n(\alpha n) < 0$ d'où $|C_1| \leq \alpha n$

Si $\beta < \alpha^*$, $g(\beta) > 0$, donc pour n assez grand, $f_n(\beta n) > 0$ d'où $|C_1| \geq \beta n$

Soit $\varepsilon > 0$, pour n assez grand, $(\alpha^* - \varepsilon)n \leq |C_1| \leq (\alpha^* + \varepsilon)n$

$$\boxed{|C_1| \sim \alpha^* n}$$

Si $\lambda < 1$

pour n assez grand, $f_n(\ln n) < 0$ d'où

$$\boxed{\text{pour } n \text{ assez grand, } |C_1| \leq \ln n}$$

Cela donne la taille du cluster 1 et donc la proportion de la population infectée par l'agent pathogène.

On a donc montré que si $\lambda > 1$, on obtient un pourcentage α^* infecté et si $\lambda < 1$, une quantité négligeable est atteinte

III. Quelques propriétés d'une marche aléatoire

9. D'après la définitions de K , $S_K \leq 0$ et $S_{K-1} > 0$

Or $S_K = S_{K-1} + X_K$ et X_K prend ses valeurs dans $\{-1, 0, 1, \dots, m-1\}$

Donc $S_K \leq 0$ et $S_{K-1} > 0$ si et seulement si $S_{K-1} = 1$ et $S_K = 0$

Donc K est le premier entier k tel que $S_k = 0$

10. $\Phi(x) = \mathbb{E}(\exp(xX_k)) = e^{-x} \mathbb{E}(\exp(x(X_k + 1)))$

Par la formule de transfert : $\mathbb{E}(\exp(x(X_k + 1))) = \sum_{j=0}^m e^{jx} \binom{m}{j} q^j p^{m-j} = (p + qe^x)^m$

$$\Phi(x) = e^{-x} (p + qe^x)^m$$

$$\Phi(x) \leq \Psi(x) \iff m \ln(p + qe^x) \leq mq(e^x - 1) \iff \ln(1 + q(e^x - 1)) \leq q(e^x - 1)$$

Or, $\forall u > -1$, $\ln(1 + u) \leq u$ donc en prenant $u = q(e^x - 1)$, on a bien l'inégalité demandée

11. On suppose $mq < 1$

a) $P(K = +\infty) = P(\forall k, S_k \geq 1) = P(\cap_{k \geq 1} (S_k \geq 1))$

D'où $1 \geq P(K < +\infty) \geq P(S_k = 0)$, $\forall k \geq 1$

Or $S_k = 1 - k + \sum_{i=1}^k (X_i + 1)$

$S_k - 1 + k = \sum_{i=1}^k (X_i + 1)$ somme de var i.i.d. admettant une espérance égale à mq et une variance, par la loi faible des grands nombres :

$$\forall \varepsilon > 0, \lim_{k \rightarrow \infty} P\left(\left|\frac{S_k - 1 + k}{k} - mq\right| \geq \varepsilon\right) = 0 \text{ ou } \lim_{k \rightarrow \infty} P\left(\left|\frac{S_k - 1 + k}{k} - mq\right| < \varepsilon\right) = 1$$

$$P\left(\left|\frac{S_k - 1 + k}{k} - mq\right| < \varepsilon\right) = P(-k\varepsilon < S_k - 1 + k - kmq < k\varepsilon) = P(1 - k(1 + \varepsilon - mq) < S_k < 1 - k(1 - \varepsilon - mq))$$

$1 - mq > 0$, choisissons $\varepsilon = 1 - mq > 0$

$$P(1 - k(1 + \varepsilon - mq) < S_k < 1 - k(1 - \varepsilon - mq)) = P(1 - 2k(1 - mq) < S_k < 1) \longrightarrow 1$$

D'où $\lim_{k \rightarrow \infty} P(S_k \geq 1) = 0$ et $\forall k$, $0 \leq P(K = +\infty) \leq P(S_k \geq 1)$

D'où $P(K = +\infty) = 0$

$$P(K < +\infty) = 1$$

b) $\Psi'(x) = (-1 + mqe^x)\Psi(x)$

$$\Psi \text{ est minimale en } x^* = -\ln(mq) > 0$$

$$\Psi(x^*) = e^{-\theta} \text{ avec } \theta = mq - 1 - \ln(mq)$$

$\theta > 0$ (étude de $t \mapsto t - 1 - \ln t$)

- c) $M_k - M_{k-1} = \exp(x^* S_k + \theta k) - \exp(x^* S_{k-1} + \theta k - \theta) = M_{k-1} [\exp(x^* (S_k - S_{k-1}) + \theta) - 1]$
 $\mathbb{E}(1_{(K \geq k-1)} (M_k - M_{k-1})) = \mathbb{E}(1_{(K \geq k-1)} M_{k-1} [\exp(x^* (S_k - S_{k-1}) + \theta) - 1])$
 M_{k-1} fonction de $X_1, \dots, X_{k-1}$
 $(K \geq k-1) = (S_1 > 0, \dots, S_{k-2} > 0)$ d'où $1_{(K \geq k-1)}$ fonction de $X_1, \dots, X_{k-2}$
D'où $1_{(K \geq k-1)} M_{k-1}$ fonction de $X_1, \dots, X_{k-1}$
 $\exp(x^* (S_k - S_{k-1}) + \theta) - 1$ fonction de X_k
D'où par indépendance :

$$\mathbb{E}(1_{(K \geq k-1)} (M_k - M_{k-1})) = \mathbb{E}(1_{(K \geq k-1)} M_{k-1}) \mathbb{E}(\exp(x^* (S_k - S_{k-1}) + \theta) - 1)$$

Pour les mêmes raisons, comme $(K = k-1) = (S_1 > 0, \dots, S_{k-2} > 0, S_{k-1} = 0)$

$$\mathbb{E}(1_{(K=k-1)} (M_k - M_{k-1})) = \mathbb{E}(1_{(K=k-1)} M_{k-1}) \mathbb{E}(\exp(x^* (S_k - S_{k-1}) + \theta) - 1)$$

- d) $\mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) = \mathbb{E}(1_{(K \geq k-1)} (M_k - M_{k-1})) - \mathbb{E}(1_{(K=k-1)} (M_k - M_{k-1}))$
 $\mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) = \mathbb{E}(\exp(x^* (S_k - S_{k-1}) + \theta) - 1) [\mathbb{E}(1_{(K \geq k-1)} M_{k-1}) - \mathbb{E}(1_{(K=k-1)} M_{k-1})]$
 $\mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) = \mathbb{E}(\exp(x^* (S_k - S_{k-1}) + \theta) - 1) \mathbb{E}(1_{(K \geq k)} M_{k-1})$
 $\mathbb{E}(1_{(K \geq k)} M_{k-1}) \geq 0$
 $\mathbb{E}(\exp(x^* (S_k - S_{k-1}) + \theta) - 1) = \Phi(x^*) e^\theta - 1 \leq \Psi(x^*) e^\theta - 1 = 0$

$$\mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) \leq 0$$

- e) $\sum_{k=1}^N \mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) = \sum_{k=1}^N \mathbb{E}(1_{(K \geq k)} M_k) - \sum_{k=1}^N \mathbb{E}(1_{(K \geq k)} M_{k-1})$
 $\sum_{k=1}^N \mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) = \sum_{k=1}^N \mathbb{E}(1_{(K \geq k)} M_k) - \sum_{k=0}^{N-1} \mathbb{E}(1_{(K \geq k+1)} M_k)$
 $\sum_{k=1}^N \mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) = \sum_{k=1}^N \mathbb{E}(1_{(K=k)} M_k) - \mathbb{E}(1_{(K \geq 1)} M_0) + \mathbb{E}(1_{(K \geq N+1)} M_N)$
 $\sum_{k=1}^N \mathbb{E}(1_{(K=k)} M_k) = \mathbb{E}(M_0) + \sum_{k=1}^N \mathbb{E}(1_{(K \geq k)} (M_k - M_{k-1})) - \mathbb{E}(1_{(K \geq N+1)} M_N) \leq \mathbb{E}(M_0)$
On admet que $\mathbb{E}(M_K) = \mathbb{E}[\sum_{k=1}^{\infty} 1_{(K=k)} M_k] = \sum_{k=1}^{\infty} \mathbb{E}[1_{(K=k)} M_k]$ (intervention espérance et sommation)

$$\mathbb{E}(M_K) \leq \mathbb{E}(M_0)$$

- f) $M_K 1_{(K \geq k)} = \sum_{i=k}^{\infty} 1_{(K=i)} M_i$
 $M_K 1_{(K \geq k)} = \sum_{i=k}^{\infty} 1_{(K=i)} \exp(x^* S_i + \theta i) = \sum_{i=k}^{\infty} 1_{(K=i)} \exp(\theta i)$ car si $K = i$, $S_i = 0$
 $M_K 1_{(K \geq k)} \geq \sum_{i=k}^{\infty} 1_{(K=i)} \exp(\theta k) = \exp(\theta k) 1_{(K \geq k)}$
 $\mathbb{E}(M_0) \geq \mathbb{E}(M_K) \geq \mathbb{E}(M_K 1_{(K \geq k)}) \geq \exp(\theta k) P(K \geq k)$
 $\mathbb{E}(M_0) = \mathbb{E}(\exp(x^*)) = \exp(x^*)$

$$P(K \geq k) \leq \exp(x^* - \theta k)$$

S_0 est maintenant une variable supérieure à j :

$$P(K < +\infty) = \sum_{k=j}^{+\infty} P(S_0 = k) P(K < +\infty / S_0 = k) \text{ avec } P(K < +\infty / S_0 = k) \leq \exp(-y^*(mq)k) \leq \exp(-y^*(mq)j)$$

$$P(K < +\infty) \leq \exp(-y^*(mq)j) \sum_{k=j}^{+\infty} P(S_0 = k) = \exp(-y^*(mq)j)$$

$$\boxed{P(K < +\infty) \leq e^{-y^*(mq)j}}$$

13. On suppose $mq > 1$

a) $S_k - S_0 + k = (X_1 + 1) + \dots + (X_k + 1) \hookrightarrow \mathcal{B}(km, q)$ comme somme de binomiales indépendantes

$$\mathbb{E}(\exp(x(S_k - S_0 + k))) = \sum_{i=0}^{km} \binom{km}{i} q^i p^{km-i} e^{xi} \quad (\text{formule de transfert})$$

$$\mathbb{E}(\exp(x(S_k - S_0 + k))) = (qe^x + p)^{km} = \exp(km \ln(qe^x + p))$$

$$\ln(qe^x + p) = \ln(1 + q(e^x - 1)) \leq q(e^x - 1) \text{ car } \ln(1 + u) \leq u \text{ pour } u > -1$$

$$\text{D'où } \mathbb{E}(\exp(x(S_k - S_0 + k))) \leq \exp(kmq(e^x - 1))$$

$$\boxed{\mathbb{E}(\exp(x(S_k - S_0 + k))) \leq \exp(mqk(e^x - 1))}$$

b) Inégalité de Markow : Si $Y \geq 0$, et $a > 0$, $P(Y \geq a) \leq \frac{\mathbb{E}(Y)}{a}$

On pose ici $Y = \exp(x(S_k - S_0 + k)) \geq 0$ et $a = \exp(2xmqk) > 0$

$$P(Y \geq a) = P(S_k - S_0 + k \geq 2mqk) \text{ car } x > 0$$

$$\mathbb{E}(Y) \leq \exp(mqk(e^x - 1))$$

$$\text{D'où } P(S_k - S_0 + k \geq 2mqk) \leq \frac{\exp(mqk(e^x - 1))}{\exp(2xmqk)}$$

$$\boxed{\text{Pour } x > 0, \quad P(S_k - S_0 + k \geq 2mqk) \leq \exp(mqk(-2x + e^x - 1))}$$

c) Soit $x = \ln 2$, $-2x + e^x - 1 = 1 - 2\ln 2 \leq -1/3$

$$\exp(mqk(-2x + e^x - 1)) \leq \exp\left(-\frac{mqk}{3}\right)$$

$$\boxed{P(S_k - S_0 + k \geq 2mqk) \leq \exp\left(-\frac{mqk}{3}\right)}$$

d) Par la formule de Taylor-Lagrange : $\forall x \leq 0, \quad e^x \leq 1 + x + \frac{x^2}{2}$

e) En reprenant la variable Y définie dans le b. et $a = \exp((1 - \alpha)xmqk)$

$$P(Y \geq a) = P(S_k - S_0 + k \leq (1 - \alpha)mqk) \text{ car } x < 0$$

$$\text{D'où } P(S_k - S_0 + k \leq (1 - \alpha)mqk) \leq \frac{\exp(mqk(e^x - 1))}{\exp((1 - \alpha)xmqk)}$$

$$\boxed{\text{Pour } x < 0, \quad P(S_k - S_0 + k \leq (1 - \alpha)mqk) \leq \exp(mqk(-(1 - \alpha)x + e^x - 1))}$$

Soit $x = -\alpha < 0$

$$-(1 - \alpha)x + e^x - 1 = e^{-\alpha} - (1 + (-\alpha) + (-\alpha)^2) \leq \frac{\alpha^2}{2}$$

$$\boxed{P(S_k - S_0 + k \leq (1 - \alpha)mqk) \leq \exp\left(-\frac{\alpha^2 mqk}{2}\right)}$$

IV. Taille du plus grand cluster lorsque $\lambda < 1$

14. On définit la variable X par :

Si $N = k$, alors $X = X_1 + \dots + X_k$

$X/N = k \hookrightarrow \mathcal{B}(k, q)$

$$\boxed{X \hookrightarrow \text{Bin}(N, q)}$$

15. a) $\star S_0 = A_0 = 1$

$\star$ On peut remarquer que la formule donnée pour $k < K_0$ est aussi valable pour $k = K_0$ car alors $A_k = 0$ et $A_{k-1} = 0$

Si $k \leq K_0$, $\sum_{i=1}^k (S_i - S_{i-1}) = S_k - 1 = \sum_{i=1}^k (A_i - A_{i-1}) + \sum_{i=1}^k \sum_{l=1}^{A_{i-1}+i-1} X_{i,l} = A_k - 1 + \sum_{i=1}^k \sum_{l=1}^{A_{i-1}+i-1} X_{i,l}$

D'où $S_k \geq A_k$

$\star$ Si $k \leq K_0$, $S_k - S_{k-1} + 1 = B_k + \sum_{l=1}^{A_{k-1}+k-1} X_{k,l}$

$B_k \hookrightarrow \text{Bin}(n - k + 1 - A_{k-1}, \frac{\lambda}{n})$

$\sum_{l=1}^{A_{k-1}+k-1} X_{k,l} \hookrightarrow \text{Bin}(k - 1 + A_{k-1}, \frac{\lambda}{n})$ (cf 14)

et par indépendance, $B_k + \sum_{l=1}^{A_{k-1}+k-1} X_{k,l} \hookrightarrow \mathcal{B}(n, \frac{\lambda}{n})$ (cf résultat admis)

$$\boxed{S_k - S_{k-1} + 1 \hookrightarrow \mathcal{B}(n, \frac{\lambda}{n})}$$

Les $Z_{i,j}$ étant indépendantes et les B_k étant construits à partir d'une partition (aléatoire) des $Z_{i,j}$, on obtient l'indépendance des $S_k - S_{k-1} + 1$ et donc des $S_k - S_{k-1} + 1$

$\star$ Si $k \leq K_0$, $S_k - A_k = S_{k-1} - A_{k-1} + \sum_{l=1}^{A_{k-1}+k-1} X_{k,l} \geq S_{k-1} - A_{k-1}$

Donc $(S_k - A_k)$ est croissante tant que $k \leq K_0$

Remarquons que la réponse à la deuxième question découle du résultat de la quatrième

b) On a déjà montré que $S_k \geq A_k > 0$

Or K est le premier entier k tel que $S_k = 0$, donc

$$\boxed{K \geq K_0}$$

16. a) Si on note $X_k = S_k - S_{k-1}$, on a les hypothèses du III avec $m = n$ et $q = \frac{\lambda}{n}$
 $mq = \lambda < 1$, on peut donc utiliser les résultats du 11 : $P(K \geq k) \leq \exp(x^* - \theta k)$

Notons $k = \left\lceil \frac{1+\varepsilon}{\theta} \ln n \right\rceil + 1$ où $[x]$ désigne la partie entière de x

$$P(K > \frac{1+\varepsilon}{\theta} \ln n) \leq P(K \geq k) \leq \exp(x^* - \theta k) \leq \exp(x^* - \theta \frac{1+\varepsilon}{\theta} \ln n)$$

$$P(K > \frac{1+\varepsilon}{\theta} \ln n) \leq \exp(x^* - (1+\varepsilon) \ln n) = \frac{\exp(x^*)}{n^{1+\varepsilon}} = \frac{1}{\lambda n^{1+\varepsilon}}$$

$$\text{Or } |C_1| = K_0 \leq K \text{ d'où } P\left(|C_1| > \frac{1+\varepsilon}{\theta} \ln n\right) = P\left(K_0 > \frac{1+\varepsilon}{\theta} \ln n\right) \leq P(K > \frac{1+\varepsilon}{\theta} \ln n) \leq P(K \geq k)$$

$$\boxed{P\left(|C_1| > \frac{1+\varepsilon}{\theta} \ln n\right) \leq \frac{1}{\lambda n^{1+\varepsilon}}}$$

b) Par rôle symétrique des clusters, $P\left(|C_i| > \frac{1+\varepsilon}{\theta} \ln n\right) = P\left(|C_1| > \frac{1+\varepsilon}{\theta} \ln n\right)$

Notons $E = \text{"}\exists i \in \{1, \dots, n\}, \text{ tel que } |C_i| > \frac{1+\varepsilon}{\theta} \ln n\text{"}$

$$P(E) = P\left(\cup_{i=1..n} \left[|C_i| > \frac{1+\varepsilon}{\theta} \ln n\right]\right) \leq \sum_{i=1}^n P\left(|C_i| > \frac{1+\varepsilon}{\theta} \ln n\right)$$

$$0 \leq P(E) \leq nP\left(|C_1| > \frac{1+\varepsilon}{\theta} \ln n\right) = \frac{1}{\lambda n^\varepsilon}$$

Par le thm des gendarmes :

$$\boxed{\lim_{n \rightarrow \infty} P\left(\text{"}\exists i \in \{1, \dots, n\}, \text{ tel que } |C_i| > \frac{1+\varepsilon}{\theta} \ln n\text{"}\right) = 0}$$

17. Si un seul individu est infecté au départ , le nombre d'individus infectés est la taille d'un cluster.

Dans une grande population, avec une probabilité proche de 1, chaque cluster a une taille négligeable (en $\ln n$) par rapport au nombre d'individus (n) si le nombre moyen de voisins de chaque site est strictement inférieur à 1.

En effet le nombre moyen de voisins est $(n-1)p \sim \lambda < 1$