Pré-conférence : Un texte un mathématicien Darwin : le hasard et l'évolution

Manon Costa

18 janvier 2013

Qui est Darwin ?

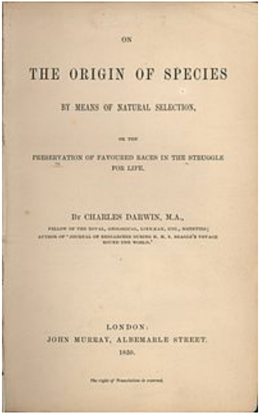
Charles Darwin (1809-1882) est un naturaliste anglais dont les travaux sur l'évolution des espèces vivantes ont révolutionné la biologie.

Voyage à bord du Beagle (1832-1836)

Lors de son voyage à bord du Beagle, il observe la diversité des espèces présentes et met au point sa théorie de l'évolution.

De l'origine des espèces (1859)

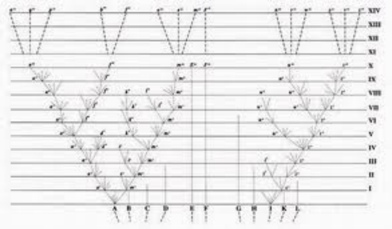
Sélection naturelle : processus permettant d'expliquer la création de différentes espèces partant d'un ancêtre commun.



Modéliser l'évolution ?

On peut distinguer deux types d'évolutions :

- L'évolution du nombre d'individus dans une population,
- L'évolution des caractéristiques des individus.



On va s'intéresser à des modèles mathématiques et plus précisément probabilistes pour comprendre ces évolutions.

Le modèle de Galton-Watson

Ce modèle décrit **l'évolution de la taille d'une population** qui se reproduit **aléatoirement**.

Chaque individu se reproduit indépendamment des autres suivant une loi de probabilité, c'est à dire :

- il a 0 descendant avec une probabilité p_0,
- il a 1 descendant avec une probabilité p_1,
- il a 2 descendants avec une probabilité p_2,
- ...
- il a q descendants avec une probabilité p_q,

On a bien sur $p_0 + p_1 + p_2 + ... + p_q = 1$.

Exemples

On considère des générations d'individus, c'est à dire qu'à chaque étape, les enfants remplacent les parents.
A la génération n, on note X_n la taille de la population (c'est une variable aléatoire).

- Un cas particulier : $\mathbf{p_1 = 1}$: chaque parent est remplacé par un descendant. La population reste **constante** : $X_n = X_0$

Exemples

On considère des générations d'individus, c'est à dire qu'à chaque étape, les enfants remplacent les parents.
A la génération n, on note X_n la taille de la population (c'est une variable aléatoire).

- Un cas particulier : $\mathbf{p_1 = 1}$: chaque parent est remplacé par un descendant. La population reste **constante** : $X_n = X_0$
- Chaque individu a deux enfants ($p_2 = 1$).
  Partant d'un individu, la taille de la population double à chaque génération : **à la génération** n **il y a** 2^n **individus**.

$$X_n = 2^n$$

Exemples aléatoires

On prend maintenant des exemples dans lesquels tous les individus n'ont pas le même nombre d'enfants.

- $\mathbf{p_0 = 1/3}$, $\mathbf{p_1 = 1/3}$, $\mathbf{p_2 = 1/3}$.
  Le nombre moyen d'enfants est donc : $0 \times 1/3 + 1 \times 1/3 + 2 \times 1/3 = 1$

Exemples aléatoires

On prend maintenant des exemples dans lesquels tous les individus n'ont pas le même nombre d'enfants.

- $\mathbf{p_0 = 1/3}$, $\mathbf{p_1 = 1/3}$, $\mathbf{p_2 = 1/3}$.
  Le nombre moyen d'enfants est donc : $0 \times 1/3 + 1 \times 1/3 + 2 \times 1/3 = 1$
- $\mathbf{p_0 = 1/5}$, $\mathbf{p_2 = 4/5}$.
  Le nombre moyen d'enfants est donc : $0 \times 1/5 + 2 \times 4/5 = 8/5 > 1$.
  On peut penser que la population va grandir.

Exemples aléatoires

On prend maintenant des exemples dans lesquels tous les individus n'ont pas le même nombre d'enfants.

- $\mathbf{p_0 = 1/3}$, $\mathbf{p_1 = 1/3}$, $\mathbf{p_2 = 1/3}$.
  Le nombre moyen d'enfants est donc : $0 \times 1/3 + 1 \times 1/3 + 2 \times 1/3 = 1$
- $\mathbf{p_0 = 1/5}$, $\mathbf{p_2 = 4/5}$.
  Le nombre moyen d'enfants est donc : $0 \times 1/5 + 2 \times 4/5 = 8/5 > 1$.
  On peut penser que la population va grandir.
- Si on regarde le cas inverse : $\mathbf{p_0 = 4/5}$, $\mathbf{p_2 = 1/5}$.
  Le nombre moyen d'enfants est donc : $0 \times 4/5 + 2 \times 1/5 = 2/5 < 1$.
  On peut penser que la population va s'éteindre.

Comment étudier ce processus ?

On introduit le polynôme :

$$f(x) = p_0 + p_1x + p_2x^2 + ... + p_qx^q$$

Comment étudier ce processus ?

On introduit le polynôme :

$$f(x) = p_0 + p_1 x + p_2 x^2 + ... + p_q x^q$$

<u>Exemples</u>

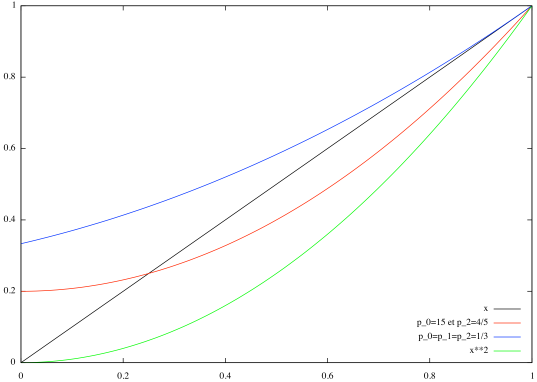
- Chaque individu a deux enfants $p_2 = 1$: alors $\mathbf{f(x) = x^2}$.
- Si $p_0 = 1/3, \quad p_1 = 1/3, \quad p_2 = 1/3$, alors

$$\mathbf{f(x) = \frac{1}{3} + \frac{1}{3}x + \frac{1}{3}x^2}$$

- Si $p_0 = 1/5, \quad p_2 = 4/5$, alors

$$\mathbf{f(x) = \frac{1}{5} + \frac{4}{5}x^2}$$

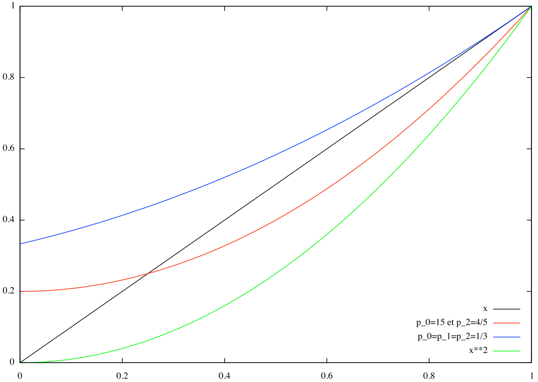
La fonction $f(x) = p_0 + p_1 x + p_2 x^2 + ... + p_q x^q$



On remarque que

- $f(0) = p_0$ et $f(1) = 1$,
- f est **croissante**
- (f est convexe)

La fonction $f(x) = p_0 + p_1 x + p_2 x^2 + ... + p_q x^q$



On remarque que

- $f(0) = p_0$ et $f(1) = 1$,
- f est **croissante**
- (f est convexe)

On peut facilement retrouver la moyenne :

$$f'(x) = p_1 + 2p_2 x + 3p_3 x^2 ... + qp_q x^{q-1}$$

alors,

$$\mathbf{f'(1) = p_1 + 2p_2 + 3p_3 ... + qp_q = m}.$$

Probabilité d'extinction à une génération

A la génération n on note $\mathbf{q_n}$ **la probabilité que la population soit éteinte à cette génération**.

$$q_n = \mathbb{P}(X_n = 0).$$

Par exemple : si on part d'un unique individu :

$$q_1 = \text{la probabilité de ne pas avoir d'enfants} = p_0.$$

Probabilité d'extinction à une génération

A la génération n on note $\mathbf{q_n}$ **la probabilité que la population soit éteinte à cette génération**.

$$q_n = \mathbb{P}(X_n = 0).$$

Par exemple : si on part d'un unique individu :

$$q_1 = \text{la probabilité de ne pas avoir d'enfants} = p_0.$$

La suite $(q_n)_n$ est croissante.

Et on peut montrer que c'est une **suite récurrente** :

$$q_n = f(q_{n-1})$$

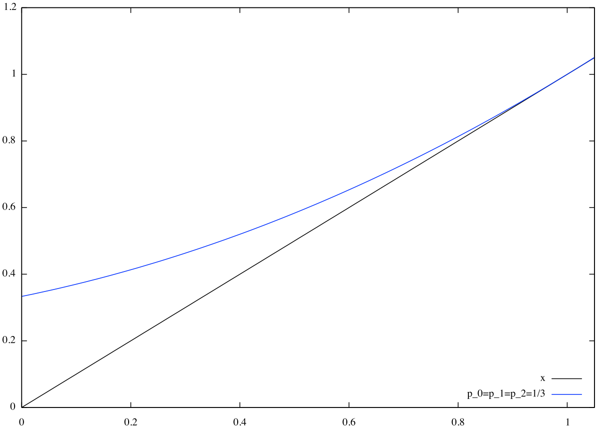
Limite de la probabilité d'extinction

On imagine que l'on regarde la population après un temps infini, **quelle est la probabilité qu'elle ait survécu ?**
C'est ce qu'on appelle la limite q de la suite $(q_n)_n$.

Limite de la probabilité d'extinction

On imagine que l'on regarde la population après un temps infini, **quelle est la probabilité qu'elle ait survécu ?**
C'est ce qu'on appelle la limite q de la suite $(q_n)_n$.

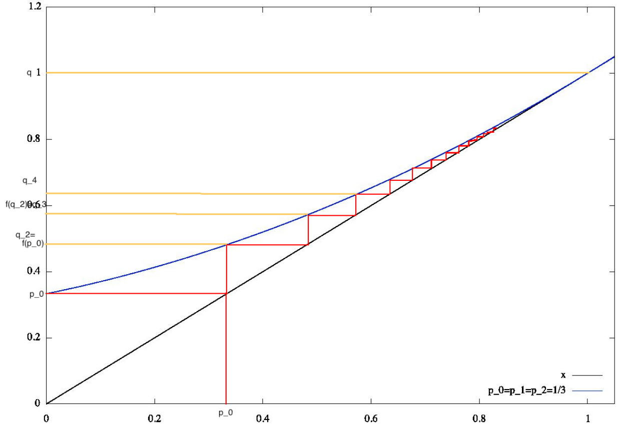
Pour calculer la limite, on utilise le graphe de f : car $q_n = f(q_{n-1})$



Limite de la probabilité d'extinction

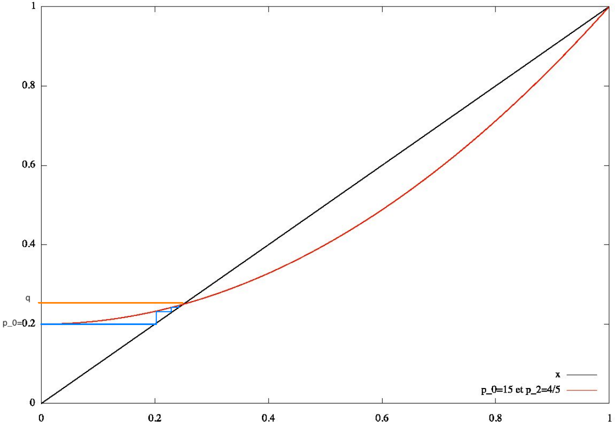
On imagine que l'on regarde la population après un temps infini, **quelle est la probabilité qu'elle ait survécu ?**
C'est ce qu'on appelle la limite q de la suite $(q_n)_n$.

Pour calculer la limite, on utilise le graphe de f : car $q_n = f(q_{n-1})$



Limite de la probabilité d'extinction

Un autre exemple : ici la moyenne est plus grande que 1 :

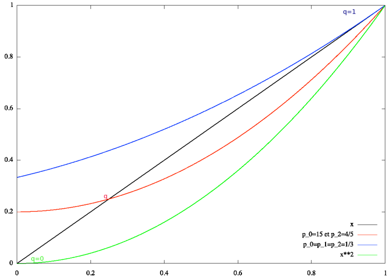


Conclusion

Si on a une limite q alors :

$$\mathbf{q} = \mathbf{f}(\mathbf{q})$$

.

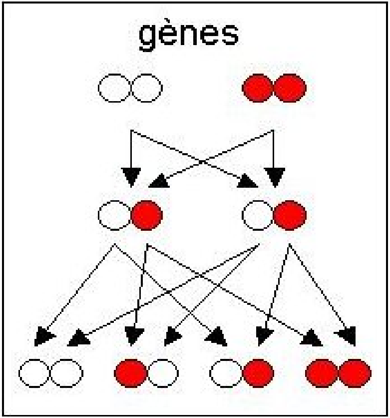


On a donc

- Si $p_0 = 0$, alors $q = 0$ et la population ne s'éteint jamais
- Si le nombre moyen d'enfants $m \leq 1$ q=1, donc la population s'éteint toujours
- Si le nombre moyen d'enfants $m > 1$, alors $0 < q < 1$ et la population ne s'éteint pas toujours

Comment ajouter la génétique ?

On veut donner un type à chaque descendant.
Le modèle le plus simple est de regarder **un gène avec deux allèles a et A**.



On a donc trois types d'individus

- les AA,
- les Aa,
- et les aa.

Passage à la seconde génération

À la première génération, on a les proportions suivantes :

- $\mathbf{p_1}$ individus AA
- $\mathbf{2q_1}$ individus Aa,
- $\mathbf{r_1}$ individus aa.

Quelles sont les proportions $\mathbf{p_2}, \mathbf{2q_2}, \mathbf{r_2}$ à la génération suivante ?

Passage à la seconde génération

À la première génération, on a les proportions suivantes :

- $\mathbf{p_1}$ individus AA
- $\mathbf{2q_1}$ individus Aa,
- $\mathbf{r_1}$ individus aa.

Quelles sont les proportions $\mathbf{p_2}, \mathbf{2q_2}, \mathbf{r_2}$ à la génération suivante ?

On raisonne pour chaque allèle, c'est à dire qu'on choisit les parents un par un ...

$$\begin{aligned}\mathbb{P}(\text{1er allèle A}) =& \mathbb{P}(\text{un parent AA}) + \frac{1}{2}\mathbb{P}(\text{un parent Aa}) + 0 \times \mathbb{P}(\text{un parent aa}) \\ =& p_1 + \frac{1}{2} 2q_1 + 0 \times r_1 \\ =& p_1 + q_1\end{aligned}$$

Passage à la seconde génération

À la première génération, on a les proportions suivantes :

- $\mathbf{p_1}$ individus AA
- $\mathbf{2q_1}$ individus Aa,
- $\mathbf{r_1}$ individus aa.

Quelles sont les proportions $\mathbf{p_2}, \mathbf{2q_2}, \mathbf{r_2}$ à la génération suivante ?

On raisonne pour chaque allèle, c'est à dire qu'on choisit les parents un par un ...

$$
\begin{aligned}
\mathbb{P}(\text{1er allèle A}) =& \mathbb{P}(\text{un parent AA}) + \frac{1}{2}\mathbb{P}(\text{un parent Aa}) + 0 \times \mathbb{P}(\text{un parent aa}) \\
=& p_1 + \frac{1}{2} 2q_1 + 0 \times r_1 \\
=& p_1 + q_1 \\
\mathbb{P}(\text{1er allèle a}) =& r_1 + q_1
\end{aligned}
$$

Passage à la seconde génération

Le choix des deux allèles revient au choix des deux parents qui se fait indépendamment :

$$\begin{aligned}\mathbb{P}(\text{être AA}) = p_2 =& \mathbb{P}(\text{1er allèle A et 2nd allèle A})\\ =& \mathbb{P}(\text{1er allèle A}) \times \mathbb{P}(\text{2nd allèle A})\\ \mathbf{p_2} =& (\mathbf{p_1} + \mathbf{q_1})^{\mathbf{2}}\end{aligned}$$

Passage à la seconde génération

Le choix des deux allèles revient au choix des deux parents qui se fait indépendamment :

$$\begin{aligned}\mathbb{P}(\text{être AA}) = p_2 =& \mathbb{P}(\text{1er allèle A et 2nd allèle A})\\ =& \mathbb{P}(\text{1er allèle A}) \times \mathbb{P}(\text{2nd allèle A})\\ \mathbf{p_2} =& (\mathbf{p_1} + \mathbf{q_1})^{\mathbf{2}}\end{aligned}$$

$$\begin{aligned}\mathbb{P}(\text{être Aa}) = 2q_2 =& \mathbb{P}(\text{1er allèle A et 2nd allèle a}) + \mathbb{P}(\text{1er allèle a et 2nd allèle A})\\ =& 2\mathbb{P}(\text{1er allèle A}) \times \mathbb{P}(\text{2nd allèle a})\\ \mathbf{2q_2} =& \mathbf{2}(\mathbf{p_1} + \mathbf{q_1})(\mathbf{q_1} + \mathbf{r_1})\end{aligned}$$

Passage à la seconde génération

Le choix des deux allèles revient au choix des deux parents qui se fait indépendamment :

$$\begin{aligned} \mathbb{P}(\text{être AA}) = p_2 &= \mathbb{P}(\text{1er allèle A et 2nd allèle A}) \\ &= \mathbb{P}(\text{1er allèle A}) \times \mathbb{P}(\text{2nd allèle A}) \\ \mathbf{p_2} &= (\mathbf{p_1} + \mathbf{q_1})^2 \end{aligned}$$

$$\begin{aligned} \mathbb{P}(\text{être Aa}) = 2q_2 &= \mathbb{P}(\text{1er allèle A et 2nd allèle a}) + \mathbb{P}(\text{1er allèle a et 2nd allèle A}) \\ &= 2\mathbb{P}(\text{1er allèle A}) \times \mathbb{P}(\text{2nd allèle a}) \\ \mathbf{2q_2} &= \mathbf{2}(\mathbf{p_1} + \mathbf{q_1})(\mathbf{q_1} + \mathbf{r_1}) \end{aligned}$$

$$\begin{aligned} \mathbb{P}(\text{être aa}) = r_2 &= \mathbb{P}(\text{1er allèle a et 2nd allèle a}) + \mathbb{P}(\text{1er allèle a et 2nd allèle A}) \\ \mathbf{r_2} &= (\mathbf{r_1} + \mathbf{q_1})^2 \end{aligned}$$

Loi de Hardy Weinberg

On a calculé les fréquences à la deuxième génération :

$$p_2 = (p_1 + q_1)^2$$

$$2q_2 = 2(p_1 + q_1)(r_1 + q_1)$$

$$r_2 = (r_1 + q_1)^2$$

On remarque que

$$q_2^2 = p_2 \times r_2$$

Conséquences de la Loi de Hardy Weinberg

Comme

$$q_2^2 = p_2 \times r_2,$$

Ceci va nous permettre d'en dire plus sur les fréquences des générations suivantes :

Conséquences de la Loi de Hardy Weinberg

Comme
$$q_2^2 = p_2 \times r_2,$$

Ceci va nous permettre d'en dire plus sur les fréquences des générations suivantes :
Pour les AA :
$$\begin{aligned} p_3 = (p_2 + q_2)^2 = & p_2^2 + 2p_2q_2 + q_2^2 \\ = & p_2^2 + 2p_2q_2 + \color{red}{p_2 \times r_2} \\ = & p_2 \times \underbrace{(p_2 + 2q_2 + r_2)}_{=1} \\ = & p_2 \end{aligned}$$

On conserve la mêmes fréquences pour les générations suivantes.

Ajout de la sélection natuelle ?

Ici tous les individus ont la même probabilité de se reproduire, **comment ajouter le fait que certains gènes peuvent améliorer la survie ?**

Merci de votre attention !