

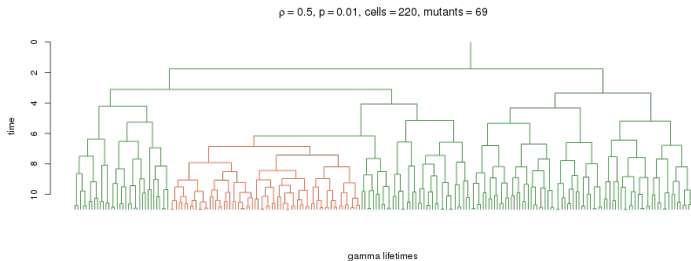
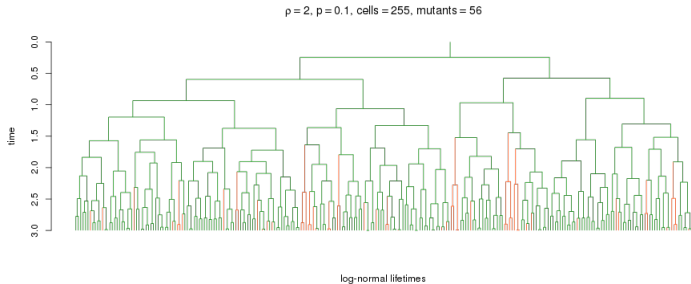
# Modèles de mutation : étude probabiliste et estimation paramétrique

MAZOYER Adrien  
encadrant de thèse : YCART Bernard (IPS)

4 décembre 2014

## Importance des modèles de mutations

- But principal : construire des estimateurs de probabilités individuelles de mutation ou d'autres paramètres.
- S'applique à des colonies de très grandes tailles, pour des taux de mutations très faibles.
- Utilisés dans de nombreux domaines de la médecine et de la biologie (cancer, microbiologie,...).



(source : <http://ljk.imag.fr/membres/Bernard.Ycart/>)

## 3 niveaux de constructions

- Apparition de mutations aléatoires au cours d'un processus de croissance.
- Développement de chaque clone issu de cellules mutantes.
- Nombre de cellules produites par un clone pendant une durée de développement donnée.

# Modèle classique de Luria-Delbrück ( $LD$ )

[1, Luria-Delbrück(1943)] ; [2, Lea-Coulson(1949)] ; [3, Hamon-Ycart(2012)]

## Hypothèses

- Population initiale homogène de  $n$  cellules normales.
- La durée de développement d'une cellule normale suit une loi exponentielle de paramètre  $\nu$ .
- La division d'une cellule normale donne :
  - Une cellule normale et une mutante avec probabilité  $\pi$ .
  - Deux cellules normales avec probabilité  $1 - \pi$ .
- La durée de développement d'une cellule mutante suit une loi exponentielle de paramètre  $\mu$ .
- La division d'une cellule mutante donne deux cellules mutantes.
- Les événements aléatoires (durée de développement et instants de mutation) sont mutuellement aléatoires.

En pratique : le nombre de divisions cellulaires est très grand et  $\pi$  est très faible  $\Rightarrow$  modélisation asymptotique.

## 3 niveaux de constructions

- Apparition de mutations aléatoires au cours d'un processus de croissance.  
 $\Rightarrow$  **Approximation Poissonnienne (Loi des Faibles Nombres)**
- Développement de chaque clone issu de cellules mutantes.  
 $\Rightarrow$  **Suite de  $\nu$  exponentielles indépendantes pour chaque clone**
- Nombre de cellules produites par un clone pendant une durée de développement donnée.  
 $\Rightarrow$  **Suite de  $\nu$  géométriques indépendantes**

## 3 niveaux de constructions

- Apparition de mutations aléatoires au cours d'un processus de croissance.

⇒ **Approximation Poissonnienne (Loi des Faibles Nombres)**

- Développement de chaque clone issu de cellules mutantes.

⇒ **Suite de  $\nu$  exponentielles indépendantes pour chaque clone**

- Nombre de cellules produites par un clone pendant une durée de développement donnée.

⇒ **Suite de  $\nu$  géométriques indépendantes**

⇒ **Composée poissonnienne d'un mélange exponentiel de lois géométriques**

## Cadre asymptotique

Soient  $t_n$  et  $\pi_n$  deux suites et  $\alpha > 0$  tels que :

$$\lim_{n \rightarrow \infty} \pi_n = 0, \quad \lim_{n \rightarrow \infty} t_n = +\infty, \quad \lim_{n \rightarrow \infty} \pi_n n e^{\nu t_n} = \alpha$$

## Convergence de la loi du nombre final de mutantes

Lorsque  $n \rightarrow \infty$ , le nombre de mutantes à l'instant  $t_n$  converge en loi vers une  $va$  dont la fonction génératrice est

$$g_{\alpha, \rho}(z) = \exp(\alpha(h_{\rho}(z) - 1))$$

$$\text{où } h_{\rho}(z) = \rho z \int_0^1 \frac{v^{\rho}}{1 - z + zv} dv = \int_0^{+\infty} \frac{ze^{-\mu s}}{1 - z + ze^{-\mu s}} \nu e^{-\nu s} ds$$

[3, Hamon-Ycart(2012)] [4, Athreya-Ney(1972)]



## Estimation paramétrique

- Loi du nombre final de mutantes explicite.
- 2 paramètres :  $\alpha$  et  $\rho$ .
- Estimation de  $\alpha$  et  $\rho$  selon différentes méthodes :
  - Maximum de vraisemblance (difficile car loi à queue lourde).
  - Estimation par fonction génératrice.
  - “Méthode  $p_0$ ” (s’appuie sur le fait que  $\mathbb{P}[0 \text{ mutantes}] = e^{-\alpha}$ ).

## Des hypothèses de modélisation non-réalistes

- Durées de développement exponentielles.
- Indépendance des durées de développement.
- Distribution identique des durées de développement.
- Absence de mort cellulaire.
- Nombre final de cellules constant.





⇒ **Estimations biaisées**

## Théorique

- Terminer l'étude probabiliste des sources de biais.
- Étude par simulation des méthodes d'estimation.
- Identification de l'estimation la plus fiable possible selon le modèle employé.

## Pratique

- Implémentation des méthodes d'estimations en un package R (collaboration avec Stéphane Desprésaux).

-  S.E. Luria and M. Delbrück.  
Mutations of bacteria from virus sensitivity to virus resistance.  
*Genetics*, 28(6):491–511, 1943.
-  D.E. Lea and C.A. Coulson.  
The distribution of the number of mutants in bacterial populations.  
*Journal of Genetics*, 49(3):264–285, 1949.
-  A. Hamon and B. Ycart.  
Statistics for the Luria-Delbrück distribution.  
*Elect. J. Statist.*, 6:1251–1272, 2012.
-  K.B. Athreya and P.E. Ney.  
*Branching processes*.  
Springer-Verlag, 1972.