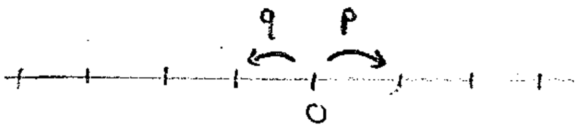
CHAP 2 — ETUDE DE 2 PROCESSUS SIMPLES : LA MARCHE ALEATOIRE EN DIMENSION 1 ET LE PROCESSUS DE GALTON-WATSON

I). Marche aléatoire en dimension 1 :

1). Présentation du problème :



$X_n = \sum_{i=1}^{n} \Delta_i \in \mathbb{Z}$ position de la particule au temps i

$X_0 = 0$ particule partant de l'origine

$(\Delta_i)_{i\in\mathbb{N}}$ suite des VAs indépendantes identiquement distribuées sur $\{-1,1\}$,
proba $(\Delta_i = +1) = p$, proba $(\Delta_i = -1) = q = 1-p$ (cf chap 1, II 1).)

On veut répondre aux questions suivantes :

1. Etant donné un site $a \in \mathbb{Z} - \{0\}$ fixé, la particule va-t-elle forcément visiter ce site en un temps fini ? Si oui, quelle est la loi de proba. du temps minimum au bout duquel la particule visite ce site ?
2. Quelle est la proba. que la particule retourne à l'origine (en un temps fini) ? Plus généralement, quelle est la loi de proba du temps minimum de retour à l'origine ?
3. La particule peut-elle s'éloigner arbitrairement loin (à gauche ou à droite) de l'origine ?

Pour répondre à la question 1, on définit le temps (discret) T_a de premier passage de la particule au site $a \in \mathbb{Z}$, $a \neq 0$,

$$\boxed{T_a = \inf \{ n \in \mathbb{N} \ ; \ X_n = a \}}$$

Remarquons que si le site a n'est jamais visité, alors $T_a = \infty$. Donc T_a est une VA à valeurs dans $\mathbb{N} \cup \{\infty\}$ (en fait $\mathbb{N}^* \cup \{\infty\}$)

Rappels sur les fonctions caractéristiques : $g_T(\eta) \stackrel{def}{=} \langle \eta^T \rangle = \sum_{t=0}^{\infty} \eta^t \,\text{proba}(T=t) + \overbrace{\eta^\infty}^{=0} \cancel{\text{proba}(T=\infty)}$ $\forall \eta \in]0,1[$

- $g_T(1^-) = \lim_{\eta \to 1^-} g_T(\eta) = \text{proba}(T<\infty) = 1 - \text{proba}(T=\infty)$
- .. $\text{proba}(T=t) = \frac{1}{t!} \left.\frac{d^t g_T}{d\eta^t}\right|_{0+}$ $\forall t \in \mathbb{N}$ (en effet, en dérivant la somme infinie termes à termes, $\frac{d^t g_T}{d\eta^t} = \sum_{s=t}^{\infty} s(s-1)\dots(s-t+1)\eta^{s-t} \,\text{proba}(T=s)$)
- ... $0 \leq g_T(\eta) \leq 1$ $\forall \eta \in]0,1[$
- .... $\left.\frac{d g_T}{d\eta}\right|_{1^-} = \sum_{s=1}^{\infty} s \,\text{proba}(T=s)$ $(\neq \langle T \rangle = \infty$ si $\text{proba}(T=\infty) > 0)$

2). Détermination de la fonction caractéristique de T_a :

On remarque que pour arriver au site $a+1$, si $a>0$, la particule doit tout d'abord passer par le site a. Donc $T_a < T_{a+1}$ et

$T_{a+1} = T_a + \inf\{m\in\mathbb{N}^* ;\ X_{T_a+m} - X_{T_a} = 1\}$ (*). Soit $\eta \in]0,1[$.

$$g_{T_{a+1}}(\eta) = \sum_{t=1}^{\infty} \eta^t p_{T_{a+1}}(t) = \sum_{\substack{s,t=0\\ s\geq t}}^{\infty} \eta^t p_{T_a,T_{a+1}}(s,t)$$ (cf chap 1, I 4) : $p_{T_{a+1}}(t)=\sum_{s=0}^{\infty} p_{T_a T_{a+1}}(s,t)$

proba du couple de VAs (T_a, T_{a+1})

$$= \sum_{s=1}^{\infty}\sum_{u=1}^{\infty} \eta^{s+u}\, p_{T_a,\, T_{a+1}-T_a}(s,u)$$ ← changement de variables $u=t-s\geq 0$

$$\underset{(*)}{=} \sum_{s=1}^{\infty}\sum_{u=1}^{\infty} \eta^{s+u}\ \text{proba}\left[\inf\{n\in\mathbb{N}^*;\ X_n=a\}=s \text{ et } \inf\{m\in\mathbb{N}^*;\ X_{s+m}-X_s=1\}=u\right]$$

or les VAs $X_n = \sum_{i=1}^{n}\Delta_i$ et $X_{s+m}-X_s=\sum_{j=s+1}^{s+m}\Delta_j$ sont indépendantes si $n\leq s$ (puisque les Δ_i sont indépendants). De plus $X_{s+m}-X_s$ a même loi de proba que $X_m=\sum_{j=1}^{m}\Delta_j$ (car les Δ_j sont identiquement distribués), donc les VAs $\inf\{m\in\mathbb{N}^*;\ X_{s+m}-X_m=1\}$ et $\inf\{m\in\mathbb{N}^*;\ X_m=1\}$ ont même loi

$$\Rightarrow g_{T_{a+1}}(\eta) = \sum_{s=1}^{\infty}\sum_{u=1}^{\infty}\eta^{s+u}\ \underbrace{\text{proba}\left[\inf\{n\in\mathbb{N}^*;\ X_n=a\}=s\right]}_{\equiv\, p_{T_a}(s)}\ \underbrace{\text{proba}\left[\inf\{m\in\mathbb{N}^*;\ X_m=1\}=u\right]}_{\equiv\, p_{T_1}(u)}$$

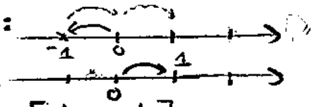
$$= \sum_{s=1}^{\infty}\eta^s p_{T_a}(s)\ \sum_{u=1}^{\infty}\eta^u p_{T_1}(u)$$

Par conséquent, l'on a $\boxed{g_{T_{a+1}}(\eta) = g_{T_a}(\eta)\, g_{T_1}(\eta)}$ $\forall \eta\in]0,1[$, $\forall a>0$. (1)

Nous allons maintenant dériver une autre relation entre $g_{T_2}(\eta)$ et $g_{T_1}(\eta)$. Pour cela, on distingue deux cas, suivant le 1ier saut effectué par la particule : soit ce saut est vers la gauche (c-à-d $\Delta_1=-1$), auquel cas l'on doit avoir $X_{T_1}-X_1=\sum_{j=2}^{T_1}\Delta_j = 2$ et $T_1\geq 2$; soit ce saut est vers la droite (c-à-d $\Delta_1=+1$), auquel cas l'on est sûr que $T_1=1$

$T_1\geq 2$:

$T_1=1$:



$$g_{T_1}(\eta) = \sum_{u=1}^{\infty}\eta^u p_{T_1}(u) = \sum_{u=1}^{\infty}\eta^u\left\{p_{\Delta_1,T_1}(-1,u) + p_{\Delta_1,T_1}(+1,u)\right\}$$

$$= \sum_{u=2}^{\infty}\eta^u\ \text{proba}\Big[\Delta_1=-1 \text{ et } \inf\{n\in\mathbb{N}^*;\ \underbrace{X_n-X_1}_{\text{même loi que } X_{n-1}}=2\}=u\Big] + \eta^1\ \text{proba}[\Delta_1=+1]$$

VAs indépendantes

$$= \sum_{u=2}^{\infty}\eta^u\ \text{proba}[\Delta_1=-1]\ \text{proba}\left[\inf\{n\in\mathbb{N}^*;\ X_{n-1}=2\}=u\right] + \eta\ \text{proba}[\Delta_1=1]$$

$$= q\sum_{u=2}^{\infty}\eta^u\ \underbrace{\text{proba}\left[\inf\{m\in\mathbb{N}^*;\ X_m=2\}=u-1\right]}_{\equiv\, p_{T_2}(u-1)} + \eta p$$

on pose $m=n-1$
$\inf\{n\in\mathbb{N}^*,\dots\} = \inf\{m\in\mathbb{N}^*,\dots\}+1$

$$= q\sum_{s=1}^{\infty}\eta^{s+1}p_{T_2}(s) + \eta p = q\, g_{T_2}(\eta)$$

D'où $\boxed{g_{T_1}(\eta) = q\eta\, g_{T_2}(\eta) + \eta p}$ $\forall \eta\in]0,1[$ (2)

En remplaçant (1) dans (2) l'on trouve l'équation du 2nd degré $q\eta[g_{T_1}(\eta)]^2 - g_{T_1}(\eta) + \eta p = 0$

qui a pour solution $\boxed{g_{T_1}(\eta) = \dfrac{1-\sqrt{1-4pq\eta^2}}{2q\eta}}$ $\forall\eta\in]0,1[$ (l'autre racine $\to\infty$, impossible)

En remplaçant cette valeur dans (1) l'on obtient

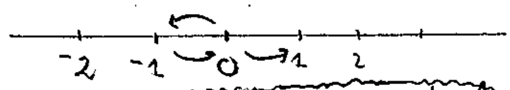
$$\boxed{g_{T_a}(\eta) = \left[g_{T_1}(\eta)\right]^a = \left[\frac{1-\sqrt{1-4pq\eta^2}}{2q\eta}\right]^a} \quad \forall \eta \in]0,1[, \ \forall a>0 \qquad (3a)$$

La fonction caractéristique de T_{-a} s'obtient à partir de $g_{T_a}(\eta)$ par symétrie : il suffit d'inverser le sens des coordonnées $X_n \to -X_n$, et donc d'échanger p et q

$$\boxed{g_{T_a}(\eta) = \left[g_{T_{-1}}(\eta)\right]^{|a|} = \left[\frac{1-\sqrt{1-4pq\eta^2}}{2p\eta}\right]^{|a|}} \quad \forall \eta \in]0,1[, \ \forall a<0 \qquad (3b)$$

Connaissant $g_{T_a}(\eta)$, l'on peut déterminer la loi de T_a et la proba que $T_a=\infty$. On trouve après calculs pour $a=1$

$$\boxed{p_{T_1}(t) = \begin{cases} p & \text{si } t=1 \\ \dfrac{1\times3\times\dots\times(t-2)}{2q}\,\dfrac{(2pq)^{\frac{t+1}{2}}}{\left(\frac{t+1}{2}\right)!} & \text{si } t \text{ est impair}, t\geq 3 \\ 0 & \text{si } t \text{ est pair} \end{cases}}$$



pour arriver au site $a=1$ la particule doit effectuer un nombre impair de sauts

et $p_{T_{-1}}(t)$ s'obtient en échangeant $p \leftrightarrow q$ dans cette formule.

La proba. que la particule arrive au site $a=1$ au bout d'un temps fini vaut

$$\text{proba}(T_1<\infty) = g_{T_1}(1-) = \frac{1-\sqrt{1-4pq}}{2q}$$

<u>or</u> $p+q=1 \Rightarrow 1-4pq = (p+q)^2-4pq = (p-q)^2 = (2p-1)^2 \Rightarrow 1-\sqrt{1-4pq} = \begin{cases} 2(1-p) & \text{si } 2p-1\geq 0 \\ 2p & \text{si } 2p-1\leq 0 \end{cases}$

$$\boxed{\text{proba}(T_1<\infty) = \begin{cases} 1 & \text{si } p\geq \frac{1}{2} \\ \frac{p}{q} & \text{si } p\leq\frac{1}{2} \end{cases}}$$

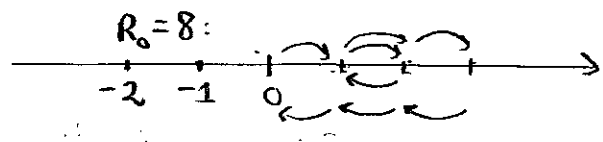
De même, $\boxed{\text{proba}(T_{-1}<\infty) = \begin{cases} \frac{q}{p} & \text{si } p\geq\frac{1}{2} \\ 1 & \text{si } p\leq\frac{1}{2} \end{cases}}$

<u>Conclusion:</u> En l'absence de "drift" ($p=q=\frac{1}{2}$) ou en présence d'un drift vers la droite ($p>q$), la particule visite le site $a=1$ (plus généralement le site $a>0$ fixé quelconque) au bout d'un temps fini avec probabilité 1, mais en présence d'un drift vers la gauche il y a une proba non nulle $\frac{q-p}{p}>0$ pour que le site ne soit jamais visité.

3). Temps de premier retour

On définit le temps de 1ier retour à l'origine par

$$\boxed{R_0 = \inf\{n\in\mathbb{N}^*;\ X_n=0\}}$$

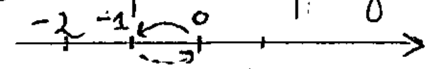


Comme pour T_a il s'agit d'une VA à valeurs dans $\mathbb{N}^*\cup\{\infty\}$

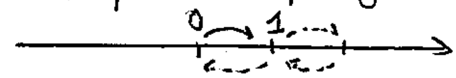
On va montrer que la fonction caractéristique $g_{R_0}(\eta)$ est simplement reliée à $g_{T_1}(\eta)$ et $g_{T_{-1}}(\eta)$.

Comme précédemment l'on distingue 2 cas, suivant le 1^{ier} saut de la particule :

→ si ce saut est vers la gauche ($\Delta_1=-1$), alors R_0 correspond au temps de 1^{ier} passage T_1 de la marche aléatoire $(\tilde{X}_m)_{m\in\mathbb{N}}$ partant de $\tilde{X}_0=0$, avec $\tilde{X}_m = X_{m+1}+1$



→ si ce saut est vers la droite ($\Delta_1=1$), alors R_0 correspond au temps de 1^{ier} passage T_{-1} de la marche aléatoire $(\tilde{\tilde{X}}_m)_{m\in\mathbb{N}}$ partant de $\tilde{\tilde{X}}_0=0$, avec $\tilde{\tilde{X}}_m = X_{m+1}-1$



$$g_{R_0}(\eta)=\sum_{n=1}^{\infty}\eta^n\left(p_{\Delta_1,R_0}(-1,n)+p_{\Delta_1,R_0}(1,n)\right)=\sum_{n=1}^{\infty}\eta^n\Big(\text{proba}\Big[\Delta_1=-1 \text{ et } \inf\{n\geq 2\,;\ \sum_{i=2}^{n}\Delta_i=1\}=n\Big]$$
$$+\ \text{proba}\Big[\Delta_1=+1 \text{ et } \inf\{n\geq 2\,;\ \sum_{i=2}^{n}\Delta_i=-1\}=n\Big]$$

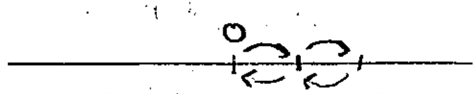
Comme $Z_1=\inf\{n\geq 2\,;\ \sum_{i=2}^{n}\Delta_i=\pm 1\}$ a même loi que $\inf\{n\geq 2\,;\ \sum_{i=1}^{n-1}\Delta_i=\pm 1\}=\inf\{m\geq 1\,;\ \sum_{i=1}^{m}\Delta_i=\pm 1\}+1=T_{\pm 1}+1$ et Z_1 est indépendante de Δ_1, on en déduit que

$$g_{R_0}(\eta)=\sum_{n=1}^{\infty}\eta^n\left\{q\,p_{T_1}(n-1)+p\,p_{T_{-1}}(n-1)\right\}$$
$$=\eta\left\{q\,g_{T_1}(\eta)+p\,g_{T_{-1}}(\eta)\right\}$$

$$\underset{(3)}{\Rightarrow}\quad \boxed{g_{R_0}(\eta)=1-\sqrt{1-4pq\eta^2}}\qquad \forall\,\eta\in\,]0,1[\qquad (4)$$

En dérivant cette fonction, on trouve

$$\boxed{p_{R_0}(n)=\begin{cases}2pq & \text{si } n=2\\ 1\times 3\times\ldots\times(n-3)\,\dfrac{(2pq)^{\frac{n}{2}}}{(\frac{n}{2})!} & \text{si } n \text{ est pair},\ n\geq 4\\ 0 & \text{sinon}\end{cases}}$$



pour retourner au site 0 la particule doit effectuer un nombre pair de sauts

De plus,

$$\boxed{\text{proba}(R_0<\infty)=g_{R_0}(1-)=1-\sqrt{1-4pq}=\begin{cases}2q & \text{si } p\geq\frac{1}{2}\\ 2p & \text{si } p\leq\frac{1}{2}\end{cases}}$$

(cf calcul de proba$(T_1<\infty)$ au 2).)

Conclusion : En l'absence de drift la particule retourne à l'origine (en un temps fini) avec probabilité 1, mais le temps moyen de retour à l'origine est infini :

$$\langle R_0\rangle=\left.\frac{d g_{R_0}}{d\eta}\right|_{1-}=\left.\frac{4pq\eta}{\sqrt{1-4pq\eta^2}}\right|_{\eta=1-}=\infty$$

exemple de VA finie qui a une moyenne infinie, c-à-d $\sum_{n=1}^{\infty}p_n=1$ mais $\langle R_0\rangle=\sum_{n=1}^{\infty}n p_n=\infty$

En présence d'un drift vers la droite ($p>\frac{1}{2}$), la particule retourne à l'origine avec une probabilité $2q=2(1-p)\geq 0$ et ne revient jamais au site 0 (et donc ne visite jamais les sites $x\leq 0$) avec probabilité $1-2q=2p-1>0$ [de même en présence d'un drift vers la gauche ($p<\frac{1}{2}$), on échange $p\leftrightarrow q$ dans les formules précédentes]

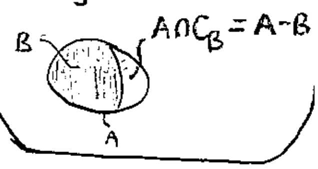
4). Elorgnement maximum de l'origine :

On définit l'éloignement maximum à droite (respectivement à gauche) de la particule par $\boxed{X_{max}=\sup_{n\in\mathbb{N}} X_n}$ (resp. $\boxed{X_{min}=\inf_{n\in\mathbb{N}} X_n}$). On a $X_{max}\in\mathbb{N}\cup\{\infty\}$ (resp. $X_{min}\in\{-\infty\}\cup\mathbb{Z}_-$)

Il est clair que si $a \geq 1$, $X_{max} \geq a$ si et seulement si $T_a < \infty$. Mais

$$p_{X_{max}}(a) = \text{proba}(X_{max} \geq a \text{ et } X_{max} < a+1) = \text{proba}(X_{max} \geq a) - \text{proba}(X_{max} \geq a+1)$$

car $\text{proba}(A-B) = \text{proba}(A) - \text{proba}(B)$ si $B \subset A$
c-à-d $\text{proba}(A) = \text{proba}(B) + \text{proba}(A \cap C_B)$ avec
ici $A = \{X_{max} \geq a\} \supset B = \{X_{max} \geq a+1\}$



D'où

$$\begin{aligned} p_{X_{max}}(a) &= \text{proba}(T_a < \infty) - \text{proba}(T_{a+1} < \infty) \\ &= g_{T_a}(1-) - g_{T_{a+1}}(1-) = g_{T_1}(1-)^a - g_{T_1}(1-)^{a+1} \\ &\underset{\substack{(\S 3).\\ \text{ci-dessus}}}{=} \begin{cases} 1^a - 1^{a+1} = 0 & \text{si } p \geq \frac{1}{2} \\ \left(\frac{p}{q}\right)^a \left(1 - \frac{p}{q}\right) & \text{si } p \leq \frac{1}{2} \end{cases} \end{aligned}$$

pour tout $a \geq 1$. Cette formule reste vraie pour $a=0$ (dans ce cas $p_{X_{max}}(0) = \text{proba}(T_1 = \infty)$)

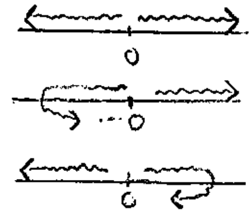
On a donc
$$\text{proba}(X_{max} < \infty) = \sum_{a=0}^{\infty} p_{X_{max}}(a) = \begin{cases} 0 & \text{si } p \geq \frac{1}{2} \\ \left(1 - \frac{p}{q}\right) \sum_{a=0}^{\infty} \left(\frac{p}{q}\right)^a = 1 & \text{si } p < \frac{1}{2} \end{cases}$$

La loi de probabilité de X_{min} s'obtient à partir de celle de X_{max} en échangeant $p \leftrightarrow q$: si $a \leq 0$

$$p_{X_{min}}(a) = \begin{cases} \left(\frac{q}{p}\right)^{|a|} \left(1 - \frac{q}{p}\right) & \text{si } p \geq \frac{1}{2} \\ 0 & \text{si } p \leq \frac{1}{2} \end{cases}$$

et donc
$$\text{proba}(X_{min} > -\infty) = \sum_{a=-\infty}^{0} p_{X_{min}}(a) = \begin{cases} 1 & \text{si } p > \frac{1}{2} \\ 0 & \text{si } p \leq \frac{1}{2} \end{cases}$$

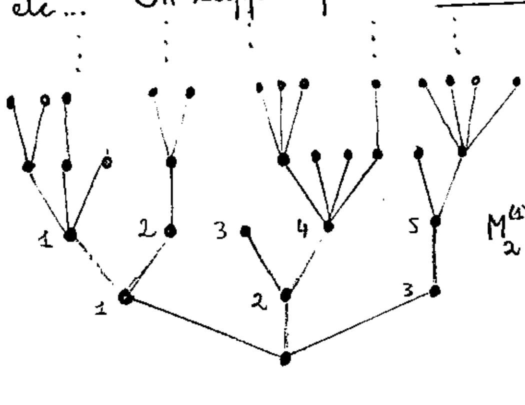
Conclusion :
- En absence de drift ($p = \frac{1}{2}$), $X_{min} = -\infty$ et $X_{max} = +\infty$ avec proba 1
- En présence d'un drift vers la droite ($p > \frac{1}{2}$), $-\infty < X_{min}$ et $X_{max} = \infty$
- En présence d'un drift vers la gauche ($p < \frac{1}{2}$), $-\infty = X_{min}$ et $X_{max} < \infty$



II). Processus de Galton-Watson

1). Le modèle d'évolution de la taille d'une population :

En 1873, les mathématiciens anglais Galton et Watson introduisent le modèle stochastique suivant pour étudier l'évolution du nombre d'individus portant le même nom appartenant à une même famille de l'aristocratie britanique. Soit $M_1^{(1)}$ le nombre de fils de l'ancêtre commun (on ne compte pas les filles qui perdent leur nom lors de leur mariage à cette époque). De même $M_2^{(1)}, \dots, M_2^{(M_1^{(1)})}$ sont les nombres de fils de chaque fils de l'ancêtre commun (2nde génération), etc... On suppose que les $M_n^{(i)}$ sont des VAs indépendantes de même loi à valeur dans $\mathbb{N}$.



$M_2^{(1)} = 2$, $M_2^{(2)} = 2$, $M_2^{(3)} = 1$

$M_1^{(1)} = 3$

$M_0^{(1)} = 1$

$Z_4 = 13$

$Z_3 = 10$

$Z_2 = 5$

$Z_1 = 3$

$Z_0 = 1$

Autrement dit, on néglige les corrélations entre les nombres de fils des différents individus (par exc. celles d'origine génétique) ainsi que le fait que certains individus ont une proba de n'avoir aucun enfant plus faible que d'autres (par exc., en France, les individus entrant dans le clergé...).

On note Z_n le nombre de descendants (masculins) de l'ancêtre commun portant son nom à la génération n. Autrement dit,

$$\boxed{Z_0 = 1 \,, \quad Z_{n+1} = \sum_{i=1}^{Z_n} M_n^{(i)} \,, \quad n \in \mathbb{N}}$$

nb de fils de l'individu N° i de la génération n

Ainsi, $Z_1 = M_1^{(1)}$, $Z_2 = M_2^{(1)} + \dots + M_2^{(Z_1)}$, etc... (cf figure ci-dessus)

$(Z_n)_{n\in\mathbb{N}}$ est un processus à temps discret à valeurs dans les entiers naturels $\mathbb{N}$; il s'agit de l'exemple le plus simple d'une famille de processus appelée processus de branchement, utilisés en particulier pour décrire l'évolution de la taille d'une population en biologie.

Le but de Galton et Watson en introduisant ce modèle était de calculer la probabilité d'extinction d'un nom de famille, c'est-à-dire, la probabilité que $Z_n = 0$ pour un n assez grand. Le temps d'extinction est la VA à valeur dans $\mathbb{N} \cup \{\infty\}$ défini par

$$\boxed{T = \inf\{n \in \mathbb{N} \,;\, Z_n = 0\}}$$

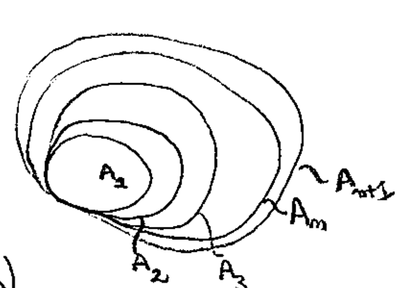
Remarquons qu'un nom de famille disparaît au bout d'un temps fini si et seulement si l'ancêtre a un nombre fini de descendants :

$T < \infty \iff Z = \sum_{n=0}^{\infty} Z_n < \infty$ [preuve mathématique : si $T<\infty$ alors $Z = \sum_{n=0}^{T} Z_n$ car $Z_n = 0 \Rightarrow Z_m = 0$ $\forall m \geq n$. Inversement, comme $Z_n \geq 1$ $\forall n \leq T$ on a $Z \geq T$ donc $T = \infty$ implique $Z = \infty$]

2). Calcul de la probabilité d'extinction

Considérons les évènements $A_n = \{Z_n = 0\}$, qui contiennent toutes les réalisations du processus telles que $T \leq n$. Comme $Z_n = 0$ implique $Z_{n+1} = Z_{n+2} = \dots = 0$, on a

$A_1 \subset A_2 \subset \dots \subset A_n \subset A_{n+1} \subset \dots$



Une telle suite d'évènements $(A_n)_{n\in\mathbb{N}^*}$ est dit croissante.

Un résultat classique des cours de proba de 1ière ou 2ième année est que pour une suite croissante d'évènements, proba$\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty}$ proba(A_n) (ce qui est intuitivement clair car les ens. A_n sont "emboîtés les uns dans les autres" et leur réunion est le plus petit ens. les contenant tous). Ainsi,

proba$(T<\infty)$ = proba$\left(\bigcup_{n=0}^{\infty}\{Z_n = 0\}\right) = \lim_{n\to\infty}$ proba$(Z_n = 0)$

Rappelons que proba$(Z_n = 0) = g_{Z_n}(0+)$ où g_{Z_n} est la fonction caractéristique de la VA Z_n.

Pour évaluer g_{Z_n} nous allons trouver une relation donnant $g_{Z_{n+1}}$ en fonction de g_{Z_n} et de la fonction caractéristique g des VAs $M_n^{(i)}$ [rappelons que ces VAs ont même loi ⇒ même fonction caract.]

$$g_{Z_{n+1}}(\eta) = \langle \eta^{Z_{n+1}} \rangle = \sum_{z_n, z_{n+1}=0}^{\infty} p_{Z_{n+1},Z_n}(z_{n+1}, z_n)\, \eta^{z_{n+1}}$$ (car $p_{Z_{n+1}}(z_{n+1}) = \sum_{z_n=0}^{\infty} p_{Z_{n+1},Z_n}(z_{n+1}, z_n)$)

$$= \sum_{z_n=0}^{\infty} \sum_{m_n^{(1)}=0}^{\infty} \sum_{m_n^{(2)}=0}^{\infty} \dots \sum_{m_n^{(z_n)}=0}^{\infty} p_{M_n^{(1)},\dots,M_n^{(z_n)},Z_n}(m_n^{(1)},\dots,m_n^{(z_n)}, z_n)\, \eta^{m_n^{(1)}+\dots+m_n^{(z_n)}}$$

car $z_{n+1} = \sum_{i=1}^{z_n} m_n^{(i)}$

VAs indépendantes car les $M_n^{(i)}$ sont indépendantes entre elles et indépendantes de Z_n qui est une fonction de $M_k^{(i)}$ avec $k \leq n-1$

Par indépendance des $M_k^{(i)}$ il vient

$$g_{Z_{n+1}}(\eta) = \sum_{z_n=0}^{\infty} \sum_{m_n^{(1)}=0}^{\infty} \cdots \sum_{m_n^{(z_n)}=0}^{\infty} p_{M_n^{(1)}}(m_n^{(1)}) \cdots p_{M_n^{(z_n)}}(m_n^{(z_n)})\, p_{Z_n}(z_n)\, \eta^{m_n^{(1)}+\dots+m_n^{(z_n)}}$$

$$= \sum_{z_n=0}^{\infty} p_{Z_n}(z_n) \left(\sum_{m_n^{(1)}=0}^{\infty} p_{M_n^{(1)}}(m_n^{(1)})\, \eta^{m_n^{(1)}} \right) \cdots \left(\sum_{m_n^{(z_n)}=0}^{\infty} p_{M_n^{(z_n)}}(m_n^{z_n})\, \eta^{m_n^{(z_n)}} \right)$$

Puisque les $M_n^{(i)}$ ont même loi et même fonction caractéristique $g(\eta)$, on obtient

$$g_{Z_{n+1}}(\eta) = \sum_{z_n=0}^{\infty} p_{Z_n}(z_n) \left[g(\eta) \right]^{z_n}$$

On reconnait l'expression de la fonction caractéristique de Z_n avec pour argument $g(\eta)$ au lieu de η :

$$g_{Z_{n+1}}(\eta) = g_{Z_n}(g(\eta)) \quad , \quad \forall n \in \mathbb{N}$$

Par itération, on trouve $g_{Z_n}(\eta) = g_{Z_{n-1}}(g(\eta)) = g_{Z_{n-2}}(g \circ g(\eta)) = \dots = g_{Z_1}(\underbrace{g \circ g \circ \dots \circ g}_{n-1 \text{ fois}}(\eta))$

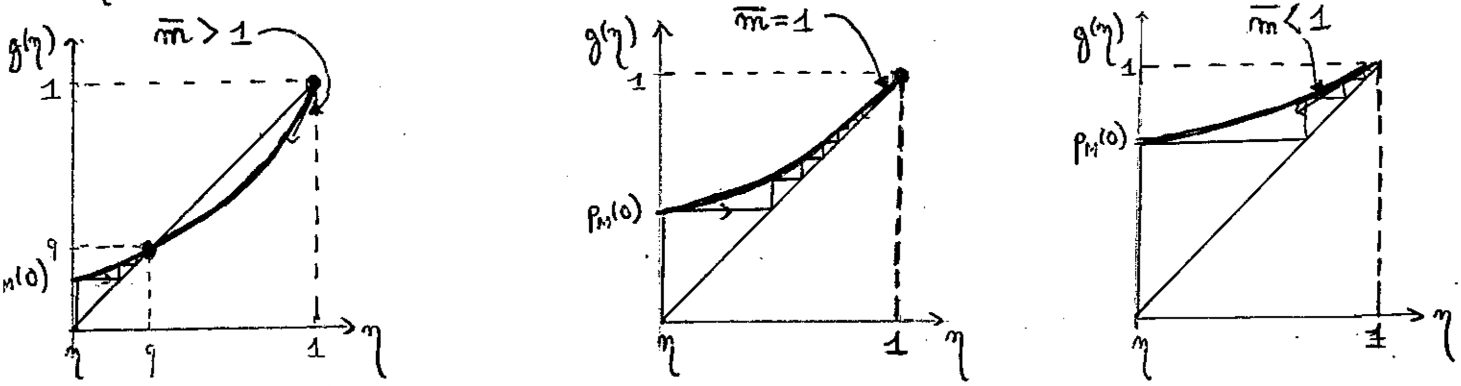
Mais $Z_1 = M_1^{(1)}$ et donc $g_{Z_1} = g$. D'où

$$\boxed{g_{Z_n}(\eta) = \underbrace{g \circ g \circ \dots \circ g}_{n \text{ fois}}(\eta) \equiv g^{(n)}(\eta)} \quad \forall \eta \in]0,1[$$

Ceci montre que $\text{proba}(T < \infty) = \lim_{n \to \infty} g^{(n)}(0+)$

On note $\bar{m} = \langle M_n^{(i)} \rangle = \frac{dg}{d\eta}(1-)$ le nombre moyen de fils de chaque individu

Rappelons qu'une fonction caractéristique g est croissante et convexe $]0,1[\to [0,1]$ (c-à-d, $\frac{dg}{d\eta} \geq 0$ et $\frac{d^2 g}{d\eta^2} \geq 0 \Rightarrow 0 \leq \frac{dg}{d\eta} \leq \frac{dg}{d\eta}(1-) = \bar{m}$)



On voit graphiquement que :

(i) si $\bar{m} > 1$ alors g a 2 points fixes sur $[0,1]$, $\eta = q \in]0,1[$ et $\eta = 1$ (càd l'équation $g(\eta) = \eta$ a 2 solutions $\eta = q$ et $\eta = 1$)

(ii) si $\bar{m} = 1$ alors
- ↗ si $p_M(0) > 0$ alors g a un seul point fixe $\eta = 1$
- ↘ si $p_M(0) = 0$ alors $g(\eta) = \eta \ \forall \eta \in]0,1[\Rightarrow M_n^{(i)} = 1$ avec proba 1

(iii) si $\bar{m} < 1$ alors g a un seul point fixe $\eta = 1$

On constate sur les graphiques que $g^{(n)}(\eta)$ converge vers le point fixe stable (c-à-d ayant une dérivée < 1) de g. C'est vrai d'après le théorème du point fixe (qui donne une manière de déterminer q numériquement en itérant la suite $(g^{(n)}(\eta))_{n \in \mathbb{N}}$). Donc, pour tout $\eta \in [0,1[$,

(i) si $\bar{m} > 1$ alors $g^{(n)}(\eta) \xrightarrow[n \to \infty]{} q$

(ii) si $\bar{m} \leq 1$ alors $g^{(n)}(\eta) \xrightarrow[n \to \infty]{} 1$ sauf si $M_n^{(i)} = 1$ avec proba 1 (dans ce cas $g(\eta) = \eta \Rightarrow g^{(n)}(\eta) \xrightarrow[n \to \infty]{} \eta$)

Conclusion : Si le nombre moyen de fils par individu est $\bar{m} \leqslant 1$ et si les individus n'ont pas systématiquement 1 fils, alors le nom s'éteint avec probabilité 1 :

proba$(T<\infty)$ = proba$(Z<\infty) = 1$

Sinon, la probabilité d'extinction vaut proba$(T<\infty) = q$ où q est le point fixe de g sur $]0,1[$. Si $M_n^{(i)} = 1$ avec proba 1 et, plus généralement, si la proba$(M_n^{(i)}=0)$ est nulle, alors proba$(T<\infty) = 0$ (ce qui n'est pas très surprenant !)

3). Loi de probabilité du nombre de descendants :

Le nombre moyen $\langle Z \rangle$ de descendants de l'ancêtre est simple à calculer.

En effet, l'on a $\langle Z_n \rangle = \frac{dg_{Z_n}}{d\eta}(1^-)$ (car $Z_n \in \mathbb{N}$) d'où, par différentiation d'une fonction composée,

$$\langle Z_{n+1}\rangle = \frac{d\, g_{Z_n}\circ g}{d\eta}(1) = \frac{dg_{Z_n}}{d\eta}\underbrace{(g(1))}_{=1}\times\underbrace{\frac{dg}{d\eta}(1)}_{=\bar{m}} = \bar{m}\,\frac{dg_{Z_n}}{d\eta}(1) = \bar{m}\,\langle Z_n\rangle$$

$$\Rightarrow \langle Z_n\rangle = \bar{m}\,\langle Z_{n-1}\rangle = \dots = \bar{m}^n\,\langle Z_0\rangle = \bar{m}^n$$

D'où

$$\langle Z\rangle = \sum_{n=0}^{\infty}\langle Z_n\rangle = \sum_{n=0}^{\infty}\bar{m}^n = \begin{cases}\dfrac{1}{1-\bar{m}} & \text{si } \bar{m}<1\\ \infty & \text{si } \bar{m}\geqslant 1\end{cases}$$

exemple de VA finie avec probabilité 1 mais ayant une moyenne infinie, comme au I).3).

Si l'on veut déterminer les moments d'ordre supérieur de Z ou des moyennes de fonctions quelconques de Z, on a besoin de calculer la loi de proba de Z ou, de manière équivalente, sa fonction caractéristique $g_Z(\eta)$.

Pour cela, notons que si $Z_n^{(i)}$ est le nombre de descendants de l'individu N° i de la $1^{\text{ère}}$ génération alors la VA $Z_n^{(i)}$ a même loi que Z_{n-1} car les $M_n^{(i)}$ ont même loi, pour tout $i=1,\dots z_1$ avec $Z_1 = z_1$. De plus $Z_n^{(i)}$ est indépendante de Z_1 et les $Z_n^{(i)}$, $i=1,\dots,z_1$ sont indépendantes entre elles. Or

$Z = 1 + \sum_{i=1}^{Z_1} Z^{(i)}$ avec $Z^{(i)} = \sum_{n=1}^{\infty} Z_n^{(i)}$ le nb total de descendant de l'individu N° i

$$\Rightarrow g_Z(\eta) = \sum_{z,z_1=0}^{\infty} P_{Z,Z_1}(z,z_1)\,\eta^{z} = \sum_{z_1=0}^{\infty}\ \sum_{z^{(1)},\dots,z^{(z_1)}=0}^{\infty} P_{Z^{(1)},\dots,Z^{(z_1)},Z_1}(z^{(1)},\dots,z^{(z_1)},z_1)\,\eta^{1+z^{(1)}+\dots+z^{(z_1)}}$$

$$= \sum_{z_1=0}^{\infty}\ \sum_{z^{(1)},\dots,z^{(z_1)}=0}^{\infty} P_{Z_1}(z_1)\,P_{Z^{(1)}}(z^{(1)})\cdots P_{Z^{(z_1)}}(z^{(z_1)})\,\eta^{1+z^{(1)}+\dots+z^{(z_1)}}$$

$$= \sum_{z_1=0}^{\infty} P_{Z_1}(z_1)\,\eta\, g_{Z^{(1)}}(\eta)\times\dots\times g_{Z^{(z_1)}}(\eta) = \eta\sum_{z_1=0}^{\infty} P_{Z_1}(z_1)\,[g_Z(\eta)]^{z_1}$$

car les VAs $Z^{(i)} = \sum_{n=1}^{\infty} Z_n^{(i)}$ ont même loi que $Z = \sum_{n=1}^{\infty} Z_{n-1}$ (même loi que $Z_n^{(i)}$). Mais Z_1 a pour fonction caractéristique g, d'où

$$\boxed{g_Z(\eta) = \eta\; g\circ g_Z(\eta)} \qquad \forall\, \eta \in\,]0,1[$$

On peut montrer que l'équation $s = \eta\, g(s)$ admet une unique solution dans $[0,1[$ pour tout $\eta \in [0,1[$ fixé. Cette solution ne peut être en général déterminée analytiquement, mais on peut la calculer numériquement (ou graphiquement) de manière approchée, elle nous donne alors $g_Z(\eta)$.

Dans le cas particulier où $M_n^{(i)}$ prend uniquement les valeurs 0 ou 2 avec probabilité $\frac{1}{2}$, l'on a $g(\eta) = \frac{1}{2}(1+\eta^2)$ et $\bar{m} = 1$

$\Rightarrow \quad g_Z(\eta) = \frac{1}{2}\eta\left(1 + g_Z(\eta)^2\right) \Rightarrow \eta[g_Z(\eta)]^2 - 2g_Z(\eta) + \eta = 0$

d'où $\quad g_Z(\eta) = \dfrac{1-\sqrt{1-\eta^2}}{\eta} \quad \forall \eta \in]0,1[$ (la solution $\dfrac{1+\sqrt{1-\eta^2}}{\eta} \underset{\eta \to 0}{\longrightarrow} \infty$ est à rejeter)

Il s'agit de la même fonction caractéristique que le temps de 1er passage d'une marche aléatoire symétrique en dim. 1 (cf I). 2).). Donc, dans ce cas particulier,

$$p_Z(z) = \begin{cases} \frac{1}{2} & \text{si } z = 1 \\ 1\times 3\times\ldots\times(z-2)\, \dfrac{(1/2)^{\frac{z+1}{2}}}{\left(\frac{z+1}{2}\right)!} & \text{si } z \text{ est impair, } z \geq 3 \\ 0 & \text{sinon} \end{cases}$$

<u>Lien avec les transitions de phase</u> (cf cours de Léonie)

Paramètre d'ordre $\quad \text{proba}(Z=\infty) = \begin{cases} 0 & \text{si } \bar{m} \leq 1 \\ 1-q & \text{si } \bar{m} > 1 \end{cases}$

$\langle Z \rangle$ diverge à la transition $\bar{m} = 1$