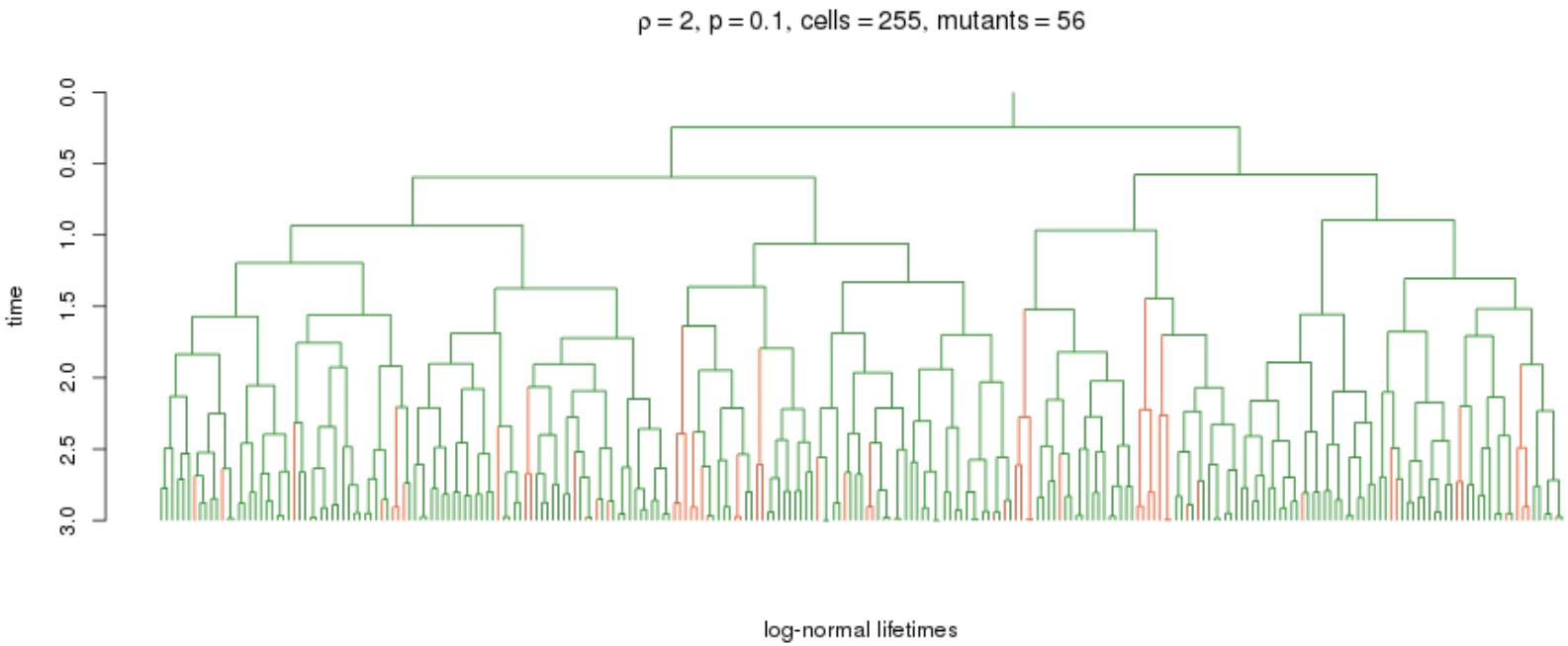
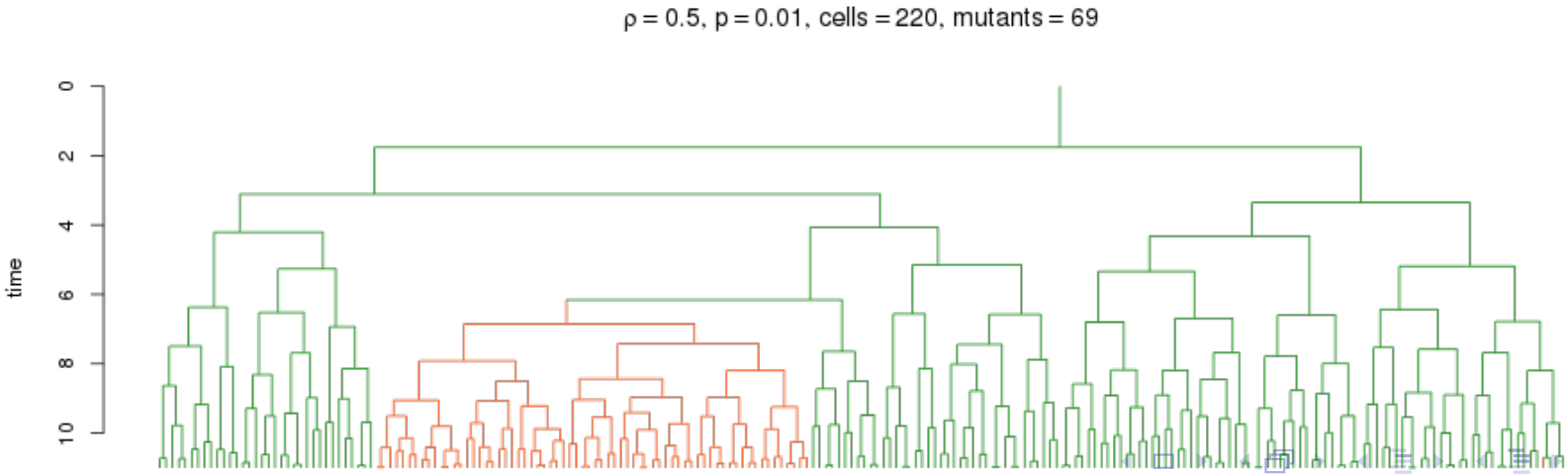
Mutation models: probabilistic study and parameter estimation

MAZOYER Adrien
Supervisor : YCART Bernard (SVH)

17th December 2014

Example


ρ = 2, p = 0.1, cells = 255, mutants = 56
time
0.0
0.5
1.0
1.5
2.0
2.5
3.0
log-normal lifetimes


ρ = 0.5, p = 0.01, cells = 220, mutants = 69
time
0
2
4
6
8
10

Motivations

N_{mut}	N_f
2	1.36e9
3	1.05e9
0	4.28e8
0	6.24e8
5	7.36e8
6	4.90e8
110	1.36e9
1	9.56e8
0	6.82e8

Parameters of interest:

$\rightarrow$ π : Probability of mutation

$\rightarrow$ α : Mean number of mutations

$\rightarrow$ ρ : "Fitness"

...

Classical model: Luria-Delbrück (LD)

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

Assumptions

- At time 0 a homogeneous culture of n normal cells.
- The generation time of any normal cell is exponentially distributed with parameter ν.
- A splitting normal cell is replaced by :
  - One normal and one mutant cell with probability π
  - Two normal cells with probability $1 - \pi$.
- The generation time of any mutant cell is exponentially distributed with parameter μ.
- A splitting mutant cell is replaced by two mutant cells.
- All random variables and events (division times and mutations) are mutually independent.

Probabilistic model of LD

$$\left.\begin{array}{l}\text{Number of divisions tends to } \infty \\ \text{Mutation probability tends to } 0\end{array}\right\} \Rightarrow \text{ Asymptotic model}$$

3 ingredients

- Occurrences of random mutations during a growth process.
- Growth of each clone starting from a single mutant cell.
- Number of cells in a mutant clone that develops for a finite time.

Probabilistic model of LD

$$\left.\begin{array}{l}\text{Number of divisions tends to } \infty \\ \text{Mutation probability tends to } 0\end{array}\right\} \Rightarrow \text{ Asymptotic model}$$

3 ingredients

- Occurrences of random mutations during a growth process.

  $\Rightarrow$ **Poisson distribution (law of small numbers)**
- Growth of each clones starting from a single mutant cell.

  $\Rightarrow$ **Sequence of independent exponential times for each clone.**
- Number of cells in a mutant clone that develops for a finite time.

  $\Rightarrow$ **Sequence of independent geometric numbers (Yule process).**

Results

Asymptotic assumptions

Let t_n et π_n two sequences and $\alpha > 0$ such that :

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

Initial result

As $n \to \infty$, the final number of mutants at time t_n, starting with n normal cells, converges to the distribution with probability generating function

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

where $h_\rho(z)$ is the probability generating function of the Yule distribution with parameter $\rho = \nu/\mu$.

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

Advantages - drawbacks

Advantages

- Explicit asymptotic distribution for the final number of mutants, with 2 parameters :
  - α : mean number of mutations ;
  - ρ : "fitness" parameter ($\rho = \nu/\mu$).
- Several estimation methods for α and ρ :
  - Maximum Likelihood estimators (but heavy tail distribution...) ;
  - Generating function estimators ;
  - p_0 estimators (relies on $\mathbb{P}[0 \text{ mutants}] = \mathrm{e}^{-\alpha}$).

Drawbacks: bias sources

- The developing times are i.i.d. exponentially distributed.
- Ignore cell deaths.
- No fluctuation of final number of cells.

"New" model

New assumptions

The generation time of a normal (resp. mutant) cell born at time $s \geq 0$ has cumulative distribution function

$$F_\nu(s,t) = 1 - \mathrm{e}^{-\nu(s,t)} \qquad \left(\text{resp. } F_\mu(s,t) = 1 - \mathrm{e}^{-\mu(s,t)}\right)$$

where ν (resp. μ), is positive, differentiable and increasing, such that

$$\lim_{t\to\infty} \nu(s,t) = +\infty \quad \text{and} \quad \forall\, t \in [0\,;\, s]\,,\ \nu(s,t) = 0\,.$$

Results

Main Result (Under similar asymptotic assumptions)

As $n \to \infty$, the final number of mutants at time t_n, starting with n normal cells, converges to the distribution with probability generating function

$$g(z) = \exp\left\{\alpha(\mathcal{I}(z, +\infty) - 1)\right\} ,$$

where $\mathcal{I}$ depends on ν and μ.

Results

Main Result (Under similar asymptotic assumptions)

As $n \to \infty$, the final number of mutants at time t_n, starting with n normal cells, converges to the distribution with probability generating function

$$g(z) = \exp\left\{\alpha(\mathcal{I}(z, +\infty) - 1)\right\},$$

where $\mathcal{I}$ depends on ν and μ.

Particular case : $\mu(s,t) = \frac{\nu(s,t)}{\rho}$

As $n \to \infty$, the final number of mutants at time t_n, starting with n normal cells, converges to the distribution with probability generating function

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

where $h_\rho(z)$ is the probability generating function of the Yule distribution with parameter ρ.

Current and future works

Theoretical work

- Studying non identical constant lifetimes case.
- Multitype branching process.
- Including cell deaths.

Current and future works

Theoretical work

- Studying non identical constant lifetimes case.
- Multitype branching process.
- Including cell deaths.

Practical work

Implementation of a R package including

- simulation ;
- estimation ;
- statistical tests.

Thank you for your attention !

Références


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 Berlin Heidelberg, 1972.